Adversarial Attacks to Distributed Voltage Control in Power Distribution Networks with DERs

Peizhong Ju, and Xiaojun Lin

School of Electrical and Computer Engineering, Purdue University.

Email: jup@purdue.edu, linx@ecn.purdue.edu

Abstract

It has been recently proposed that the reactive power injection of distributed energy resources (DERs) can be used to regulate the voltage across the power distribution network, and simple distributed control laws have been recently developed in the literature for performing such distributed Volt/VAR control. However, enabling the reactive-power injection capability of DERs also opens the door for potential adversarial attacks. Specifically, the adversary can compromise a subset of the DERs and use their reactive power to disrupt the voltage profile across the distribution network. In this paper, we study the potential damage (in terms of the voltage disruption) of such adversarial attacks and how to mitigate the damage by controlling the allowable range of reactive power injection at each bus. Somewhat surprisingly and contrary to the intuition that the reactive power injection at legitimate buses should help mitigating the voltage disruption inflicted by the adversary, we demonstrate that an intelligent attacker can actually exploit the response of the legitimate buses to amplify the damage by two times. Such a higher level of damage can be attained even when the adversary has no information about the network topology. We then formulate an optimization problem to limit the potential damage of such adversarial attacks. Our formulation sets the range of the reactive power injection on each bus so that the damage by the adversary is minimized, subject to the constraint that the voltage mismatch (without attack) can still be maintained within a given threshold under an uncertainty set of external inputs. Numerical results demonstrate the validity of our analysis and the effectiveness of our approach to mitigate the damage caused by such attacks.

I. Introduction

The increasing penetration of renewable generation from distributed energy resources (DERs) (such as solar panels) poses significant challenges to the reliable operation of the power system [10]. At the distribution level, reliability requirements dictate that the voltage across the distribution network must be maintained close to targeted levels [4]. However, the highly variable renewable generation causes uncertain disruption to the voltage profile. Traditional approaches of voltage control utilize utility-own devices, such as capacitor banks and on-load tap-changing (OLTC) transformers. These devices incur significant wear-and-tear cost and can only sustain a limited number of switching operation during their life span. As a result, these traditional approaches are no longer adequate when the voltage fluctuation caused by DERs are frequent and uncertain.

Due to this reason, there have been significant recent interests in utilizing the reactive-power (VAR) injection themselves to perform voltage control [29]. By formulating Volt/VAR control as an optimization problem, recent studies have led to the development of distributed control schemes [14]. Under such distributed Volt/VAR control schemes, each DER device only needs to measure the local voltage and accordingly adjust the amount of reactive power injection. Together, their reactive power injection will converge to an equilibrium value that solves the Volt/VAR optimization problem, without the need of any real-time communications. Therefore, they can be easily implemented in real systems.

However, the introduction of DER-based distributed VAR/Volt control also opens the door for potential adversarial attacks. Specifically, the adversary can compromise a subset of the DERs and use their reactive power to disrupt the voltage profile across the distribution network. Intuitively, if the compromised DER devices are allowed to inject large values of reactive power, they can potentially cause significant voltage disruption in the network. Therefore, there is a pressing need to understand the potential level of damage that can be inflicted by this type of adversarial attack, and how to mitigate them. To the best of our knowledge, this problem has not been studied in the literature.

In this paper, we take the first step towards answering this question. We first study how much damage (in terms of the voltage disruption) a given set of compromised DER devices can inflict on the voltage of a target bus. Note that since the DER devices on the legitimate buses (including the target bus) still follow the distributed control laws to compensate for the observed voltage mismatch, a naive thinking would be that the actions of the legitimate DER devices would lower the damage inflicted by the adversary. Somewhat surprisingly, our study shows the opposite. Specifically, we demonstrate that an intelligent attacker can actually exploit the response of the legitimate DER devices to amplify the damage by two times. Further, such a higher level of damage can be attained even when the adversary has no information about the network topology. Therefore, this type of more efficient attack is very simple to implement. We then solve the optimal attack strategy of the attacker, and we find in our numerical results that the damage inflicted by the above topology-agnostic strategy is often close to the optimal.

Based on this damage analysis, we then formulate an optimization problem to control the damage caused by the potential adversary, by appropriately setting the reactive power range of DERs on each bus. Note that the allowable range of the DERs' reactive power cannot be too small. Otherwise, even without attacks, the DER devices will lack sufficient reactive power to regulate the voltage under the presence of uncertain external inputs (e.g., the uncertain renewable generation). However, if these ranges are too large, a potential adversary will be able to inflict significant damage by compromising a few DER devices. The optimization problem that we formulate precisely capture this trade off. Specifically, given an uncertainty set $\mathcal{U}$ of external inputs and an upper bound M on the number of DER devices that the adversary can compromise, our formulation will minimize the maximum damage (i.e., the voltage disruption) that the adversary can inflict across a set of target buses, subject to the constraint that the voltage mismatch without attack can be maintained below a threshold for all possible external inputs within the uncertainty set $\mathcal{U}$. Since this formulation has the flavor of adjustable robust optimization (ARO) [6] and is often intractable, we use tools from affinely adjustable robust optimization (AARO) to develop a tractable solution.

Related Work

Our work is related to the large body of work for voltage control in power distribution networks. Traditional approaches of voltage control relay on capacitor banks and OLTC transformers (e.g., [1], [4]). Recently, utilizing the reactive power of the inverters on DERs is viewed as a new and important way of Volt/VAR control [12], [14], [27], [29]. Various DER-based Volt/VAR control schemes have been proposed, many of which represent distributed control schemes [10], [11], [14], [15], [19], [22], [23], [28], [29]. However, none of the above related works consider potential adversarial attacks that can utilize the reactive power injection capability of DERs to disrupt the voltage profile across the distribution network. To the best of our knowledge, our work is the first to study efficient attack strategies for such an adversary, as well as how to control the damage due to such adversarial attacks.

Our work is also related to the literature of cyber-physical system security. Many studies in this area focus on compromising the communication, sensing, and/or monitoring capability of the system [7], [16]–[18], [20], [21], [24]–[26]. For example, *false data injection attacks* are designed to inject false sensing data so that the state-estimator of the power system cannot obtain the true state of the system [16]. *Denial of service attacks* compromise the availability the communication channel, and *deception attacks* compromise the integrity of control packets or measurements [2]. The main goal of such attacks (as well as *replay attacks* [18], *covert attacks* [24], and *stealthy deception attacks* [7]) is to alter the sensing data or control packets while avoiding detection by the system [21]. Note that in centralized power systems, actuators (such as generators, capacitor banks, transformers) are often difficult to reach by an adversary. Thus, it makes sense to focus the studies of adversarial attacks on the communication, sensing and monitoring infrastructure. In contrast, in the adversarial attacks that we study in this paper, the target is not on the communication, sensing and monitoring part of the system. Instead, the attacker directly manipulates the physical signal (i.e., the reactive power injection) to induce large voltage disruptions. No direct communication or sensing is needed here because DER devices can inject reactive power distributively (and cheaply) and their effect on the voltage is through physical laws. Therefore, our study is very different from the above related work. We note that [17] also studies attacks to Volt/VAR control with DERs. However, it still focuses on compromising data integrity. In summary, both our study of the optimal attack strategy against distributed voltage control and our formulation of how to optimize the allowable range of reactive power injection under uncertain external inputs represent new contributions to this literature.

The rest of the paper is organized as follows. Section II introduces the system model for the power distribution network, the distributed Volt/VAR control scheme, and the adversary model. Section III analyzes intelligent attack strategies that the potential adversary can use and the maximum damage that it can inflict. Section IV formulates an optimization problem to control the potential damage under a given uncertainty set $\mathcal{U}$ of external inputs. Section V presents the simulation results to validate our analysis and our proposed approach for controlling the potential damage. Then, we conclude.

II. System Model

A. Power distribution network

Consider a radial power distribution network with $N+1$ buses forming a tree. We use the set $\mathcal{S} := \{0, 1, \cdots, N\}$ to denote the $(N+1)$ buses, with the bus 0 denoting the substation at the root of the tree. Let (i, j) denote the line connecting bus i to bus j, with the convention that bus i is closer to the substation than bus j (i.e., bus i is the parent of bus j on the tree). Let $\mathcal{L} \subset \mathcal{S} \times \mathcal{S}$ denote the set of lines. We define V_j, p_j and q_j as the voltage magnitude, the real power injection, and the reactive power injection, respectively, of bus j. The real and reactive power that flow from i to j on line (i, j) are denoted by P_{ij} and Q_{ij}, respectively. Let r_{ij} and x_{ij} denote the resistance and reactance, respectively, of the line (i, j). Let $\mathcal{S}_j \subset \mathcal{S}$ denote the set of bus j's neighboring buses that are further down from the substation. We use a linearized DistFlow model [29], which establishes a linear relationship between the power flow and bus voltage. This linearized model serves as an approximation when the voltage profile is relatively flat, i.e., V_i is not too far away from 1 for all buses $i \in \mathcal{S}$ when we have used unit voltage of 1 to denote the desired standard voltage. Specifically, assume that the voltage of the substation is always 1, i.e., $V_0 = 1$. The linearized DistFlow model states that for all $(i, j) \in \mathcal{L}$, we have

$$P_{ij} - \sum_{k \in \mathcal{S}_j} P_{jk} = -p_j, \tag{1}$$

$$Q_{ij} - \sum_{k \in \mathcal{S}_j} Q_{jk} = -q_j, \tag{2}$$

$$V_i - V_j = r_{ij} P_{ij} + x_{ij} Q_{ij}. \tag{3}$$

We can condense the above linearized DistFlow equations into a simpler matrix form [29]. Let matrix $\mathbf{M}^o = \left[M^o_{i,l}\right]$ of size $(N+1) \times N$ denote the graph incidence matrix for the tree topology. Its l-th column corresponds to the l-th line segment in $\mathcal{L}$ and its i-th row corresponds to bus i. Specifically, if bus $a(l)$ and bus $b(l)$ are two ends of the l-th line, and $a(l)$ is closer to the substation, we have $M^o_{a(l),l} = 1$ and $M^o_{b(l),l} = -1$. Further, $M^o_{i,l} = 0$ for $i \neq a(l), b(l)$. Let $\mathbf{m}_0$ be the first row of $\mathbf{M}^o$, which corresponds to the substation. Let $\mathbf{M}$ be the sub-matrix containing the rest of rows of $\mathbf{M}^o$. Then, we have $\mathbf{M}^o = \left[\begin{smallmatrix}\mathbf{m}_0 \\ \mathbf{M}\end{smallmatrix}\right]$. Because the tree topology is connected, the rank of $\mathbf{M}^o$ equals N [9], and thus $\mathbf{M}$ is invertible. Further, let $\mathbf{D}_r$ be an $N \times N$ diagonal resistance matrix with the l-th diagonal element equal to $r_{a(l),b(l)}$. Similarly, let $\mathbf{D}_x$ be an $N \times N$ diagonal reactance matrix with the l-th diagonal element equal to $x_{a(l),b(l)}$. Assume that $V_0 = 1$, i.e., the voltage of the substation bus always equals to 1. Using these matrices, the voltage equations (3) can be written as

$$\mathbf{m}_0^T + \mathbf{M}^T\mathbf{V} = (\mathbf{M}^o)^T \begin{bmatrix} V_0 \\ \mathbf{V} \end{bmatrix} = \mathbf{D}_r\mathbf{P} + \mathbf{D}_x\mathbf{Q}, \tag{4}$$

where $\mathbf{V}$, $\mathbf{P}$, and $\mathbf{Q}$ are $N \times 1$ vectors formed by v_j, P_{ij}, and Q_{ij} (for all $j \in \mathcal{S}\backslash\{0\}, (i,j) \in \mathcal{L}$) respectively. Similarly, the power balance equations (1) and (2) can be respectively represented by

$$-\mathbf{MP} = -\mathbf{p},$$

$$-\mathbf{MQ} = -\mathbf{q},$$

where $\mathbf{p}$ and $\mathbf{q}$ are $N \times 1$ vectors formed by p_j and q_j (for all $j \in \mathcal{S}\backslash\{0\}$), respectively. Solving for $\mathbf{P}$ and $\mathbf{Q}$ and substituting them in (4), we get

$$\mathbf{M}^T\mathbf{V} = \mathbf{D}_r\mathbf{M}^{-1}\mathbf{p} + \mathbf{D}_x\mathbf{M}^{-1}\mathbf{q} - \mathbf{m}_0^T,$$

or, equivalently,

$$\mathbf{V} = \mathbf{Rp} + \mathbf{Xq} - \mathbf{M}^{-T}\mathbf{m}_0^T, \tag{5}$$

where $\mathbf{R} := \mathbf{M}^{-T}\mathbf{D}_r\mathbf{M}^{-1}$, $\mathbf{X} := \mathbf{M}^{-T}\mathbf{D}_x\mathbf{M}^{-1}$. Note that each element $\mathbf{X}_{ij}$ of $\mathbf{X}$ represents the input on the voltage of bus i by one unit of reactive power injected on bus j. It is not difficult to see (from (1)-(3)) that $\mathbf{X}_{ij}$ is equal to the sum of the reactance on those links that are on both the path from the substation to bus i and the path from the substation to bus j. The elements of $\mathbf{R}$ can be interpreted in a similar manner. With the help of the next lemma, we have

$$\mathbf{V} = \mathbf{Rp} + \mathbf{Xq} + \mathbf{1}, \tag{6}$$

where $\mathbf{1}$ denote the $N \times 1$ vector with all 1's.

Lemma II.1. $-\mathbf{M}^{-T}\mathbf{m}_0^T = \mathbf{1}$.

Remark: One way to interpret Lemma II.1 is as follows. Suppose that there is no real or reactive power injection, i.e., $\mathbf{p} = \mathbf{q} = \mathbf{0}$. Then, all power flow must also be zero. Thus, the voltage of all buses should be equal to $V_0 = 1$. Therefore, we can let $\mathbf{p} = \mathbf{q} = \mathbf{0}$ in (5), apply the above property, and get the result $-\mathbf{M}^{-T}\mathbf{m}_0 = \mathbf{V} = \mathbf{1}$. A more mathematical proof is in Appendix.

In this paper, we are interested in scenarios with both normal load and DERs, both of which contribute to the injected real/reactive power on each bus. We use $\mathbf{p}^c$ and $\mathbf{q}^c$ to denote the consumed real and reactive power, respectively, of the normal load. Similarly, we use $\mathbf{p}^g$ and $\mathbf{q}^g$ denote the generated real and reactive power, respectively, of DERs. Then, the total injected real and reactive power are given by $\mathbf{p} = \mathbf{p}^g - \mathbf{p}^c$ and $\mathbf{q} = \mathbf{q}^g - \mathbf{q}^c$, respectively. Note that these vectors are all $N \times 1$ vectors. Further, in this paper, we are mainly interested in controlling $\mathbf{q}^g$, i.e., the reactive power of the DER, while the other variables $\mathbf{p}^c$, $\mathbf{q}^c$, and $\mathbf{p}^g$ are determined by external inputs and are uncontrollable. For ease of exposition, throughout the rest of the paper we make the simplifying assumption that every bus has only one DER device. Thus, in the rest of the paper we will refer to a DER device by the corresponding bus. We emphasize that there is no loss of generality due to this assumption because, if a bus has multiple DER devices, we can replace this bus as multiple buses (each one of them has one DER device) in an expanded tree topology. Then, we have

$$\begin{aligned}\mathbf{V} &= \mathbf{R}\mathbf{p} + \mathbf{X}\mathbf{q} + \mathbf{1}\\ &= \mathbf{R}\left(\mathbf{p}^g - \mathbf{p}^c\right) + \mathbf{X}\left(\mathbf{q}^g - \mathbf{q}^c\right) + \mathbf{1}\\ &= \mathbf{X}\mathbf{q}^g + \mathbf{1} + \mathbf{U},\end{aligned} \tag{7}$$

where $\mathbf{U} = \mathbf{R}\mathbf{p}^g - \mathbf{R}\mathbf{p}^c - \mathbf{X}\mathbf{q}^c$.

B. Volt/VAR control

In distribution systems, the objective of Volt/VAR control is to control the reactive power injection to minimize the voltage mismatch between the real-time voltage profile $\mathbf{V}$ and the standard voltage of $\mathbf{1}$. In the context of this paper, we are mainly interested in controlling the reactive power $\mathbf{q}^g$ of DER. (Note that in practice, reactive power can also be injected by other equipments, such as capacitor banks. However, capacitor banks incur wear-and-tear costs and their reactive power cannot be changed as frequently as that of DER. Hence, in this paper, we assume that the effect of capacitor banks is incorporated in the uncontrollable term $\mathbf{q}^c$.)

Define $\mathbf{q}^g(t)$ as the value of $\mathbf{q}^g$ at time slot t. Suppose that $\mathbf{q}^g$ is bounded in the range $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$. Define $\mathbf{V}(t)$ as the value of $\mathbf{V}$ at time slot t. We assume the following distributed VAR control scheme proposed in [29],

$$\mathbf{q}^g(t+1) = \left[\mathbf{q}^g(t) - \mathbf{D}(\mathbf{V}(t) - \mathbf{1})\right]^+, \tag{8}$$

where $[\cdot]^+$ denotes the projection when the value is outside of the range $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$, i.e., $[x]^+ = \arg\min_{y \in [\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]} \{||y - x||\}$. The matrix $\mathbf{D}$ is a $N \times N$ diagonal matrix with all positive diagonal elements, which correspond to step sizes. It has been shown in [29] that this VAR control scheme can be viewed as a gradient-projection method to solve the following optimization problem,

$$\min_{\mathbf{q}^g \in [\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]} \quad (\mathbf{X}\mathbf{q}^g + \mathbf{U})^T \mathbf{X}^{-1} (\mathbf{X}\mathbf{q}^g + \mathbf{U}). \tag{9}$$

Note that $\mathbf{X}$ is positive definite, which has been proved in [29]. As a special case, when every bus has abundant VAR resources (i.e., $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}] = (-\infty, \infty)$ and thus $\mathbf{q}^g$ is unbounded), the optimal solution to (9) is $\mathbf{q}^{g*} = -\mathbf{X}^{-1}\mathbf{U}$. Substituting $\mathbf{q}^{g*}$ into (7), one gets $\mathbf{V} = \mathbf{1}$, i.e., the iteration (8) converges to a solution that perfectly regulates all voltage to $\mathbf{1}$. If $\mathbf{q}^g$ is bounded, the convergent point may deviate from $\mathbf{1}$.

C. The adversary model

While distributed VAR control similar to (8) has been shown to regulate the voltage across the power distribution grid, it also opens the door for potential adversarial attacks. In particular, once an adversary compromises a number of DER devices, it can use them to inject reactive power in undesirable ways to disrupt the distribution-level voltage. As we will show in Section III, if the adversary knows that the legitimate DER devices will follow the control law (8), it can then leverage their (legitimate) response to further increase the voltage disruption. Thus, it is critical to understand (i) how adversary can launch such attacks effectively; (ii) how to limit the damage of such attacks.

In this paper, we take the first step to answer these questions. Note that if the reactive power injected by a compromised DER device is unbounded, than there is little we can do to limit the resulting voltage disruption. To make progress, in this paper we make the assumption that system designer can enforce the range $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$ of the reactive power injection $\mathbf{q}^g$ of each DER device. (Recall that we assume that there is only one DER device on a bus. Hence, this range also bounds the total DER reactive power on each bus.) In other words, even if a DER device has been compromised, it still cannot inject reactive power outside of this range. We expect that such enforcement may be achieved by building specialized circuits in the DER device that are more difficult to compromise.

Without loss of generality, we divide the set $\mathcal{S} \backslash \{0\}$ of all buses except the substation into two subsets $\mathcal{S}_b$ and $\mathcal{S}_a$. The subset $\mathcal{S}_b$ contains benign/legitimate buses where DER devices are not compromised. The subset $\mathcal{S}_a$ contains attacker buses where DER devices are compromised. We assume that the adversary can also compromise the smart meters on the compromised buses. Therefore, the system can not closely monitor the reactive power injected on these buses, which reflects most current power distribution systems where no additional monitoring capability exists. We now divide the vector $\mathbf{q}^g$ into two parts $\mathbf{q}_a^g$ and $\mathbf{q}_b^g$, which denote the sub-vectors of reactive power injected on the attacker buses and benign buses, respectively. We also divide the voltage vector $\mathbf{V}$ and $\mathbf{U}$ in the same way and divide the matrices $\mathbf{X}$ and $\mathbf{D}$ correspondingly. Specifically, we have

$$\mathbf{q}^g = \begin{bmatrix} \mathbf{q}_a^g \\ \mathbf{q}_b^g \end{bmatrix}, \mathbf{V} = \begin{bmatrix} \mathbf{V}_a \\ \mathbf{V}_b \end{bmatrix}, \mathbf{U} = \begin{bmatrix} \mathbf{U}_a \\ \mathbf{U}_b \end{bmatrix}, \mathbf{1} = \begin{bmatrix} \mathbf{1}_a \\ \mathbf{1}_b \end{bmatrix},$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{aa} & \mathbf{X}_{ab} \\ \mathbf{X}_{ba} & \mathbf{X}_{bb} \end{bmatrix}, \mathbf{D} = \begin{bmatrix} \mathbf{D}_a & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_b \end{bmatrix}. \tag{10}$$

Then, according to (7), the voltage of the benign buses is given by

$$\mathbf{V}_b = \mathbf{X}_{ba}\mathbf{q}_a^g + \mathbf{X}_{bb}\mathbf{q}_b^g + \mathbf{U}_b + \mathbf{1}_b. \tag{11}$$

The adversary can control the values of $\mathbf{q}_a^g(t)$ at each time slot t in arbitrary ways, as long as it is within the range $[\underline{\mathbf{q}_a^g}, \overline{\mathbf{q}_a^g}]$. On the other hand, $\mathbf{q}_b^g$ follows the distributed control law (8), which responds to the voltage seen at these buses, i.e.,

$$\mathbf{q}_b^g(t+1) = \left[\mathbf{q}_b^g(t) - \mathbf{D}_b(\mathbf{V}_b(t) - \mathbf{1}_b)\right]^+. \tag{12}$$

D. Objectives

Based on the above adversarial model, we aim to answer the following questions. First, given the reactive power range $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$ on each bus, the set of compromised buses $\mathcal{S}_a$, and a target bus c, we would like to find the sequence of reactive power injection $\mathbf{q}_a(t)$ by the adversary that can lead to the largest voltage disruption on the target bus c. We will focus on this first question in Section III.

Second, we wish to study how to choose the reactive power range $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$ so that the maximum damage is controlled when a given number of buses can be compromised. Specifically, suppose that there are at most M number of buses that may be compromised by the adversary, but we do not know which buses they are (since we assume each bus has one DER device, the value M is also the number of compromised DER devices). Note that if the reactive power range is too large, the potential adversary will be able to cause much voltage fluctuation. On the other hand, if the range is too small, even legitimate DER devices lack enough reactive power to regulate the voltage. Thus, the reactive power range needs to be carefully chosen. In Section IV, we will formulate an optimization problem to capture this trade-off.

III. Damage Analysis

In this section, we first focus on analyzing the maximum amount of damage (i.e., voltage disruption) that an adversary can inflict by controlling a given set of compromised DER devices (i.e., a given set of compromised buses).

A. Linearization and centering

In this section, we make the additional simplifying assumption that the reactive power injection is unbounded for the legitimate buses, i.e., $\mathbf{q}_b^g$ is unbounded. This assumption allows us to treat the system dynamics as a linear system, with the reactive power $\mathbf{q}_b^g$ at legitimate buses as state variables and the reactive power $\mathbf{q}_a^g(t)$ at malicious buses as inputs. (The system becomes linear because the nonlinear projection in (12) for the legitimate buses is removed.) Using (11) and (12), we have

$$\begin{aligned}\mathbf{q}_b^g(t+1) &= \mathbf{q}_b^g(t) - \mathbf{D}_b(\mathbf{V}_b(t) - \mathbf{1}_b)\\ &= \mathbf{q}_b^g(t) - \mathbf{D}_b\mathbf{X}_{ba}\mathbf{q}_a^g(t) - \mathbf{D}_b\mathbf{X}_{bb}\mathbf{q}_b^g - \mathbf{D}_b\mathbf{U}_b\\ &= (\mathbf{I} - \mathbf{D}_b\mathbf{X}_{bb})\mathbf{q}_b^g(t) - \mathbf{D}_b\mathbf{X}_{ba}\mathbf{q}_a^g(t) - \mathbf{D}_b\mathbf{U}_b, \end{aligned} \tag{13}$$

where $\mathbf{I}$ denotes the identity matrix. Next, for ease of exposition, we introduce the following centering procedure for $\mathbf{q}_b^g$ and $\mathbf{q}_a^g$. For $\mathbf{q}_a^g$ on the attacker buses, let $\mathbf{q}_a^{g,0}$ be the middle point in its range, i.e., $\mathbf{q}_a^{g,0} = (\overline{\mathbf{q}_a^g} + \underline{\mathbf{q}_a^g})/2$. For $\mathbf{q}_b^g$ on the legitimate buses, let $\mathbf{q}_b^{g,0}$ be the point such that

$$\mathbf{X}_{ba}\mathbf{q}_a^{g,0} + \mathbf{X}_{bb}\mathbf{q}_b^{g,0} = -\mathbf{U}_b. \tag{14}$$

Note that (14) implies that $\mathbf{q}_b^{g,0}$ is the fixed point for (13) when $\mathbf{q}_a^g(t) = \mathbf{q}_a^{g,0}$. Such a $\mathbf{q}_b^{g,0}$ always exists because $\mathbf{X}_{bb}$ can be shown to be invertible as in Proposition 1 of [29]. (Note that $\mathbf{X}_{bb}$ corresponds to the reactance matrix of a sub-network where the malicious buses are removed, and thus Proposition 1 of [29] also applies to $\mathbf{X}_{bb}$.) Now, let $\check{\mathbf{q}}_b^g(t) = \mathbf{q}_b^g(t) - \mathbf{q}_b^{g,0}$ and $\check{\mathbf{q}}_a^g(t) = \mathbf{q}_a^g(t) - \mathbf{q}_a^{g,0}$. Then, by (13), we have

$$\check{\mathbf{q}}_b^g(t+1) = (\mathbf{I} - \mathbf{D}_b\mathbf{X}_{bb})\check{\mathbf{q}}_b^g(t) - \mathbf{D}_b\mathbf{X}_{ba}\check{\mathbf{q}}_a^g(t). \tag{15}$$

Subtracting (14) from (11) and letting $\check{\mathbf{V}}_b(t) = \mathbf{V}_b(t) - \mathbf{1}_b$, we have

$$\check{\mathbf{V}}_b(t) = \mathbf{X}_{ba}\check{\mathbf{q}}_a^g(t) + \mathbf{X}_{bb}\check{\mathbf{q}}_b^g(t). \tag{16}$$

In this way, we obtain a typical linear model (15) and (16) with $\check{\mathbf{q}}_b$ as the state, $\check{\mathbf{q}}_a$ as the input, and $\check{\mathbf{V}}_b(t)$ as the output. Further, the bound on the range of $\check{\mathbf{q}}_a^g(t)$ is symmetric, i.e., $\check{\mathbf{q}}_a^g(t) \in [-\overline{\check{\mathbf{q}}}, \overline{\check{\mathbf{q}}}]$, where $\overline{\check{\mathbf{q}}} = (\overline{\mathbf{q}_a^g} - \underline{\mathbf{q}_a^g})/2$.

The first question that one may be concerned about is whether the adversary can make the system unstable, i.e., causing unbounded value of $\check{\mathbf{V}}_b(t)$. In linear control theory, a system is bounded input bounded output stable (BIBO stable) if bounded input can only lead to bounded output [3]. Since the system derived by (15) and (16) is a typical linear system, we can apply the condition of input-output stability in [3], and conclude that the system is BIBO stable if and only if the system without those malicious inputs is asymptotically stable. Here, by "without those malicious inputs," we mean that those malicious buses always have zero generation and load, but are still connected to the network (i.e., the topology remains the same). In [29], when all buses are legitimate, the authors have shown that the system must be asymptotically stable if

$$\lambda_{\max}\left(\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\right) < 2, \tag{17}$$

where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of a matrix. We wish to use this condition to verify both asymptotic stability and BIBO stability. However, note that when there is an adversary, the matrix $\mathbf{X}$ is changed to $\mathbf{X}_{bb}$. Thus, it appears that we need to verify the condition (17) for every possible set of malicious buses, which would have been quite computationally intensive. Fortunately, the following proposition shows that, as long as (17) holds for the original system without attack, it must also hold for the system under attack, regardless of which set of buses are malicious.

Proposition 1. *Regardless of the set of malicious buses, the system under attack (described by (15) and (16)) is guaranteed to be BIBO stable if $\lambda_{\max}(\mathbf{D}\mathbf{X}) < 2$, where $\mathbf{X}$ is given below (5) and is for the system when there is no attack. (Note that $\lambda_{\max}(\mathbf{D}\mathbf{X}) = \lambda_{\max}\left(\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\right)$, and hence the condition above is the same as (17).)*

Proof. See Appendix. □

B. A topology-agnostic attacker with twice the immediate damage

In this subsection, we present one example of an intelligent attack strategy (of choosing the reactive power injection $\mathbf{q}_a^g(t)$) that can leverage the response of legitimate nodes to increase the damage, i.e., the voltage disruption $\check{\mathbf{V}}(t)$. For simplicity, let us first focus on the setting with one attacker bus. In this case, the vector $\check{\mathbf{q}}_a^g(t)$ reduces to a scalar $\check{q}_a^g(t)$ in the interval $[-\overline{\check{q}}, \overline{\check{q}}]$. Assume that the initial state of the system is at the equilibrium, i.e., $\check{q}_a^g(0) = \mathbf{0}, \check{\mathbf{q}}_b^g(0) = \mathbf{0}$, and consequently, $\check{\mathbf{V}}_b(0) = \mathbf{0}$. If the adversary changes $\check{q}_a^g(1)$, while the reactive power injection $\check{\mathbf{q}}_b^g(1)$ at legitimate buses remains at 0, then the voltage disruption at the target bus c is simply $[\mathbf{X}_{ba}]_c \check{q}_a^g(1)$ (from (16)), where $\mathbf{X}_{ba}$ is a column vector corresponding to the one attacker bus a, and $[\cdot]_c$ denotes the row corresponding to the target bus c. Since $\check{q}_a^g(1)$ is within $[-\overline{\check{q}}, \overline{\check{q}}]$, the absolute value $|[\mathbf{X}_{ba}]_c \check{q}_a^g(1)|$ is at most $|[\mathbf{X}_{ba}]_c| \, \overline{\check{q}}$. Thus, we refer to this value as the *immediate damage* (from the attacker bus a to the target bus c). Of course, the values of $\check{\mathbf{q}}_b^g(t)$ will not remain at zero, but instead will follow (15). A naive thinking is that, since the change in $\check{\mathbf{q}}_b^g(t)$ tends to regulate the voltage closer to $\mathbf{1}$, the dynamics of the legitimate bus will make the ultimate damage smaller than the immediate damage. Our first result, described below, shows that this naive thinking is incorrect. Somewhat surprisingly, by leveraging the "corrective" dynamics of $\check{\mathbf{q}}_b^g(t)$, the attacker can actually inflict a damage that is twice the immediate damage. This increased damage can be attained even if the adversary knows nothing about the network topology.

More precisely, by (15) and (16), we can write down the relationship between the reactive power injection $\check{q}_a^g(t)$ of the attacker bus from time slot $t = 1$ to T and the damage to any bus $j \in \mathcal{S}_b$ at the time slot T as

$$\left[\check{\mathbf{V}}_b(T)\right]_j = \Bigg[-\sum_{t=1}^{T-1} \mathbf{X}_{bb}(\mathbf{I} - \mathbf{D}_b \mathbf{X}_{bb})^{T-t-1} \mathbf{D}_b \mathbf{X}_{ba} \check{q}_a^g(t) \\ + \mathbf{X}_{ba} \check{q}_a^g(T) \Bigg]_j, \text{ for all } j \in \mathcal{S}_b. \qquad (18)$$

Now, consider the following attack strategy. At time $t = 1, \cdots, T$, let $\check{q}_a^g(t) = \overline{\check{q}}$, i.e., the adversary always inject the largest possible reactive power. As one would expect, the dynamics (15) of the reactive power injection at legitimate buses will drive $\check{\mathbf{V}}_b(t)$ to zero. In other words, we have, from (18),

$$0 = \lim_{T\to\infty} \left[-\sum_{i=0}^{T-1} \mathbf{X}_{bb}(\mathbf{I} - \mathbf{D}_b\mathbf{X}_{bb})^i \mathbf{D}_b \mathbf{X}_{ba}\overline{\check{q}} + \mathbf{X}_{ba}\overline{\check{q}} \right]_j, \text{ for all } j \in \mathcal{S}_b$$

$$\implies \mathbf{X}_{ba} = \lim_{T\to\infty} \sum_{i=0}^{T-1} \mathbf{X}_{bb}(\mathbf{I} - \mathbf{D}_b\mathbf{X}_{bb})^i \mathbf{D}_b \mathbf{X}_{ba}. \tag{19}$$

Obviously, this strategy of keeping the same amount of reactive power injection is unwise for the attacker bus. Instead, the adversary can flip the sign of $\check{q}^g_a(t)$ at $t = T+1$. We then get,

$$\left[\check{\mathbf{V}}_b(T+1)\right]_c = \left[-\sum_{i=1}^{T} \mathbf{X}_{bb}(\mathbf{I} - \mathbf{D}\mathbf{X}_{bb})^i \mathbf{D}\mathbf{X}_{ba}\overline{\check{q}} + \mathbf{X}_{ba}(-\overline{\check{q}}) \right]_c.$$

Using (19), we then have

$$\lim_{T\to\infty} \left[\check{\mathbf{V}}_b(T+1)\right]_c = -2\left[\mathbf{X}_{ba}\right]_c \overline{\check{q}}. \tag{20}$$

In other words, when T is large, the damage at $T+1$ is almost twice the immediate damage. To summarize, the adversary can simply hold $\check{q}^g_a(t)$ at one extreme value (either $\overline{\check{q}}$ or $-\overline{\check{q}}$) for a sufficient long period of time T, wait for the legitimate buses to approach their equilibrium, and then flip the sign of $\check{q}^g_a(t)$ in the next time slot. In this way, the maximum voltage disruption is twice the immediate damage.

When there are multiple attacker buses, the same strategy can be applied for every attacker bus. Since the system is linear, the total damage simply adds up and can be written as $2\left[|\mathbf{X}_{ba}|\right]_c \overline{\check{\mathbf{q}}}$, where $\overline{\check{\mathbf{q}}}$ is now a vector, and $|\cdot|$ takes the absolute value for each element in a matrix. It is remarkable that this strategy does not need any network topology information, which is easy to be implemented by the adversary.

C. The optimal attack strategy

The above example naturally leads to the following question of whether a more intelligence attack strategy can introduce an even larger damage. In this section, we characterize the optimal attack strategy with the largest possible damage, under the assumption that the adversary knows the full network topology. For simplicity, we again first assume that there is only one attacker bus and the initial state of the system is at the equilibrium. Then, we can formulate the problem for the optimal attack strategy as

$$\max_{T, \check{\mathbf{q}}^g_a(t), t\in\{1,2,\cdots,T\}} \left[\left|\check{\mathbf{V}}_b(T)\right|\right]_c$$

$$\text{subject to} \quad (18) \text{ and } -\overline{\check{q}} \le \check{q}^g_a(t) \le \overline{\check{q}}, \text{ for all } t \in \{1,2,\cdots,T\}.$$

This problem has a standard solution [3]. Notice that the relationship (18) of $\check{\mathbf{V}}_b(t)$ and $\check{q}_a^g(t)$ is linear. Then, we can get, for a fixed T,

$$\begin{aligned}
&\left[\left|\check{\mathbf{V}}_b(T)\right|\right]_c \\
=&\left[\left|-\sum_{t=1}^{T-1}\mathbf{X}_{bb}(\mathbf{I}-\mathbf{D}_b\mathbf{X}_{bb})^{T-t-1}\mathbf{D}_b\mathbf{X}_{ba}\check{q}_a^g(t)+\mathbf{X}_{ba}\check{q}_a^g(T)\right|\right]_c \\
\leq&\sum_{t=1}^{T-1}\left[\left|-\mathbf{X}_{bb}(\mathbf{I}-\mathbf{D}_b\mathbf{X}_{bb})^{t-1}\mathbf{D}_b\mathbf{X}_{ba}\right|\right]_c\cdot\bar{\check{q}}+\left[\left|\mathbf{X}_{ba}\right|\right]_c\cdot\bar{\check{q}}, \qquad (21)
\end{aligned}$$

where equality can be achieved by setting

$$\check{q}_a^g(t)=\begin{cases}\bar{\check{q}} & \text{, if } \left[-\mathbf{X}_{bb}(\mathbf{I}-\mathbf{D}_b\mathbf{X}_{bb})^{T-t-1}\mathbf{D}_b\mathbf{X}_{ba}\right]_c\geq 0\\ -\bar{\check{q}} & \text{, otherwise,}\end{cases}$$

for time $t=1,\cdots,T-1$, and

$$\check{q}_a^g(T)=\begin{cases}\bar{\check{q}} & \text{, if } \left[\mathbf{X}_{ba}\right]_c\geq 0\\ -\bar{\check{q}} & \text{, otherwise,}\end{cases}$$

for time T. Since (21) is monotonically increasing in T, we get the maximum damage over T as

$$\begin{aligned}
&\max_{T,\check{q}_a^g(t),t\in\{1,2,\cdots,T\}}\left[\left|\check{\mathbf{V}}_b(t)\right|\right]_c \\
=&\lim_{T\to\infty}\left(\sum_{t=1}^{T-1}\left[\left|-\mathbf{X}_{bb}(\mathbf{I}-\mathbf{D}_b\mathbf{X}_{bb})^{t-1}\mathbf{D}_b\mathbf{X}_{ba}\right|\right]_c\bar{\check{q}}+\left[\left|\mathbf{X}_{ba}\right|\right]_c\bar{\check{q}}\right) \\
=&\beta\left[\left|\mathbf{X}_{ba}\right|\right]_c\bar{\check{q}},
\end{aligned}$$

where

$$\beta=\frac{\lim\limits_{T\to\infty}\left(\sum_{t=1}^{T-1}\left[\left|-\mathbf{X}_{bb}(\mathbf{I}-\mathbf{D}_b\mathbf{X}_{bb})^{t-1}\mathbf{D}_b\mathbf{X}_{ba}\right|\right]_c+\left[\left|\mathbf{X}_{ba}\right|\right]_c\right)}{\left[\left|\mathbf{X}_{ba}\right|\right]_c}. \qquad (22)$$

Since the system is BIBO stable, the limit in (22) always converges and β is always bounded.

Unfortunately, it seems difficult to get a closed-form expression for β. We have conducted numerical experiments based on various topologies. A representative numerical result in Section V shows that, when the step size $\mathbf{D}$ is reasonably small, the value of β is only slightly larger than 2. In other words, even if an attacker knows the full network topology, its damage is usually still very close to that of the simple strategy discussed in Section III-B (which does not require topology information at all). Thus, in the next section, we will assume that the attacker always uses the topology-agnostic strategy in Section III-B, in which case the damage is $2\left[|\mathbf{X}_{ba}|\right]_c \bar{\check{\mathbf{q}}}$. Recall that $\bar{\check{\mathbf{q}}} = (\overline{\mathbf{q}_a^g} - \underline{\mathbf{q}_a^g})/2$. Thus, going back to the system equations before the centering transformation of Section III-A, this maximum damage can simply be written as $\left[|\mathbf{X}_{ba}|\right]_c (\overline{\mathbf{q}_a^g} - \underline{\mathbf{q}_a^g})$.

Finally, we recall that the analysis of this section assumes that the reactive power injection of the legitimate buses is unbounded. In the next section, we will impose bounds on the reactive power of all buses. Although we do not know the optimal attack strategy in this more general setting, we expect that the maximum damage in the bounded setting will likely not be larger (which will be confirmed by the numerical results in Section V). Thus, we take the above expression of $\left[|\mathbf{X}_{ba}|\right]_c (\overline{\mathbf{q}_a^g} - \underline{\mathbf{q}_a^g})$ as an approximate and conservative (i.e., upper bound) estimate of the maximum damage inflicted by the adversary.

IV. Setting the Reactive Power Range

In this section, we study how to limit the damage of the potential adversary by setting the reactive power range of the DER devices. As we discussed at the end of Section II-D, the reactive power range $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$ needs to be carefully chosen. If the range is too large, when there is an attack, the adversary will be able to cause much voltage fluctuation. If the range is too small, even without attack the legitimate DER devices will lack enough reactive power to regulate the voltage under varying and uncertain external inputs. In this section, we will formulate an optimization problem to capture this trade-off. Naturally, this formulation needs to consider both scenarios (i.e., with and without attacks). Specifically, in the following, we aim to minimize the voltage disruption *under attack* subject to the constraint that, *when there is no attack*, the voltage can be regulated within a certain threshold under a given uncertainty set of external inputs (including real and reactive power consumption of normal load, and the real power injection of DER).

A. Without attack

Even without attack, the voltage for each bus should be within a certain range (e.g., 10% away from $V_0 = 1$). Let $\overline{\Delta V}$ denote the maximum allowable voltage mismatch when there is no attack. Then, by (7), we wish to control the reactive power injection $\mathbf{q}^g$ such that

$$-\overline{\Delta V} \leq \mathbf{X}\mathbf{q}^g + \mathbf{U} \leq \overline{\Delta V}, \tag{23}$$

where $\mathbf{U}$ is given below Eq. (7). However, in reality, the values of $\mathbf{p}^g$, $\mathbf{p}^c$, and $\mathbf{q}^c$ are varying and uncertain, and so is $\mathbf{U}$. Thus, we assume that we are given an uncertainty set $\mathcal{U}$ of possible values of $\mathbf{U}$. This uncertainty set then produces a constraint on $\underline{\mathbf{q}^g}$ and $\overline{\mathbf{q}^g}$, i.e.,

$$\text{for all } \mathbf{U} \in \mathcal{U}, \text{ there exists } \mathbf{q}^g \in [\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}] \text{ such that } (23) \text{ holds.} \tag{24}$$

In this paper, for simplicity we use an uncertainty set $\mathcal{U}$ of a box form, $\mathcal{U} = [\mathbf{U}^0 - \overline{\mathbf{U}}, \mathbf{U}^0 + \overline{\mathbf{U}}]$, where $\mathbf{U}^0$ and $\overline{\mathbf{U}}$ are two given $N \times 1$ vectors. (We note that our approach can also be applied to more complex forms of uncertainty sets.)

B. With attack

We assume that a limited number M of DER devices can be compromised at the same time, but the exact set of DER devices is unknown. Since we assume that each bus has one DER device, this translates into M attacker buses. As in Section II-C, denote the set of attacker buses by $\mathcal{S}_a$. Then, we have $\mathcal{S}_a \subset \mathcal{S}\backslash\{0\}$, and $|\mathcal{S}_a| = M$, where $|\mathcal{S}_a|$ denotes the number of elements in $\mathcal{S}_a$. Let Λ denote the set of target buses c that we wish to protect. Clearly, the malicious reactive power injection of the attacker will produce additional voltage mismatch on top of the voltage mismatch without attack (which is supposed to be within $\overline{\Delta V}$ as in (23)). In the following, we will reserve the term "voltage disruption" for this additional voltage mismatch due to the attacker, and we will use the term "voltage mismatch without attack" for the case when there is no attack. Our goal is thus to minimize the maximum voltage disruption across all buses in Λ. Recall from the end of Section III that we use the expression of the voltage disruption under the topology-agnostic scheme. Thus, the above goal can be written as minimizing δ subject to

$$\sum_{j \in \mathcal{S}_a} |\mathbf{X}_{cj}| \left(\overline{\mathbf{q}^g}_j - \underline{\mathbf{q}^g}_j \right) \le \delta, \text{ for all } \mathcal{S}_a \subset \mathcal{S}, |\mathcal{S}_a| = M, \text{ for all } c \in \Lambda, \tag{25}$$

where $\mathbf{X}_{cj}$ denotes the element of $\mathbf{X}$ at the c-th row, j-th column.

Putting these altogether, we arrive at the following formulation

$$\min_{\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}} \delta, \text{ subject to } (24) \text{ and } (25), \tag{26}$$

which is an instance of adjustable robust optimization problems [6]. Note that under this formulation, when there is an attack the worst-case voltage deviation/mismatch from $\mathbf{1}$ on any bus $c \in \Lambda$ is $(\delta + \overline{\Delta V})$, because δ is on top of the allowable voltage mismatch $\overline{\Delta V}$ without attack.

C. A tractable solution

Unfortunately, the above formulation can be intractable. First, the number of all possible $\mathcal{S}_a$ (and also the number of constraints in (25)) grows exponentially with M. Second, even with the boxed uncertainty set, the adjustable robust optimization problem like (26) is usually computational intractable [5] [6]. To resolve the first difficulty, we make the constraint (25) stricter as

$$|\mathbf{X}_{cj}|\left(\overline{\mathbf{q}^g}_j - \underline{\mathbf{q}^g}_j\right) \le \frac{\delta}{M}, \text{ for all } j \in \mathcal{S}, \text{ for all } c \in \Lambda. \tag{27}$$

In Section V, we will show via simulation that the stricter security constraint (27) has similar performance compared with the original one. To resolve the second difficulty, we use the idea of affinely adjustable robust counterparts (AARC) [6]. Specifically, we enforce an affine relationship between $\mathbf{q}^g$ and $\mathbf{U}$ as

$$\mathbf{q}^g = \mathbf{W}\mathbf{U} + \mathbf{w}, \text{ for all } \mathbf{U} \in \mathcal{U}, \tag{28}$$

where $\mathbf{W}$ denotes a $N \times N$ matrix, and $\mathbf{w}$ denotes a $N \times 1$ vector. we will choose $\mathbf{W}$ and $\mathbf{w}$ such that $\mathbf{q}^g$ given by (28) is always within $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$. Note that the affine restriction leads to a stricter constraint than (24). In summary, the AARC formulation becomes

$$\begin{aligned} &\min_{\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}, \mathbf{W}, \mathbf{w}} \quad \delta \\ &\text{subject to} \quad (27) \\ &\qquad\qquad \mathbf{W}\mathbf{U} + \mathbf{w} \in [\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}], \quad \text{for all } \mathbf{U} \in \mathcal{U}. \end{aligned} \tag{29}$$

Although the last constraint seems to involve an infinite number of inequality constraints (one for every $\mathbf{U}$), the above formulation is a standard AARC and can be transformed to a standard linear programming (LP) problem, which can then be solved effectively [5]. Due to page limits, we omit this transformation here. Interest readers can refer to [5].

Remark: We note that, when there is no attack, the distributed control law (8) generally converges to an equilibrium point that is different from the value of $\mathbf{q}^g$ given by the affine policy in (28). As a result, even when the affine policy meets the constraint (23) for all $\mathbf{U} \in \mathcal{U}$, the distributed control law (8) may not. However, in our experiments, we have found that the latter type of violation rarely happens (see Section V for details). The reason is that the affine policy is already conservative in nature. Thus, the resulting reactive power range $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$ often provides enough room for the distributed control law (8) to regulate the voltage within (23).

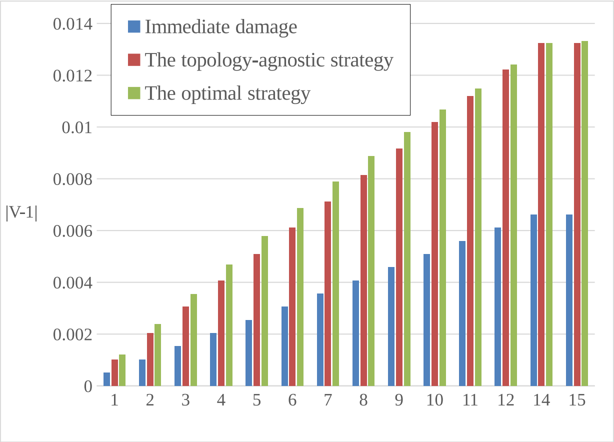

Fig. 1. Voltage disruption at the target bus 13 under different kinds of attack strategies as the attacker bus is varied.

V. Simulation Results

In this section, we use simulation to verify our damage analysis in Section III and to demonstrate the effectiveness of our proposed approach in Section IV for mitigating the damage of the potential adversary. Note that in all of the figures, we will use the unit voltage of 1 to denote the desired voltage throughout the distribution grid (also called “feeder”), even though the real voltage on these feeders differs. Further, as in Sections II-IV, we assume one DER device per bus.

A. Experiments based on a single-phase 16-bus feeder with a line topology

First, we use a simple distribution network with a line topology (which is also used in [29]). This is a 12kV distribution feeder with 16 buses on a line topology. Each line segment between two adjacent buses has the identical impedance of $(0.466 + j0.733)\Omega$. Each bus has a constant load of $(100 + j50)$kVA. The step size of the distributed control law (8) is chosen to be 15 for each bus. (This step size is reasonably away from the threshold for stability. As a value for comparison, if the step size is below 40.3175, condition (17) is satisfied, which guarantees that the system is stable. If the step size is above 40.3175, our simulation shows that the system becomes unstable.)

We focus on one target bus 13, which is further away from the substation. We assume that there is only one attacker bus, but we vary this attacker bus from 1 to 15 (except 13). The range of reactive power injection of the DER at the attacker bus is $[-100, 100]$kVAR. The DER devices on the legitimate buses have unlimited reactive power (note that the same assumption is used in Section III). In Fig. 1, we show the damage to the target bus 13 by different attack strategies, as we vary the attacker bus. Here, the "immediate damage" refers to the voltage disruption inflicted by the attacker assuming that the reactive power injection at the legitimate buses is not changed. The "topology-agnostic strategy" refers to the attacker in Section III-B that does not need information on network topology. The "optimal strategy" refers to the attacker in Section III-C that requires topology information. From Fig 1, we observe that, by leveraging the response of legitimate buses, the topology-agnostic strategy can indeed inflict twice the immediate damage. Further, although the optimal strategy inflicts slightly higher damage than the topology-agnostic strategy, the difference is small. Notice that the immediate damage and the damage of the topology-agnostic strategy are independent of the step size in the distributed control law (8). On the other hand, we find that the damage of the optimal attack strategy may increase with the step size. (This is expected because, if the step size is too large, the system will eventually become unstable. Thus, the damage of the adversary becomes arbitrarily large.) However, in our simulations we find that, when the step size is relatively small (less than 20 in this case), the damage of the optimal attack strategy is not very sensitive to the step size. Our experiments with other more realistic topologies show similar results. Thus, these results justify the use of the topology-agnostic strategy in Section IV for an estimate of the maximum damage inflicted by the adversary. Fig. 1 also shows that, if the attacker bus is closer to the substation than the target bus (i.e., the attacker on bus 1∼12), then the damage becomes larger when the attacker bus gets further away from the substation. On the other hand, if the target bus is closer to the substation than the attacker bus (i.e., the attacker on bus 14∼15), then the damage is less dependent on the location of the attacker. This behavior is because $[\mathbf{X}_{ba}]_c$ in (20) depends on the overlap between the path from the substation to the attacker bus and that to the target bus. Thus, a larger overlap leads to a larger damage.

Fig. 2 illustrates a typical attack sequence by the attacker bus 14 with the topology-agnostic strategy. The top sub-figure shows the voltage disruption on the target bus 13 over time. The bottom sub-figure shows the reactive power injection q_a^g of the attacker bus 14 by the topology-agnostic strategy. The attack starts at $t = 100$ when the attacker bus injects 100kVAR reactive power. Notice that the voltage disruption at $t = 100$ is actually the immediate damage. Due to the actions of other legitimate buses, the voltage disruption on the target bus gradually goes to zero over $100 \leq t \leq 199$. However, at $t = 200$, the attacker bus suddenly changes the sign of q_a^g

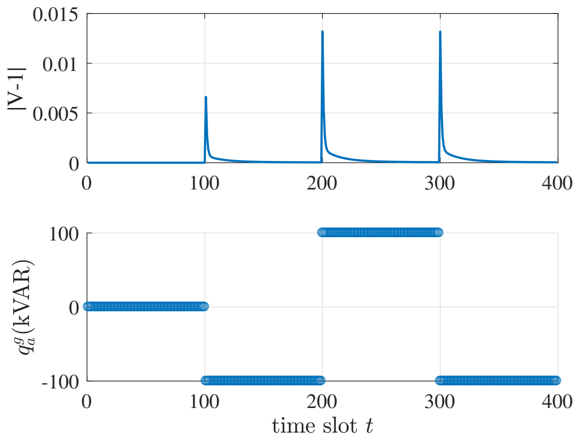

Fig. 2. Voltage disruption to the target bus 13 over time (top figure) when the attacker bus 14 uses the topology-agnostic strategy (bottom figure) to attack.

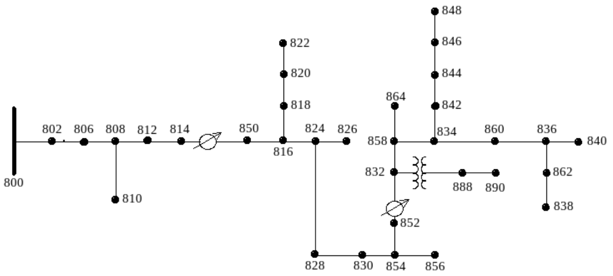

Fig. 3. IEEE 34 node test feeder.

to achieve the peak voltage disruption. Note this peak disruption is twice the immediate damage at $t = 100$. From then on, the attacker bus keeps flipping the sign of q_a^g every 100 time slots. As a result, the topology-agnostic strategy can bring twice the immediate damage at $t = 200, 300$ and so on.

B. Experiments based on the IEEE 34-node test feeder

Next, we use a more realistic topology shown in Fig. 3, which is based on an actual feeder located in Arizona [8]. It has a nominal voltage of 24.9kV. We use the configuration of phase 1 of this feeder [8] to test our proposed scheme in Section IV. Since we are mainly interested in using the reactive power injection of DERs to control voltage, we do not use voltage regulators at line 814-850 and line 852-832. To simplify our experiments, without loss of generality, we

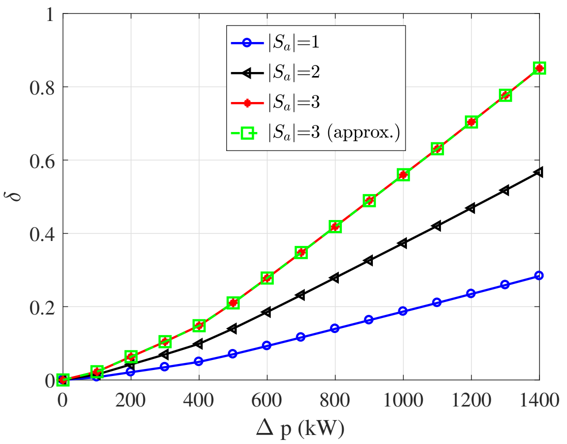

Fig. 4. Relationship between δ and uncertainty Δp with different number M of attacker buses $|\mathcal{S}_a|$.

assume that all uncertainty comes from the real power injection of DER, while the real and reactive power consumption of the normal load are fixed. To model the uncertainty set $\mathcal{U}$ when there is no attack, we assume that the level of uncertainty is the same for every bus. Specifically, the real power injection $[\mathbf{p}^g]_k$ of the DER on each bus k is within the range $[0, 2\Delta p]$. Thus, by varying Δp, we vary the level of uncertainty. When there is an attack, we aim to protect bus 890 and bus 844 because they have the highest and the second highest real-power loads, which are 150kW and 135kW, respectively. We let the maximum allowable voltage deviation $\overline{\Delta V}$ to be 10% and solve for the maximum damage δ for different size M of the set $\mathcal{S}_a$ of attackers buses.

Fig. 4 shows the relationship between δ and the amount of uncertainty Δp for different sizes of the set $\mathcal{S}_a$ of attacker buses. The three curves labeled with "$|\mathcal{S}_a| = 1$", "$|\mathcal{S}_a| = 2$", and "$|\mathcal{S}_a| = 3$" are obtained when we apply affine policy with the original security constraint (25), which exhaustively enumerates all subsets $\mathcal{S}_a$ of size M. Clearly, as either the uncertainty Δp or the size M of the set of attacker buses increases, the maximum damage δ increases. Fig. 4 can be used by the utility to determine the amount of uncertainty that it can handle while tolerating a given level of damage. For example, suppose that we want to limit the total voltage mismatch within 20%. Since $\overline{\Delta V} = 10\%$, we have $\delta \leq 0.2 - \overline{\Delta V} = 0.1$. Fig. 4 shows that when $\delta \leq 0.1$, the uncertainty Δq for $|\mathcal{S}_a| = 3$, $|\mathcal{S}_a| = 2$, and $|\mathcal{S}_a| = 1$ should be at most about 300kW, 400kW, and 600kW, respectively. The fourth curve, labeled with "$|\mathcal{S}_a| = 3$ (approx.)," is obtained when we use the stricter but simpler security constraint (27) for $|\mathcal{S}_a| = 3$. We can observe that the

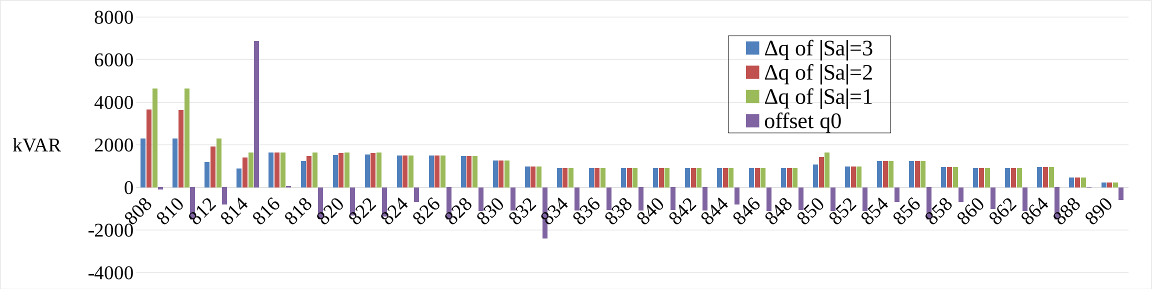

Fig. 5. The range of $\mathbf{q}^g$ for each bus when $|\mathcal{S}_a| = 3$, 2, and 1. The y-axis represents either $\Delta\mathbf{q} = (\overline{\mathbf{q}^g} - \underline{\mathbf{q}^g})/2$ (the first 3 bars) or $\mathbf{q}0 = (\overline{\mathbf{q}^g} + \underline{\mathbf{q}^g})/2$ (the forth bar). The values for buses 802 and 806 are not plotted because they are very large. They are reported here. The values for bus 802 are (30293.33, 46209.22, 65532.76, -801.50) kVAR. The values for bus 806 are (18252.31, 28936.99, 39230.87, -801.50) kVAR.

result using (27) almost perfectly matches the result that uses the original security constraint (25). This result demonstrates the effectiveness of replacing the original security constraint (25) by the simpler one (27), which reduces the complexity. A further important benefit of using the simpler constraint (27) is that δ becomes proportional to $|\mathcal{S}_a|$. This relation holds because, when we use the simpler constraint (27), the optimal solution of the range of reactive power injection for the optimization problem (29) becomes independent of $|\mathcal{S}_a|$. Hence, the utility only needs to solve (29) for $|\mathcal{S}_a| = 1$, and then extrapolate the solution for any value of $|\mathcal{S}_a|$. This is much more efficient when the maximum number of compromised DER devices is high.

To provide more details of the resulting range of the DER reactive power, we show in Fig. 5 the range of reactive power injection of DER for each bus when the uncertainty level is $\Delta p = 800\text{kW}$. Here, for each bus the first three bars represent the value of $\Delta\mathbf{q} = (\overline{\mathbf{q}^g} - \underline{\mathbf{q}^g})/2$ for $|\mathcal{S}_a| = 3$, $|\mathcal{S}_a| = 2$, and $|\mathcal{S}_a| = 1$, respectively. The last bar represents the value of $\mathbf{q}0 = (\overline{\mathbf{q}^g} + \underline{\mathbf{q}^g})/2$, which we refer to as the “offset.” From the figure, we can see that buses closer to the substation have larger reactive-power range. That is because adversaries that are closer to the substation have less influence on the voltage at the target buses (which can also be seen from Fig. 1). As a result, those buses can be allowed to have larger ranges of $\mathbf{q}^g$ without violating the security constraint. Notice that we report the values for bus 802 and bus 806 (two closest buses to the substation) in the caption instead of in the figure.

We also use simulations to verify some of the assumptions that we made in our analysis. Recall that the damage analysis in Section III is based on the assumption that legitimate buses have unlimited reactive power. Next, we evaluate the damage under the topology-agnostic strategy

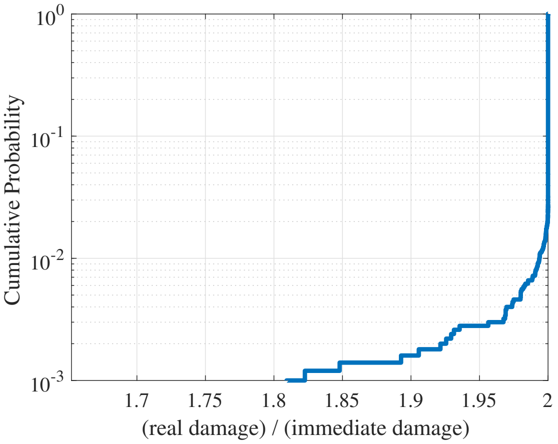

Fig. 6. The empirical cumulative distribution function of the ratio between the real damage and the immediate damage under the topology-agnostic strategy, when all buses have restricted reactive power.

when this assumption is removed. Specifically, we choose $|\mathcal{S}_a| = 3$, $\Delta p = 800$kW and obtain the optimal range of reactive power $[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$ by solving (29). Then, we uniformly randomly pick three attacker buses, one target bus, and the value of real-power $[\mathbf{p}^g]_k \in [0, 2\Delta p]$ for each bus k. Next, let attacker buses inject zero reactive power, i.e., $\mathbf{q}_a^g = \mathbf{0}$, and let the system reach the equilibrium of the distributed control law (12). We then let three attacker buses apply the topology-agnostic strategy and we record the voltage deviation from the equilibrium state on the target bus. We use this voltage deviation to quantify the damage caused by attackers. Fig. 6 plots the empirical distribution function of the ratio between this damage and the immediate damage (defined in Section III-B). From Fig. 6, we observe that, for most cases, the topology-agnostic strategy can still bring twice the immediate damage. In other cases ($\leq 2\%$ of the random cases), having bounds on the reactive power injection at legitimate buses makes the damage smaller. Thus, the figure justifies the use of our damage estimate (25) in Section IV even when the range of the reactive power at legitimate buses is also enforced.

Finally, recall the discussion at the end of Section IV that the affine policy uses different reactive power injection from that chosen by the equilibrium point of the distributed control law (8). As a result, the constraint (24) may still be violated by the distributed control law (8) even if the affine policy satisfies the constraint (23). We now use simulation to evaluate how likely this mismatch can happen. Specifically, we choose $|\mathcal{S}_a| = 1$, $\Delta p = 800$kW, and target buses being bus 890 and bus 844. After we determine the optimal range of reactive power injection

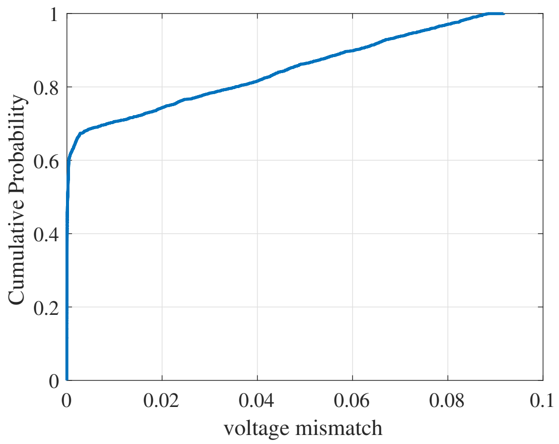

Fig. 7. The empirical cumulative distribution function of the voltage mismatch without attack.

$[\underline{\mathbf{q}^g}, \overline{\mathbf{q}^g}]$ from (29), we uniformly randomly pick the value of $[\mathbf{p}^g]_k \in [0, 2\Delta p]$ for each bus k and record the largest voltage mismatch among all buses when the local voltage control scheme (8) is used. Fig. 7 shows the empirical distribution function of that voltage mismatch. Recall that we have used $\overline{\Delta V} = 10\%$, i.e., the affine policy ensures that the voltage mismatch is always below 10% when there is no attack. From Fig. 7, we observe that the voltage mismatch under the distributed control law (8) is also below 10% for nearly all cases. In summary, the use of affine policy provides a convenient and fairly accurate way to control the voltage mismatch without attacks.

VI. Conclusions and Future Work

In this paper, we study adversarial attacks to voltage control in the power distribution network. We first study how intelligent attacker can vary the reactive power injection to inflict large voltage disruption. We demonstrate that a topology-agnostic attack strategy that can leverage the response of legitimate buses to amplify the damage by two times. We also solve the optimal attack strategy when network topology information is known to the adversary. Then, we formulate an adjustable robust optimization to control the potential damage of the adversary by optimally setting the reactive power range of the DER devices.

There are a number of interesting directions for future work. First, although we have shown by simulation that the gap between the damage of the optimal strategy and the damage of the topology-agnostic strategy is often small (when the step size of (8) is not too large), it would be desirable if we can obtain a more rigorous theoretical bound on the difference. Second, it would also be of interest to understand the optimal attack strategy when the reactive power injection of legitimate DERs is bounded (which makes the system non-linear). Further, throughout this paper we have assumed that there is no additional monitoring capability by the system to monitor reactive power injections. It would be interesting to study how to optimally launch attacks when there is additional monitoring capability, so that the attacks cannot be detected (e.g., in the manner of [21]), as well as how to best deploy such monitoring capability to limit the potential damage.

References

[1] Ahmadi, H., Martí, J. R., and Dommel, H. W. A framework for Volt-VAR optimization in distribution systems. *IEEE Transactions on Smart Grid 6*, 3 (May 2015), 1473–1483.

[2] Amin, S., Cárdenas, A. A., and Sastry, S. S. Safe and secure networked control systems under denial-of-service attacks. In *Hybrid Systems: Computation and Control* (2009), pp. 31–45.

[3] Antsaklis, P. J., and Michel, A. N. *Linear Systems*. McGraw-Hill, 1997.

[4] Baran, M. E., and Wu, F. F. Optimal capacitor placement on radial distribution systems. *IEEE Transactions on Power Delivery 4*, 1 (Jan. 1989), 725–734.

[5] Ben-Tal, A., Ghaoui, L. E., and Nemirovski, A. *Robust Optimization*. Princeton University Press, 2009.

[6] Ben-Tal, A., Goryashko, A., Guslitzer, E., and Nemirovski, A. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming 99*, 2 (2004), 351–376.

[7] Dán, G., and Sandberg, H. Stealth attacks and protection schemes for state estimators in power systems. In *2010 First IEEE International Conference on Smart Grid Communications (SmartGridComm)* (Oct. 2010), pp. 214–219.

[8] Distribution System Analysis Subcommittee. IEEE 34 Node Test Feeder, 2010. Available: http://ewh.ieee.org/soc/pes/dsacom/testfeeders/index.html.

[9] Douglas B. West. *Introduction to Graph Theory*, 2nd ed. Upper Saddle River: Prentice hall, 2001.

[10] Evangelopoulos, V. A., Georgilakis, P. S., and Hatziargyriou, N. D. Optimal operation of smart distribution networks: A review of models, methods and future research. *Electric Power Systems Research 140* (2016), 95–106.

[11] Farivar, M., and Low, S. Equilibrium and dynamics of local voltage control in distribution systems. In *2013 52nd IEEE Conference on Decision and Control* (Florence, Italy, Dec. 2013), pp. 4329–4334.

[12] Farivar, M., Neal, R., Clarke, C., and Low, S. Optimal inverter VAR control in distribution systems with high PV penetration. In *IEEE Power and Energy Society General Meeting* (San Diego, CA, USA, July 2012), pp. 1–7.

[13] Hwang, S.-G. Cauchy's interlace theorem for eigenvalues of hermitian matrices. *The American Mathematical Monthly 111*, 2 (Feb. 2004), 157–159.

[14] Jahangiri, P., and Aliprantis, D. C. Distributed Volt/VAr control by PV inverters. *IEEE Transactions on Power Systems 28*, 3 (Aug. 2013), 3429–3439.

[15] Liu, H. J., Shi, W., and Zhu, H. Decentralized dynamic optimization for power network voltage control. *IEEE Transactions on Signal and Information Processing over Networks 3*, 3 (Sept. 2017), 568–579.

[16] Liu, Y., Ning, P., and Reiter, M. False data injection attacks against state estimation in electric power grids. In *Proceedings of the 16th ACM conference on Computer and Communications Security* (Chicago, Illinois, USA, Nov. 2009), pp. 21–32.

[17] Majumdar, A., Agalgaonkar, Y., Pal, B. C., and Gottschalg, R. Centralized volt-var optimization strategy considering malicious attack on distributed energy resources control. *IEEE Transactions on Sustainable Energy 9*, 1 (Jan. 2018), 148–156.

[18] Mo, Y., and Sinopoli, B. Secure control against replay attacks. In *2009 47th Annual Allerton Conference on Communication, Control, and Computing* (Monticello, IL, USA, 2009), pp. 911–918.

[19] Moghe, R., Tholomier, D., Divan, D., Schatz, J., and Lewis, D. Grid edge control: A new approach for volt-var optimization. In *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference* (Dallas, TX, USA, May 2016), pp. 1–5.

[20] Mohsenian-Rad, A. H., and Leon-Garcia, A. Distributed internet-based load altering attacks against smart power grids. *IEEE Transactions on Smart Grid 2*, 4 (Dec. 2011), 667–674.

[21] Pasqualetti, F., Dörfler, F., and Bullo, F. Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control 58*, 11 (Nov. 2013), 2715–2729.

[22] Robbins, B. A., Hadjicostis, C. N., and Dominguez-Garcia, A. D. A two-stage distributed architecture for voltage control in power distribution systems. *IEEE Transactions on Power Systems 28*, 2 (May 2013), 1470–1482.

[23] Singh, S., and Singh, S. P. A smart volt-var optimization engine for energy distribution system. In *International Conference on Emerging Trends in Electrical, Electronics and Sustainable Energy Systems (ICETEESES)* (Sultanpur, India, Mar. 2016), pp. 35–41.

[24] Smith, R. S. A decoupled feedback structure for covertly appropriating networked control systems. *IFAC Proceedings Volumes 44*, 1 (Jan. 2011), 90–95.

[25] Sridhar, S., Hahn, A., and Govindarasu, M. Cyber–physical system security for the electric power grid. *Proceedings of the IEEE 100*, 1 (Jan. 2012), 210–224.

[26] Sridhar, S., and Manimaran, G. Data integrity attack and its impacts on voltage control loop in power grid. In *IEEE Power and Energy Society General Meeting* (Detroit, MI, USA, July 2011), pp. 1–6.

[27] Turitsyn, K., Šulc, P., Backhaus, S., and Chertkov, M. Options for control of reactive power by distributed photovoltaic generators. *Proceedings of the IEEE 99*, 6 (June 2011), 1063–1073.

[28] Zhu, H., and Li, N. Asynchronous local voltage control in power distribution networks. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Shanghai, China, Mar. 2016), pp. 3461–3465.

[29] Zhu, H., and Liu, H. J. Fast local voltage control under limited reactive power: Optimality and stability analysis. *IEEE Transactions on Power Systems 31*, 5 (Sept. 2016), 3794–3803.

Appendix

A. Proof of Lemma 2.1

Proof. Because the sum of elements in each column of the incidence matrix $\mathbf{M}^o$ is zero, we have

$$\overbrace{\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}}^{\text{N+1}} \begin{bmatrix}\mathbf{m}_0\\ \mathbf{M}\end{bmatrix} = \overbrace{\begin{bmatrix}0 & 0 & \cdots & 0\end{bmatrix}}^{\text{N+1}}$$

$$\begin{aligned}
&\Longrightarrow \mathbf{m}_0 + \overbrace{\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}}^{\text{N}} \cdot \mathbf{M} = \overbrace{\begin{bmatrix}0 & 0 & \cdots & 0\end{bmatrix}}^{\text{N}}\\
&\Longrightarrow \mathbf{M}^T \cdot \mathbf{1} = -\mathbf{m}_0^T\\
&\Longrightarrow \; - (\mathbf{M}^T)^{-1} \cdot \mathbf{m}_0^T = \mathbf{1}\\
&\Longrightarrow \; - \mathbf{M}^{-T}\mathbf{m}_0^T = \mathbf{1}.
\end{aligned}$$

□

B. Proof of Proposition 3.1

Proof. The result in [29] shows that $\mathbf{X}$ is positive definite. Recall that $\mathbf{X} = \mathbf{M}^{-T}\mathbf{D}_x\mathbf{M}^{-1}$ and $\mathbf{D}_x$ is diagonal. Thus, $\mathbf{X}$ and $\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}$ are real, positive definite, and symmetric. We can show that $\mathbf{DX}$ and $\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}$ have the same set of eigenvalues. To see this, let λ be an eigenvalue of $\mathbf{DX}$ with the corresponding eigenvector s. Then, we have

$$\begin{aligned}
&\mathbf{DXs} = \lambda\mathbf{s}\\
\Longleftrightarrow\; &\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{-\frac{1}{2}}\mathbf{s} = \mathbf{D}^{-\frac{1}{2}}\lambda\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{-\frac{1}{2}}\mathbf{s}\\
\Longleftrightarrow\; &\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\left(\mathbf{D}^{-\frac{1}{2}}\mathbf{s}\right) = \lambda\left(\mathbf{D}^{-\frac{1}{2}}\mathbf{s}\right),
\end{aligned}$$

which means that λ is also an eigenvalue of $\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}$. Because $\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}$ is real, positive definite, and symmetric, the eigenvalues of $\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}$ (and also the eigenvalues of $\mathbf{DX}$) are all real and positive. By the assumption of the proposition, we then have $\lambda_{\max}\left(\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\right) < 2$. Now, consider the matrix $\mathbf{X}_{bb}$ when some buses are malicious. Using a similar argument, we conclude that $\mathbf{D}_b\mathbf{X}_{bb}$ has the same eigenvalues with $\mathbf{D}_b^{\frac{1}{2}}\mathbf{X}_{bb}\mathbf{D}_b^{\frac{1}{2}}$ We now use Cauchy Interlace Theorem [13], which is stated below for readers' convenience:

Theorem A.1 (Cauchy Interlace Theorem)**.** *Let A be a Hermitian matrix of order n, and let B be a principle sub-matrix of A of order $n-1$. If $\lambda_n \leq \lambda_{n-1} \leq \cdots \leq \lambda_1$ are the eigenvalues of*

A and $\mu_n \leq \mu_{n-1} \leq \cdots \mu_2$ are the eigenvalues of B, then $\lambda_n \leq \mu_n \leq \lambda_{n-1} \leq \mu_{n-1} \leq \cdots \leq \lambda_2 \leq \mu_2 \leq \lambda_1$.

Since $\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}$ is symmetric and $\mathbf{D}_b^{\frac{1}{2}}\mathbf{X}_{bb}\mathbf{D}_b^{\frac{1}{2}}$ is a principle sub-matrix of $\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}$, by repeatedly applying Cauchy Interlace Theorem, we have $\lambda_{\max}\left(\mathbf{D}_b^{\frac{1}{2}}\mathbf{X}_{bb}\mathbf{D}_b^{\frac{1}{2}}\right) \leq \lambda_{\max}\left(\mathbf{D}^{\frac{1}{2}}\mathbf{X}\mathbf{D}^{\frac{1}{2}}\right)$, which implies that $\lambda_{\max}\left(\mathbf{D}_b\mathbf{X}_{bb}\right) = \lambda_{\max}\left(\mathbf{D}_b^{\frac{1}{2}}\mathbf{X}_{bb}\mathbf{D}_b^{\frac{1}{2}}\right) < 2$. Using Theorem 10.17 of [3, pp. 508], we conclude that the linear system described by (15) and (16) is BIBO stable. □