

Online Voltage Control for Unbalanced Distribution Networks Using Projected Newton Method

Rui Cheng, *Graduate Student Member, IEEE*, Zhaoyu Wang, *Senior Member, IEEE*, Yifei Guo, *Member, IEEE*, Qianzhi Zhang, *Graduate Student Member, IEEE*

Abstract—This paper proposes an online voltage control strategy of distributed energy resources (DERs), based on the projected Newton method (PNM), for unbalanced distribution networks. The optimal Volt/VAr control (VVC) problem is formulated as an optimization program, with the goal of maintaining the voltage profile across the network by coordinating the VAr outputs of DERs. To overcome the slow convergence rate of conventional gradient-based methods, a PNM-based solution algorithm is developed to solve this VVC problem. It utilizes a non-diagonal symmetric positive definite matrix, developed from the Hessian matrix of the objective, to scale the gradient, and thus a fast convergence performance can be expected in this Newton-like algorithm. Moreover, taking advantage of the instantaneous feedback of voltage measurements, the online implementation of the PNM-based VVC is further designed to deal with fast system variations. In this online PNM-based VVC scheme, each bus agent communicates the instantaneous voltage measurements to the central agent, and the central agent communicates the VAr output commands of DERs back to each bus agent. The fast convergence performance of PNM results in a stronger capability to track the system variations in real time. Finally, numerical case studies are performed to validate the effectiveness, superiority, and scalability of the proposed method.

Index Terms—Online voltage control, distributed energy resources (DERs), unbalanced distribution networks, projected Newton method (PNM).

I. INTRODUCTION

THE increasing proliferation of distributed energy resources (DERs), such as photovoltaic (PV) generators, has posed new challenges to Volt/VAr control (VVC) problems in distribution networks due to the intermittent nature of renewable energy resources. However, meanwhile, it also provides promising opportunities to utilize DERs to resolve the VVC problems due to the great advance in the inverter-based technologies of DERs [1].

The VVC problems are often cast as optimization programs with various control goals. Some studies have made efforts to explore different VVC methods to coordinate DERs in distribution networks. In [2], the VVC problem is formulated as an optimal power flow (OPF) problem through a Second-Order Cone Program to minimize line losses and energy consumption in single-phase radial distribution networks. A sequential

This work was supported in part by the U.S. Department of Energy Wind Energy Technologies Office under Grant DE-EE0008956, and in part by the National Science Foundation under ECCS 1929975 (Corresponding author: Zhaoyu Wang).

Rui Cheng, Zhaoyu Wang, Yifei Guo and Qianzhi Zhang are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: ruicheng@iastate.edu; wzy@iastate.edu; yfguo.sdu@gmail.com; qianzhi@iastate.edu).

convex programming is applied to minimize total injected reactive power while satisfying the operational constraints of unbalanced distribution networks [3]. However, the works [2], [3] are solved in the centralized manner, entailing large amounts of computational burdens and time.

Considering the high proliferation of DERs in distribution networks, various distributed VVC methods have attracted growing attentions, which can be roughly classified into two categories: hierarchical and decentralized VVC strategies. Hierarchical VVC strategies establish on the communication between the central agent and local agents. In [4]-[5], the hierarchical VVC strategies, based on the Alternating Direction Method of Multipliers (ADMM), are applied to coordinate the electric vehicle charging and wind turbines, respectively. In [6], an ADMM-based hierarchical management strategy is developed between the utility and customers to optimally dispatch real and reactive power set-points for residential PV inverters. In contrast, decentralized VVC strategies mainly rely on the communication between neighboring buses. The study by Zhang et al. [7] proposes a dual decomposition-based decentralized VVC method to minimize power losses in single-phase networks, relying on the utilization of reactive-power-capable DERs. The authors in [8] propose a voltage-constrained ADMM-based algorithm that optimally coordinates the VAr outputs of DERs in unbalanced distribution systems to regulate bus voltages. The ADMM-based strategies are utilized in [9]-[11] to solve VVC problems as well. Those VVC problems in [2]-[11] are solved in an *offline* manner. That is, the solution in the offline voltage control cannot be applied to distribution networks till the iteration convergences. Thus, they might not be able to capture the fast fluctuations and disturbances in distribution networks caused by the high variability of DERs, potentially leading to undesirable and unacceptable results.

In recent years, distribution networks are undergoing a fundamental architecture transformation to become more intelligent, controllable, and open due to the increasing deployment of communication, computing, information devices across distribution networks, such as advanced metering infrastructure. It offers an unprecedented promising opportunity to transform the voltage control scheme from the open-loop *offline* manner to the closed-loop *online* manner.

Different than the offline algorithms, the online algorithms naturally track changing network conditions as these changes manifest themselves in the network state that is used to compute the control solution [12]. The online VVC strategy has already triggered a wide-spread interest. One category is the

traditional “droop” control [13]-[15], which actively adjusts its VAr output as a function of local bus voltage. However, the work [16] points out that the droop control cannot maintain a feasible voltage profile under certain circumstances. Consequently, other more sophisticated VVC strategies have been proposed. The authors in [16]-[18] propose various online VVC schemes based on a dual ascent method. In [19], a primal-dual gradient algorithm is leveraged to regulate voltage in real time through the iterative update of primal and dual variables. A gradient projection (GP) method is applied to design online voltage control with the goal of minimizing the voltage deviation in [20]. Liu et al. [21] propose an ADMM-based VVC strategy in an online fashion to solve voltage regulation problems, subject to time-varying operating conditions. Those underlying methodologies in online VVC strategies [16]-[21] can be regarded as first-order gradient-based algorithms. A known drawback of gradient-based algorithms is that they may suffer from a slow convergence rate and are highly affected by the condition number of the problem [22]. However, the convergence rate of algorithms is of great importance for the online implementation since it significantly affects the tracking capabilities. Researchers and practitioners have developed some newly online VVC strategies to improve the tracking capabilities. An accelerated dual descent algorithm is proposed to solve the voltage regulation problem in single-phase networks [23] by employing the dual decomposition and accelerated gradient projected techniques. But this work is established on the assumption that the single-phase network lines have the same ratio of resistance to reactance. The works [24], [25] propose feedback-based VVC strategies by means of a diagonally scaled gradient projection (DSGP) method, adopting a diagonal scaling of gradient to improve the convergence performance; however, the DSGP method theoretically only has a linear convergence rate [26]. Moreover, although the simulations of [24] are tested on unbalanced distribution networks, the theoretical analysis of [24] only builds on single-phase distribution networks. The Newton’s method has always been a popular option to improve the convergence performance by directly scaling the gradient with the inverse Hessian matrix of the objective. However, using the inverse Hessian matrix directly could be problematic for constrained optimization problems. As indicated in [22, Sec. 2.4], the Newton’s update may fail to attain the optimal solution of a constrained program. A real-time OPF algorithm, based on quasi-Newton methods, is developed in [27] for single-phase networks. Nevertheless, it requires solving sub-optimization programs in the online implementation, thus its effectiveness is dependent upon the computational burdens of solving sub-optimization programs.

In this context, an online VVC scheme, based on the projected Newton method (PNM) [28], is proposed to adjust the VAr outputs of DERs to maintain the voltage profile across unbalanced distribution networks. In this online PNM-based scheme, each bus agent sends the local voltage measurement as feedback to the central agent; then, the central agent communicates the VAr commands back to each agent to optimally coordinate the DERs across the unbalanced distribution network. The main contributions of this paper are summarized as

follows:

- The second-order PNM is applied to solve the constrained VVC problem with the goal of maintaining the voltage profile. Instead of directly using the inverse Hessian matrix of the objective, PNM adopts a non-diagonal positive definite symmetric matrix, developed from the Hessian matrix, to scale the gradient so that the resulting Newton-like algorithm can guarantee a fast convergence rate, contributing to a great capability to track the time-varying changes in real time.
- Taking advantage of the instantaneous local voltage measurement from each bus agent, an online feedback-based PNM-based hierarchical VVC strategy is further developed. Each local bus agent sends its local voltage measurement to the central agent, and the central agent communicates back to each bus agent the VAr output command. Under this online VVC strategy, the central agent and bus agents are not required to solve any (sub) optimization program. This attribute can effectively reduce the computational burdens, facilitating the online implementation to cope with fast system variations.
- The theoretical analysis and simulations are both built on unbalanced distribution networks. We mathematically analyze this PNM-based VVC scheme, based on the linearized power flow model [29] of unbalanced distribution networks. And we test the performance of this PNM-based VVC scheme on the multi-phase nonlinear power flow in numerical case studies. Numerical comparisons with other VVC strategies are also provided to demonstrate the superiority of our proposed method.

Remaining sections are organized as follows. The system modeling and problem statement are presented in Section II. The details of leveraging PNM to solve the VVC problem are discussed in Section III. Section IV outlines the online implementation strategy of the proposed PNM-based VVC. Numerical case studies and results are provided in Section V. Concluding comments are given in Section VI.

II. SYSTEM MODELING AND PROBLEM STATEMENT

A. Unbalanced Radial Distribution Networks

Consider an unbalanced radial distribution network with $N+1$ buses. Let $\{0\} \cup \mathcal{N}$ denote the index set for these buses, where $\mathcal{N} = \{1, 2, \dots, N\}$, bus 0 is the feeder head bus. For each bus $i \in \mathcal{N}$, let $bp(i)$ denote the bus that immediately precedes bus i along the radial network headed by bus 0. Let $\mathcal{L} = \{\ell_j = (i, j) | i = bp(j), j \in \mathcal{N}\}$ denote this edge set of line segments. Also, let \mathcal{N}_i denote the set of all buses located strictly after bus i along the radial network.

Let Φ_i and n_i denote the phase set and number of bus $i \in \{0\} \cup \mathcal{N}$, Φ_{ij} and n_{ℓ_j} denote the phase set and number of line segment $\ell_j = (i, j) \in \mathcal{L}$. Note that for $\forall \ell_j \in \mathcal{L}$, $n_{\ell_j} = n_j$. For each bus $i \in \{0\} \cup \mathcal{N}$, $V_i^\phi(t)$ and $v_i^\phi(t)$ denote its phase ϕ complex voltage and squared voltage magnitude, respectively, and define the following column vectors: $\mathbf{V}_i(t) = [V_i^\phi(t)]_{\phi \in \Phi_i} \in \mathbb{C}^{n_i}$, $\mathbf{V}_i^{\Phi_{ij}}(t) = [V_i^\phi(t)]_{\phi \in \Phi_{ij}} \in \mathbb{C}^{n_{\ell_j}}$, $\mathbf{v}_i(t) = [v_i^\phi(t)]_{\phi \in \Phi_i} \in \mathbb{R}^{n_i}$, $\mathbf{v}_i^{\Phi_{ij}}(t) = [v_i^\phi(t)]_{\phi \in \Phi_{ij}} \in \mathbb{R}^{n_{\ell_j}}$.

TABLE I
NOMENCLATURE: OPERATOR

$\text{diag}(\mathbf{u})$	A diagonal matrix with the entries of \mathbf{u} in its diagonal
\mathbf{u}^H	The conjugate transpose of \mathbf{u}
$\ \boldsymbol{\mu}\ _C^2$	It denotes $\boldsymbol{\mu}^T \mathbf{C} \boldsymbol{\mu}$
$[\boldsymbol{\mu}]_i$	For vector $\boldsymbol{\mu}$, let $[\boldsymbol{\mu}]_i$ denote the i -th element in $\boldsymbol{\mu}$
$[\mathbf{U}]_{ij}$	For matrix \mathbf{U} , let $[\mathbf{U}]_{ij}$ denote the i -th row and j -th column element in \mathbf{U}
$[\cdot]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g}$	Projection operator onto the box constraint $[\underline{\mathbf{q}}^g, \bar{\mathbf{q}}^g]$
\odot	Element-wise multiplication
\oslash	Element-wise division

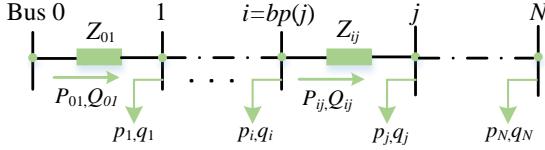


Fig. 1. A radial distribution network.

And let column vectors $\mathbf{p}_i(t) = [p_i^\phi(t)]_{\phi \in \Phi_i}$ and $\mathbf{q}_i(t) = [q_i^\phi(t)]_{\phi \in \Phi_i} \in \mathbb{R}^{n_i}$ denote the real and reactive power consumption at bus i . For each line segment $(i, j) \in \mathcal{L}$, let $\mathbf{P}_{ij}(t) = [P_{ij}^\phi(t)]_{\phi \in \Phi_{ij}}$ and $\mathbf{Q}_{ij}(t) = [Q_{ij}^\phi(t)]_{\phi \in \Phi_{ij}} \in \mathbb{R}^{n_{\ell_j}}$ denote the real and reactive power flow from bus i to j , and $\mathbf{Z}_{ij} \in \mathbb{C}^{n_{\ell_j} \times n_{\ell_j}}$ denote the impedance matrix for line segment (i, j) . The nonlinear distribution power flow can be represented as follows, for $\forall (i, j) \in \mathcal{L}$,

$$\mathbf{V}_j(t) = \mathbf{V}_i^{\Phi_{ij}}(t) - \mathbf{Z}_{ij}[(\mathbf{P}_{ij}(t) - \mathbf{Q}_{ij}(t)) \oslash \mathbf{V}_i^{\Phi_{ij}}(t)] \quad (1)$$

The nonlinear distribution power flow (1) poses challenges for optimization problems due to its nonconvex nature. Some convex relaxation approaches, e.g., semidefinite programming relaxation ([29], [30]), are leveraged to convexify the nonlinear distribution power flow. However, those convex relaxation approaches may be computationally expensive, and the exactness of convex relaxation may not be guaranteed. Instead, the linearized distribution power flow ([8], [29]) exhibits lower complexity and higher computational efficiency by assuming (i) voltages among phases are nearly balanced; (ii) line losses are small. Compared to the nonlinear distribution power flow and convex relaxation approaches, the linearized power flow model is easier to facilitate the algorithm design in optimization problems due to its linear and convex characteristics. The linearized distribution power flow generalizes the LinDistFlow equations [31] from single-phase networks to multiphase networks.¹ The linearized distribution power flow can be presented as follows, for $\forall (i, j) \in \mathcal{L}$,

$$\mathbf{v}_j(t) = \mathbf{v}_i^{\Phi_{ij}}(t) - 2(\tilde{\mathbf{R}}_{ij} \mathbf{P}_{ij}(t) + \tilde{\mathbf{X}}_{ij} \mathbf{Q}_{ij}(t)) \quad (2a)$$

$$\mathbf{P}_{ij}(t) = \sum_{k \in N_j} \mathbf{P}_{jk}(t) + \mathbf{p}_j(t) \quad (2b)$$

$$\mathbf{Q}_{ij}(t) = \sum_{k \in N_j} \mathbf{Q}_{jk}(t) + \mathbf{q}_j(t) \quad (2c)$$

¹We refer readers to [8], [29] for more details regarding the linearized distribution power flow.

with

$$\tilde{\mathbf{Z}}_{ij} = \tilde{\mathbf{R}}_{ij} + j \tilde{\mathbf{X}}_{ij} = \left[(\mathbf{a} \mathbf{a}^H)^{\Phi_{ij}} \odot \mathbf{Z}_{ij}^* \right]^* \quad (3a)$$

$$\mathbf{a} = [1, e^{-j\frac{2\pi}{3}}, e^{j\frac{2\pi}{3}}]^T \quad (3b)$$

$$\mathbf{a}^H = [1, e^{j\frac{2\pi}{3}}, e^{-j\frac{2\pi}{3}}] \quad (3c)$$

where $(\mathbf{a} \mathbf{a}^H)^{\Phi_{ij}} \in \mathbb{Z}^{n_{\ell_j} \times n_{\ell_j}}$ denotes the sub-matrix of $\mathbf{a} \mathbf{a}^H$, consisting of entries associated with Φ_{ij} .

B. Compact Form of Power Flow Model

For an unbalanced radial distribution network, let $m = \sum_{j=1}^N n_{\ell_j} = \sum_{i=1}^N n_i$. Let $\bar{\mathbf{A}} = [\mathbf{A}_0, \mathbf{A}^T]^T \in \mathbb{R}^{(n_0+m) \times m}$ be the incidence matrix of unbalanced radial distribution network, where $\mathbf{A}_0^T \in \mathbb{R}^{n_0 \times m}$ represents the connection structure between bus 0 and each of the line segments in \mathcal{L} , $\mathbf{A} \in \mathbb{R}^{m \times m}$ represents the connection structure between the remaining buses and each of the line segments in \mathcal{L} ². More precisely, the incidence matrix $\bar{\mathbf{A}}$ with an entry 1 for each “from” phase node and -1 for each “to” phase node corresponding to each phase circuit of line segments takes the following form:

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{J}(0, \ell_1) & \mathbf{J}(0, \ell_2) & \dots & \mathbf{J}(0, \ell_N) \\ \mathbf{J}(1, \ell_1) & \mathbf{J}(1, \ell_2) & \dots & \mathbf{J}(1, \ell_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{J}(N, \ell_1) & \mathbf{J}(N, \ell_2) & \dots & \mathbf{J}(N, \ell_N) \end{bmatrix} \quad (4)$$

where $\mathbf{J}(i, \ell_j) \in \mathbb{R}^{n_i \times n_{\ell_j}}$ indicates the connection structure between bus i and line segment ℓ_j ; see [32, Appendix B] for one numerical example illustrating the construction of $\bar{\mathbf{A}}$ for an unbalanced radial network.

Let the squared voltage magnitudes, real and reactive power consumption, and real and reactive power flows over line segments be denoted for the network by the following column vectors³: $\mathbf{v} = [\mathbf{v}_i]_{i \in \mathcal{N}}$, $\mathbf{p} = [\mathbf{p}_i]_{i \in \mathcal{N}}$, $\mathbf{q} = [\mathbf{q}_i]_{i \in \mathcal{N}}$, $\mathbf{P} = [\mathbf{P}_{bp(i)}]_{bp(i) \in \mathcal{L}}$, $\mathbf{Q} = [\mathbf{Q}_{bp(i)}]_{bp(i) \in \mathcal{L}}$, $\mathbf{v}_0 = [v_0^a, v_0^b, v_0^c]^T$. The linearized distribution power flow model (2) can be written in a compact form:

$$-\mathbf{A}_0 \mathbf{v}_0(t) - \mathbf{A}^T \mathbf{v}(t) = -2\mathbf{D}_r \mathbf{P}(t) - 2\mathbf{D}_x \mathbf{Q}(t) \quad (5a)$$

$$-\mathbf{A} \mathbf{P}(t) = \mathbf{p}(t) \quad (5b)$$

$$-\mathbf{A} \mathbf{Q}(t) = \mathbf{q}(t) \quad (5c)$$

with

$$\mathbf{D}_r = \text{blkdiag}[\tilde{\mathbf{R}}_{bp(1)1}, \dots, \tilde{\mathbf{R}}_{bp(N)N}] \quad (6a)$$

$$\mathbf{D}_x = \text{blkdiag}[\tilde{\mathbf{X}}_{bp(1)1}, \dots, \tilde{\mathbf{X}}_{bp(N)N}] \quad (6b)$$

We separate $\mathbf{q}(t)$ into two parts $\mathbf{q}^g(t)$ and $\mathbf{q}^c(t)$, where $\mathbf{q}(t) = \mathbf{q}^c(t) - \mathbf{q}^g(t)$, $\mathbf{q}^g(t)$ denotes the vector of reactive power generated by DERs, e.g., PV inverters, and $\mathbf{q}^c(t)$ denotes the reactive power consumption vector generated by other resources

²Note that $\bar{\mathbf{A}}$ is invertible, see [32, Appendix C] for details.

³The squared voltage magnitudes and the real and reactive loads at buses i are sorted in accordance with the ordering of these buses from small to large i . The real and reactive power flows over line segments ℓ_j are sorted in accordance with the ordering of these line segments from small to large j .

except DERs. Substituting $\mathbf{q}(t) = \mathbf{q}^c(t) - \mathbf{q}^g(t)$, (5b) and (5c) into (5a), then we have:

$$\mathbf{v}(t) = \underbrace{2[\mathbf{A}^T]^{-1}\mathbf{D}_x\mathbf{A}^{-1}}_{\mathbf{M}}\mathbf{q}^g(t) + \mathbf{c}(t) \quad (7)$$

with

$$\begin{aligned} \mathbf{c}(t) &= -\mathbf{M}\mathbf{q}^c(t) - [\mathbf{A}^T]^{-1}\mathbf{A}_0\mathbf{v}_0(t) \\ &\quad - 2[\mathbf{A}^T]^{-1}\mathbf{D}_r\mathbf{A}^{-1}\mathbf{p}(t) \end{aligned} \quad (8)$$

where vector $\mathbf{c}(t)$ is a function of $\mathbf{v}_0(t)$, $\mathbf{p}(t)$, $\mathbf{q}^c(t)$, reflecting the impacts of slack bus, real power, and other reactive power except DERs.

Remark 1: The linearized distribution power flow (7) is adopted in this work since the structure of (7) can facilitate the algorithm design and theoretical analysis. Note that our proposed voltage control scheme can also be applied to the multi-phase nonlinear power flow model. In our numerical case studies, we test the performance of our proposed voltage control scheme on the nonlinear model using the open-source simulator OpenDSS [33].

C. Problem Statement

The goal of voltage control problem is assumed to minimize the voltage deviations by optimally coordinating the VAr outputs of DERs⁴, which can be formulated as follows:

$$\min_{\mathbf{q}^g(t), \mathbf{v}(t)} \frac{1}{2}\|\mathbf{v}(t) - \mathbf{v}_r\|_2^2 \quad (9a)$$

$$\text{subject to: } \underline{\mathbf{q}}^g(t) \leq \mathbf{q}^g(t) \leq \bar{\mathbf{q}}^g(t) \quad (9b)$$

$$\mathbf{v}(t) = \mathbf{M}\mathbf{q}^g(t) + \mathbf{c}(t) \quad (9c)$$

where $\mathbf{v}_r \in \mathbb{R}^m$ is the reference of squared voltage magnitude, (9b) is the VAr limits for DERs, (9c) is the compact form of linearized distribution power flow constraints.

III. VOLTAGE CONTROL USING PROJECTED NEWTON METHOD

A. Problem Reformulation

Substituting (9c) into (9a), we have:

$$h(\mathbf{q}^g(t)) = \frac{1}{2}\|\mathbf{v}(t) - \mathbf{v}_r\|_2^2 = \frac{1}{2}\|\mathbf{M}\mathbf{q}^g(t) + \mathbf{c}(t) - \mathbf{v}_r\|_2^2 \quad (10)$$

Let $\mathcal{X}(t) = \{\mathbf{q}^g(t) | \underline{\mathbf{q}}^g(t) \leq \mathbf{q}^g(t) \leq \bar{\mathbf{q}}^g(t)\}$. Then, the problem (9) can be reformulated as follows:

$$\min_{\mathbf{q}^g(t) \in \mathcal{X}(t)} h(\mathbf{q}^g(t)) \quad (11)$$

Remark 2: With respect to single-phase radial distribution networks, the works [20] and [24] adopt $\|\cdot\|_{M^{-1}}^2$ to replace $\|\cdot\|_2^2$ in (10)⁵, which facilitates the local/decentralized voltage control design. There might be two limitations in this way. First, the surrogate VVC problem with $\|\cdot\|_{M^{-1}}^2$ assigns different

⁴For this VVC problem, we do not consider the hard voltage constraint, but instead treat the voltage constraint as a soft penalty in the objective function to facilitate the algorithm design, just like [23], [24].

⁵ \mathbf{M} is a symmetric positive definite matrix in single-phase radial distribution networks.

weights to the squared voltage magnitude deviations and their product, which is not physical interpretable. Second, it might be improper to replace $\|\cdot\|_2^2$ by $\|\cdot\|_{M^{-1}}^2$ in unbalanced radial distribution networks. Since \mathbf{D}_x is not a symmetric matrix in unbalanced radial distribution networks, \mathbf{M} is no longer a symmetric positive definite matrix in unbalanced radial distribution networks.

Proposition 1: Let \mathbf{H} denote the symmetric Hessian matrix of (10), where $\mathbf{H} = \nabla^2 h(\mathbf{q}^g(t)) = \mathbf{M}^T \mathbf{M}$. The Hessian matrix \mathbf{H} is a symmetric positive definite matrix as \mathbf{D}_x is invertible.

Proof of Proposition 1: \mathbf{H} can be represented as follows:

$$\begin{aligned} \mathbf{H} &= \mathbf{M}^T \mathbf{M} = 4(\mathbf{A}^{-T} \mathbf{D}_x^T \mathbf{A}^{-1}) \cdot (\mathbf{A}^{-T} \mathbf{D}_x \mathbf{A}^{-1}) \\ &= 4\mathbf{A}^{-T} \mathbf{D}_x^T \cdot (\mathbf{A}^{-1} \mathbf{A}^{-T}) \cdot \mathbf{D}_x \mathbf{A}^{-1} \end{aligned} \quad (12)$$

For any non-zero row vector \mathbf{u} , we have:

$$\mathbf{u}\mathbf{A}^{-1}\mathbf{A}^{-T}\mathbf{u}^T = (\mathbf{u}\mathbf{A}^{-1}) \cdot (\mathbf{u}\mathbf{A}^{-1})^T > 0 \quad (13)$$

It means that $\mathbf{A}^{-1}\mathbf{A}^{-T}$ is positive definite. Then, we have:

$$\begin{aligned} \mathbf{u}(\mathbf{M}^T \mathbf{M})\mathbf{u}^T &= \\ 4(\mathbf{u}\mathbf{A}^{-T} \mathbf{D}_x^T) \cdot (\mathbf{A}^{-1} \mathbf{A}^{-T}) \cdot (\mathbf{u}\mathbf{A}^{-T} \mathbf{D}_x^T)^T & \end{aligned} \quad (14)$$

As \mathbf{D}_x is invertible, $\mathbf{u}\mathbf{A}^{-T} \mathbf{D}_x^T$ is a non-zero row vector. Plus $\mathbf{A}^{-1}\mathbf{A}^{-T}$ is positive definite, it follows that $\mathbf{u}(\mathbf{M}^T \mathbf{M})\mathbf{u}^T > 0$ which implies that $\mathbf{M}^T \mathbf{M}$ is positive definite. Plus $\mathbf{H}^T = \mathbf{M}^T \mathbf{M} = \mathbf{H}$, it follows that $\mathbf{H} \in \mathbb{S}_{++}^m$ is a symmetric positive definite matrix. Q.E.D.

B. Projected Newton Method

In this subsection, we propose a voltage control scheme using the projected Newton method. To facilitate algorithm design and theoretic analysis, we assume $\mathbf{c}(t) = \mathbf{c}$, $\underline{\mathbf{q}}^g(t) = \underline{\mathbf{q}}^g$, $\bar{\mathbf{q}}^g(t) = \bar{\mathbf{q}}^g$ are fixed. But this is not required when applying our method, the online implementation of our proposed method is discussed in Section IV to deal with the time-varying conditions. The problem (11) is a strictly convex problem with box constraints since the Hessian matrix \mathbf{H} is positive definite (See Proposition 1). For this type of problem, it can be solved by the GP method, which iteratively updates the VAr outputs of DERs at step t in the following manner:

$$\mathbf{q}^g(t+1) = [\mathbf{q}^g(t) - \nabla h(\mathbf{q}^g(t))]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g} \quad (15)$$

with

$$\nabla h(\mathbf{q}^g(t)) = \mathbf{M}[\mathbf{v}(\mathbf{q}^g(t)) - \mathbf{v}_r] \quad (16a)$$

$$\mathbf{v}(\mathbf{q}^g(t)) = \mathbf{M}\mathbf{q}^g(t) + \mathbf{c} \quad (16b)$$

where $[\cdot]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g}$ denotes the projection operator onto the box constraint $[\underline{\mathbf{q}}^g, \bar{\mathbf{q}}^g]$. However, the condition number of \mathbf{H} could be large, the gradient-based methods, e.g., GP, always suffer from the slow convergence rate.

To resolve the slow convergence implementation dilemma, one possible method is to scale the gradient with some positive definitive matrices such that:

$$\mathbf{q}^g(t+1) = [\mathbf{q}^g(t) - \alpha(t)\mathbf{u}(t)]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g} \quad (17)$$

with

$$\mathbf{u}(t) = \mathbf{D}(t)\nabla h(\mathbf{q}^g(t)) \quad (18)$$

where $\alpha(t)$ and $\mathbf{D}(t)$ are a positive step size and a positive definite matrix. The work [24] makes use of a diagonal positive definite matrix $\mathbf{D}(t)$ to achieve a proper diagonal scaling of ∇h . However, the convergence rate of DSGP method is still not ideal, which is typically characterized by the linear convergence rate [26]. Any attempt to construct a superlinearly convergent algorithm should by necessity involve a non-diagonal scaling matrix. This motivates us to explore the use of more general, nondiagonal scaling of ∇h . The Newton's method is always a popular approach by scaling the gradient with the inverse Hessian matrix of the objective. However, the scaling manner by directly using the inverse Hessian matrix could be problematic when it comes to constrained optimization problems with a projection operator. The iterated manner is not always in general a descent iteration. That is, we may have $h(\mathbf{q}^g(t+1)) > h(\mathbf{q}^g(t))$ for all positive step size $\alpha(t)$ with an unfavorable choice of $\mathbf{D}(t)$ (see [22, Sec. 2.4] for details). In short, the scaling manner for constrained optimization problems is far more complicated than their unconstrained counterparts. The appropriate selection of $\mathbf{D}(t)$ is of great importance for the scaling manner for constrained optimization problems. To this end, the following Proposition 2 identifies a class of matrices $\mathbf{D}(t)$ for which a descent iteration is obtained. Proposition 2 can be easily proved by using [28, Proposition 1].

Definition (\mathbf{D} is diagonal with respect to \mathbf{I}): If a symmetric $m \times m$ matrix \mathbf{D} and a subset of indices $\mathbf{I} \subset \{1, 2, \dots, m\}$ satisfy:

$$[\mathbf{D}]_{ij} = 0, \text{ for } i \in \mathbf{I}, j = 1, 2, \dots, m, i \neq j \quad (19)$$

then we say that \mathbf{D} is diagonal with respect to \mathbf{I} .

And define $\hat{\mathbf{I}}(t)$:

$$\begin{aligned} \hat{\mathbf{I}}(t) = & \left\{ i | [\mathbf{q}^g(t)]_i = [\underline{\mathbf{q}}^g]_i \text{ and } \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} > 0; \right. \\ & \left. \text{or } [\mathbf{q}^g(t)]_i = [\bar{\mathbf{q}}^g]_i \text{ and } \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} < 0 \right\} \end{aligned} \quad (20)$$

Proposition 2: Assume $\mathbf{D}(t)$ is a symmetric positive definite matrix which is diagonal with respect to $\hat{\mathbf{I}}(t)$, then we have:

- **[P2.a]** If $\mathbf{q}^g(t)$ is the optimal solution \mathbf{q}^{g*} of the problem (11), then:

$$\begin{aligned} \mathbf{q}^g(t+1) &= [\mathbf{q}^g(t) - \alpha(t)\mathbf{u}(t)]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g} \\ &= \mathbf{q}^g(t), \forall \alpha(t) > 0 \end{aligned}$$

- **[P2.b]** If $\mathbf{q}^g(t)$ is not the optimal solution \mathbf{q}^{g*} of the problem (11), then there exists a scalar $\bar{\alpha}(t)$ such that:

$$h(\mathbf{q}^g(t+1)) < h(\mathbf{q}^g(t)), \forall \alpha(t) \in (0, \bar{\alpha}(t)]$$

with

$$\begin{aligned} \mathbf{q}^g(t+1) &= [\mathbf{q}^g(t) - \alpha(t)\mathbf{u}(t)]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g} \\ \mathbf{u}(t) &= \mathbf{D}(t)\nabla h(\mathbf{q}^g(t)), \text{ defined in (18).} \end{aligned}$$

Proposition 2 shows that the iteration essentially terminates at a optimal solution and is capable of descent when not at a

global minimum. Unfortunately, $\hat{\mathbf{I}}(t)$ in Proposition 2 exhibits an undesirable discontinuity at the boundary of the constraint set, i.e., $\underline{\mathbf{q}}^g$ and $\bar{\mathbf{q}}^g$, which could have an adverse effect on its rate of convergence. (This phenomenon is quite common in feasible direction algorithms and is referred to as zigzagging or jamming). Consequently, we aim to introduce $\mathbf{I}(t)$, certain enlargements of $\hat{\mathbf{I}}(t)$, to bypass those difficulties, where $\mathbf{I}(t)$ is defined as follows:

$$\begin{aligned} \mathbf{I}(t) = & \left\{ i | [\underline{\mathbf{q}}^g]_i \leq [\mathbf{q}^g(t)]_i \leq [\bar{\mathbf{q}}^g]_i + [\varepsilon(t)]_i \text{ and } \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} > 0; \right. \\ & \left. \text{or } [\bar{\mathbf{q}}^g]_i - [\varepsilon(t)]_i \leq [\mathbf{q}^g(t)]_i \leq [\bar{\mathbf{q}}^g]_i \text{ and } \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} < 0 \right\} \end{aligned} \quad (21)$$

where

$$[\varepsilon(t)]_i = \min\{\varepsilon, [\mathbf{w}(t)]_i\} \quad (22)$$

ε is a fixed positive scalar (typically small) and $[\mathbf{w}(t)]_i$ is the i -th element in $\mathbf{w}(t)$, $\mathbf{w}(t)$ is given by

$$\mathbf{w}(t) = \left| \mathbf{q}^g(t) - [\mathbf{q}^g(t) - \mathbf{C}\nabla h(\mathbf{q}^g(t))]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g} \right| \quad (23)$$

where $|\cdot|$ denotes the element-wise absolute value operation, \mathbf{C} is a fixed diagonal positive definite matrix (e.g., the identity matrix). This is an anti-zigzagging procedure commonly employed in feasible direction methods (see [34]), and is designed to counteract the possible discontinuity exhibited by the set $\hat{\mathbf{I}}(t)$. Ideally, $\mathbf{D}(t)$ is not only diagonal with respect to $\hat{\mathbf{I}}(t)$, but rather with respect to the larger set $\mathbf{I}(t)$. Meanwhile, to achieve a fast convergence performance, $\mathbf{D}(t)$ should be an adequate approximation of the inverse Hessian \mathbf{H}^{-1} , at least along a suitable subspace. At this point, Proposition 3 introduces $\mathbf{D}(t)$ such that it is diagonal with respect to $\mathbf{I}(t)$ and the portion of $\mathbf{D}(t)$ corresponding to the indices $i \notin \mathbf{I}(t)$ is the inverse of the Hessian \mathbf{H} with respect to these indices.

Proposition 3: Suppose $\mathbf{D}(t)$ satisfies:

$$\mathbf{D}(t) = \mathbf{E}(t)^{-1} \quad (24)$$

where $\mathbf{E}(t)$ is:

$$[\mathbf{E}(t)]_{ij} = \begin{cases} 0 & \text{if } i \neq j, \text{ and either } i \text{ or } j \in \mathbf{I}(t) \\ |\mathbf{H}_{ii}| & \text{if } i = j \text{ and } i \in \mathbf{I}(t) \\ \mathbf{H}_{ij} & \text{otherwise} \end{cases} \quad (25)$$

Then, $\mathbf{D}(t)$ is a symmetric positive definite matrix, which is diagonal with respect to $\mathbf{I}(t)$.

Proof of Proposition 3: Without loss of generality, let γ denote the number of elements in $\mathbf{I}(t)$. Define $\tilde{\mathbf{E}}(t)$ as follows:

$$\tilde{\mathbf{E}}(t) = \begin{bmatrix} \mathbf{B}(t) & & & \\ & |\mathbf{H}_{I_1(t)I_1(t)}| & & \\ & & \ddots & \\ & & & |\mathbf{H}_{I_\gamma(t)I_\gamma(t)}| \end{bmatrix} \quad (26)$$

where $\mathbf{B}(t)$ is a $(m-\gamma) \times (m-\gamma)$ principal sub-matrix of \mathbf{H} formed by removing any i -th row and column for $i \in \mathbf{I}(t)$. And let $I_j(t)$ denote the j -th element in $\mathbf{I}(t)$, where $j = 1, 2, \dots, \gamma$.

Since \mathbf{H} is positive definite, it follows that $\mathbf{B}(t)$ is positive definite since any principal sub-matrix of positive definite

Algorithm 1 Offline PNM-based VVC

For any time step t : Alternately update variables by the following steps until convergence:
S1: Update $\nabla h(\mathbf{q}^g(t))$ by (16).
S2: Update $\mathbf{w}(t)$ by (23).
S3: Identify the set $\mathbf{I}(t)$ by (21) and (22).
S4: Update $\mathbf{D}(t)$ by (24) and (25), which is diagonal with respect to $\mathbf{I}(t)$.
S5: Update $\mathbf{q}^g(t+1)$. Let $\mathbf{u}(t) = \mathbf{D}(t)\nabla h(\mathbf{q}^g(t))$. Starting at $\tau(t) = 0$, repeat:

$$\begin{aligned}\tau(t) &= \tau(t) + 1, \alpha(t) = \beta^{\tau(t)} \\ \mathbf{q}^g(t+1) &= [\mathbf{q}^g(t) - \alpha(t)\mathbf{u}(t)]_{\mathbf{q}^g}^{\bar{\mathbf{q}}^g}\end{aligned}$$

until:

$$\begin{aligned}h(\mathbf{q}^g(t)) - h(\mathbf{q}^g(t+1)) &\geq \delta \left\{ \beta^{\tau(t)} \sum_{i \notin \mathbf{I}(t)} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [\mathbf{u}(t)]_i \right. \\ &\quad \left. + \sum_{i \in \mathbf{I}(t)} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [[\mathbf{q}^g(t)]_i - [\mathbf{q}^g(t+1)]_i] \right\}\end{aligned}$$

where $\beta \in (0, 1)$ and $\delta \in (0, 0.5)$.

S6: Update $t \leftarrow t + 1$.

matrix is positive definite. Thus, $\tilde{\mathbf{E}}(t)$ is a positive definite matrix. For any vector $\boldsymbol{\nu} \in \mathbb{R}^m$, we have:

$$\boldsymbol{\nu}^T \mathbf{E}(t) \boldsymbol{\nu} = \bar{\boldsymbol{\nu}}^T \tilde{\mathbf{E}}(t) \bar{\boldsymbol{\nu}} > 0 \quad (27)$$

where $\bar{\boldsymbol{\nu}}$ is a vector formed by rearranging elements in $\boldsymbol{\nu}$ in the order consistent with $\tilde{\mathbf{E}}(t)$, the inequality holds since $\tilde{\mathbf{E}}(t)$ is positive definite. Thus, $\mathbf{E}(t)$ is positive definite, implying $\mathbf{E}(t)$ is invertible. In addition, $\mathbf{E}(t)$ is symmetric since \mathbf{H} is symmetric. Then, we can conclude that $\mathbf{E}(t)$ is a symmetric positive definite matrix, it follows that $\mathbf{D}(t)$ is a symmetric positive definite matrix. Besides, from (24) and (25), it follows that:

$$[\mathbf{D}(t)]_{ij} = 0, \text{ if } i \neq j, \text{ and either } i \text{ or } j \in \mathbf{I}(t). \quad (28)$$

Then, $\mathbf{D}(t)$ is diagonal with respect to $\mathbf{I}(t)$. Q.E.D.

Proposition 4: Suppose $\mathbf{D}(t) = \mathbf{E}(t)^{-1}$, we have

- [P4.a] [P2.a] holds.
- [P4.b] [P2.b] holds.

Proof of Proposition 4: See Appendix A.

As indicated in Propositions 3 and 4, they extend $\hat{\mathbf{I}}(t)$ in Proposition 2 to the larger set $\mathbf{I}(t)$. Moreover, $\mathbf{D}(t) = \mathbf{E}(t)^{-1}$ shows a suitable approximation of the inverse Hessian to some extent, and a descent iteration can be obtained as $\mathbf{D}(t) = \mathbf{E}(t)^{-1}$ holds.

Finally, we propose an offline PNM-based VVC strategy, scaling the gradient with the nondiagonal symmetric positive definite matrix $\mathbf{D}(t) = \mathbf{E}(t)^{-1}$, instead of directly making use of the inverse Hessian \mathbf{H}^{-1} . The offline PNM-based VVC is presented in Algorithm 1. Regarding the choice of step size in S4 of Algorithm 1, it could be viewed as a combination of the Armijo-like rule and the Armijo rule. The convergence proof of PNM is provided in [28].

Remark 3: Gradient-based methods, e.g., GP, are first-order methods, they are typically characterized by the linear convergence rate. As discussed before, the second-order Newton's

method cannot be directly applied to solve constrained optimization problems. There are some extensions and variants [35]-[36] of Newton's method to solve constrained optimization problems, which are capable of superlinear convergence. However, sub-optimization problems are required to be solved in these methods [35]-[36] (which could be very time consuming for large-scale systems), thus rendering these methods impractical. In contrast, PNM can be utilized without solving any (sub)optimization problems, thereby reducing the associated computational burden. It is proved in [28, Prop. 4] that the superlinear convergence rate could be achieved for PNM under mild assumptions, thus leading to a faster convergence performance compared with first-order methods.

IV. ONLINE IMPLEMENTATION

To better capture the time-varying characteristic of system, the online implementation of our proposed voltage control scheme is proposed in this section. With respect to the online implementation, $\mathbf{c}(t)$, $\bar{\mathbf{q}}^g(t)$, and $\underline{\mathbf{q}}^g(t)$ do not need to be fixed, which could be time-varying.

A. Estimation of $\mathbf{c}(t)$

$\mathbf{c}(t)$ is a function of $\mathbf{v}_0(t)$, $\mathbf{p}(t)$, $\mathbf{q}^c(t)$, where $\mathbf{c}(t)$ can be calculated by (8) for the online implementation.

B. Estimation of VAr Limits

VAr limits $\bar{\mathbf{q}}^g(t)$ and $\underline{\mathbf{q}}^g(t)$ of DERs can be updated based on the inverter capacities and the instantaneous real power output of DERs. The online update of VAr limits is helpful to guarantee inverters of DERs operate within a secure range, especially for preventing inverters from overload.

C. Estimation of Gradient

For the online implementation, the rule to update the gradient $\nabla h(\mathbf{q}^g(t)) = \mathbf{M}[\mathbf{v}(\mathbf{q}^g(t)) - \mathbf{v}_r]$ as shown in (16) can be replaced by a feedback control law based on the voltage measurement, which is:

$$\nabla h(\mathbf{q}^g(t)) = \mathbf{M}[\mathbf{v}^m(t) - \mathbf{v}_r] \quad (29)$$

where $\mathbf{v}^m(t)$ is the squared voltage measurement vector for time step t . It indicates the gradient $\nabla h(\mathbf{q}^g(t))$ can be estimated by utilizing the local voltage measurement from each bus agent in the online implementation.

D. Online PNM-based VVC

The details and framework of the online PNM-based VVC are provided in Algorithm 2 and Fig.2, respectively. As seen in Algorithm 2 and Fig.2, each bus agent i sends its voltage measurement $\mathbf{v}_i^m(t)$ and VAr limits to the central agent, and the central agent communicates the VAr output command $\mathbf{q}_i^g(t+1)$ back to each bus agent i . Note that there is no any (sub)optimization problem required to solve for both the central agent and local bus agent, the central agent only needs to implement the analytical calculation for each time step, such characteristic is suitable for online implementation. Moreover,

Algorithm 2 Online PNM-based VVC

For any time step t , repeat the following steps:

- S1:** Each bus agent i measures voltage magnitudes and updates $\mathbf{v}_i^m(t)$, and estimates the VAr limits $\bar{\mathbf{q}}_i^g(t+1)$ and $\underline{\mathbf{q}}_i^g(t+1)$.
- S2:** Each bus agent i sends $\mathbf{v}_i^m(t)$, $\bar{\mathbf{q}}_i^g(t+1)$ and $\underline{\mathbf{q}}_i^g(t+1)$ to the central agent.
- S3:** The central agent estimates $\mathbf{c}(t)$, $\nabla h(\mathbf{q}^g(t))$ by (8) and (29).
- S4:** The central agent updates the VAr limits by setting $\bar{\mathbf{q}}^g = \bar{\mathbf{q}}^g(t+1)$, $\underline{\mathbf{q}}^g = \underline{\mathbf{q}}^g(t+1)$.
- S5:** The central agent updates $\mathbf{w}(t)$ by (23).
- S6:** The central agent identifies the set $I(t)$ by (21) and (22).
- S7:** The central agent updates $D(t)$ by (24) and (25), which is diagonal with respect to $I(t)$.
- S8:** The central agent updates $\mathbf{q}^g(t+1)$. Let $\mathbf{u}(t) = D(t)\nabla h(\mathbf{q}^g(t))$. Starting at $\tau(t) = 0$, repeat:

$$\begin{aligned}\tau(t) &= \tau(t) + 1, \alpha(t) = \beta^{\tau(t)} \\ \mathbf{q}^g(t+1) &= [\mathbf{q}^g(t) - \alpha(t)\mathbf{u}(t)]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g}\end{aligned}$$

until:

$$\begin{aligned}h(\mathbf{q}^g(t)) - \hat{h}(\mathbf{q}^g(t+1)) &\geq \delta \left\{ \beta^{\tau(t)} \sum_{i \notin I(t)} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [\mathbf{u}(t)]_i \right. \\ &\quad \left. + \sum_{i \in I(t)} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [[\mathbf{q}^g(t)]_i - [\mathbf{q}^g(t+1)]_i] \right\}\end{aligned}$$

where $\hat{h}(\mathbf{q}^g(t+1)) = \frac{1}{2} \|\mathbf{M}\mathbf{q}^g(t+1) + \mathbf{c}(t) - \mathbf{v}_r\|_2^2$, $\beta \in (0, 1)$ and $\delta \in (0, 0.5)$. Note that $\mathbf{c}(t)$ is used in $\hat{h}(\mathbf{q}^g(t+1))$ to keep consistent with $h(\mathbf{q}^g(t))$. Then the central agent sends $\mathbf{q}^g(t+1)$ to bus agent i .

S9: Update $t \leftarrow t + 1$.

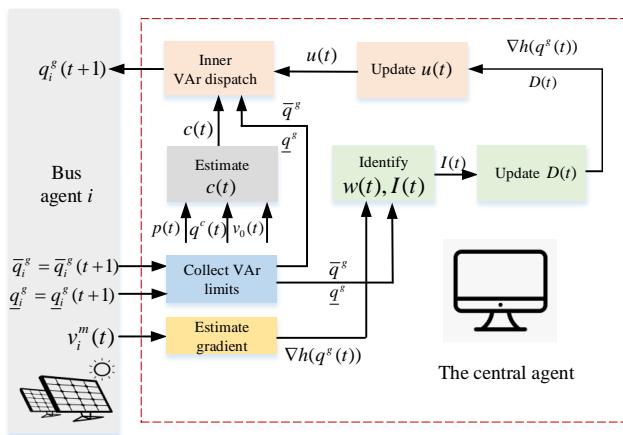


Fig. 2. Online PNM-based VVC framework.

the fast convergence performance of PNM contributes to a better capability to track the changes in distribution networks.

Remark 4: The key feature of online implementation is its own closed-loop nature, exploiting the most up-to-date voltage measurements to estimate ∇h . Though we design and analyze the algorithm under a fixed-point condition, the online implementation could asymptotically mitigate the model errors due to the closed-loop nature. This might shed light on why the online implementation can achieve a better performance than the offline implementation.

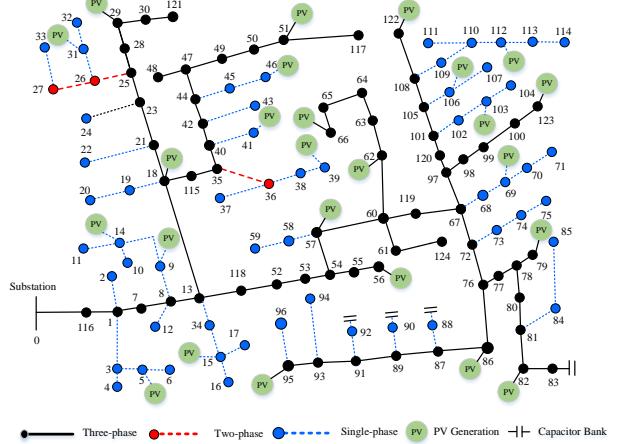


Fig. 3. IEEE 123-bus test system.

TABLE II
OBJECTIVE VALUES AND CONVERGENCE ITERATIONS COMPUTED BY DIFFERENT VVC STRATEGIES.

Strategy	Feedback	Convergence Iteration	Value
GP-based VVC	Yes	46	0.0013
DSGP-based VVC	Yes	25	0.0007
PNM-based VVC	Yes	5	0.0005
Offline OPF	No	converged	0.0015

V. CASE STUDY

A. Overview

In this section, numerical simulations are performed in the modified IEEE 123-bus test case [37] to demonstrate the proposed online voltage control scheme. As shown in Fig. 3, PV generators are distributed across the radial distribution network. Note that the proposed online PNM-based VVC strategy can be embedded into the two-layer VVC framework to coordinate the conventional discrete voltage regulation devices and DERs in different timescales. In the upper layer, the conventional discrete voltage regulation devices are scheduled over a slow timescale. In the lower layer, the VAr outputs of DERs can be adjusted by the online PNM-based VVC over a fast timescale. More details regarding the determination of the conventional devices in the upper layer are given in Appendix B.⁶ In the numerical simulations, the base voltage for the modified IEEE 123-bus network is 4.16 kV and the base power is 100 kVA. We set the reference of squared voltage magnitude as $\mathbf{v}_r = \mathbf{1}_m$, a $m \times 1$ column vector of ones. With respect to the PNM algorithm, the parameter settings are: C is the identity matrix, $\varepsilon = 0.001$, $\beta = 0.5$, $\delta = 0.1$.

To validate the effectiveness and superiority of the proposed online VVC strategy, both the static scenario and the dynamic scenario are taken into account in our simulations. The static scenario considers the static load and PV generators while the dynamic scenario considers the time-varying load and PV generators. All the simulations are conducted using MATLAB

⁶The operating status of on-line load tap changer and capacitor banks are determined in the upper layer.

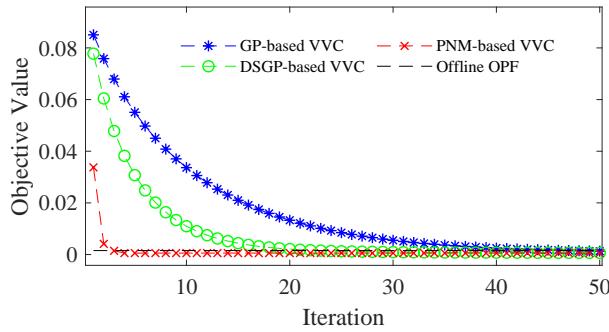


Fig. 4. Convergence performance of different VVC strategies.

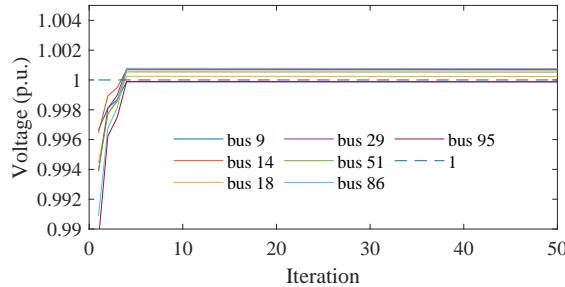


Fig. 5. Voltage magnitudes of buses 9, 14, 18, 29, 51, 86, 95 in the phase-a network for the online PNM-based VVC under the static scenario.

R2019b, the open source simulator OpenDSS [33] and the IBM ILOG CPLEX 12.9 solver [38].

B. Static Scenario

In the static scenario, each phase node has a constant load of $(6+j3)$ kVA, and each phase of PV generators supplies 20 kW real power generation to the grid. We also assume each phase of PV generator can supply or consume at most 50 kVar reactive power. For the static scenario, we consider the following four different VVC strategies:

(i) GP-based VVC ([20]): As shown in (15), the GP-based VVC replies on the gradient information and projection operation to adjust the VAr outputs of DERs.

(ii) DSGP-based VVC ([24],[25]): In the DSGP-based VVC, the diagonal entries of the Hessian matrix are leveraged to scale the gradient, thus improving the convergence performance. It adopts the scaled gradient and projection operation to adjust the VAr outputs of DERs.

(iii) PNM-based VVC: As shown in Algorithm 2, it leverages the projected Newton method to update the VAr outputs of DERs in the online manner.

(iv) Offline OPF: The problem (11) is directly solved by the CPLEX solver [38] without considering the feedback of voltage measurements.

Note that voltage magnitudes, solved by the nonlinear distribution power flow in OpenDSS, are regarded as the actual voltage measurements to update the VAr outputs of PV generators in the GP-based VVC, DSGP-based VVC, and PNM-based VVC. Table II and Fig.4 show objective values and convergence iterations computed by different VVC strategies. It can be observed that the GP-based VVC, DSGP-based VVC

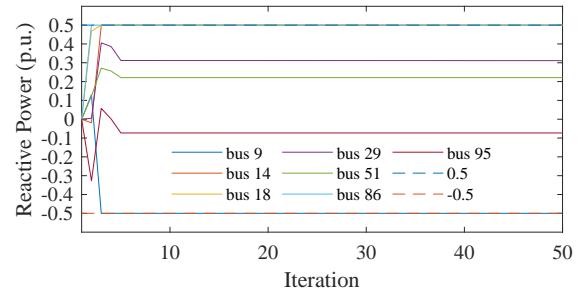


Fig. 6. VAr outputs of buses 9, 14, 18, 29, 51, 86, 95 in the phase-a network for the online PNM-based VVC under the static scenario.

and PNM-based VVC have smaller objective values than the OPF method due to the fact that they are able to asymptotically compensate the model errors via the feedback mechanism. The GP-based VVC shows a slow convergence rate since the convergence performance of the convention gradient-based method (steepest gradient) is not good. The performance of DSGP-based VVC is better than the GP-based VVC by taking advantage of the diagonal scaling. The convergence iteration of PNM-based VVC is only 5, which is far less than GP-based VVC and DSGP-based VVC, validating our previous analysis. This implies the online PNM-based VVC will have less communication costs in real-life implementations. Fig.5 and Fig.6 show that voltage magnitudes and VAr outputs of buses 9, 14, 18, 29, 51, 86, 95 in the phase-a network as the PNM-based VVC is applied. As depicted in Fig.5 and Fig.6, the voltage magnitudes are close to the target voltage 1.0 without violating the VAr limits by utilizing the PNM-based VVC.

C. Dynamic Scenario

For the dynamic scenario, we consider a more realistic time-varying system. The aggregate load and PV generation across the modified IEEE 123-bus test case are shown in Fig.7, where the total time span is one day (24 hours) and the time resolution is 10s.

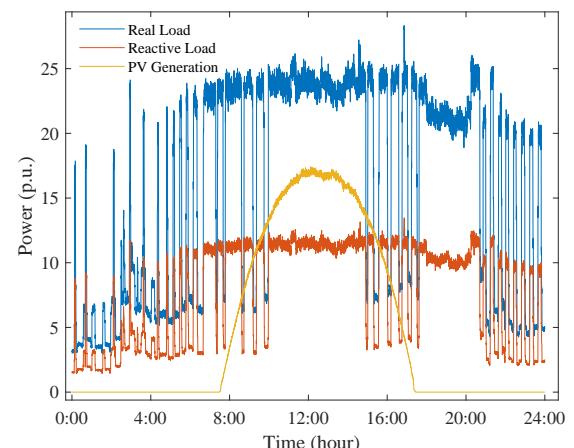


Fig. 7. Aggregate load and PV generation across the IEEE 123-bus test system.

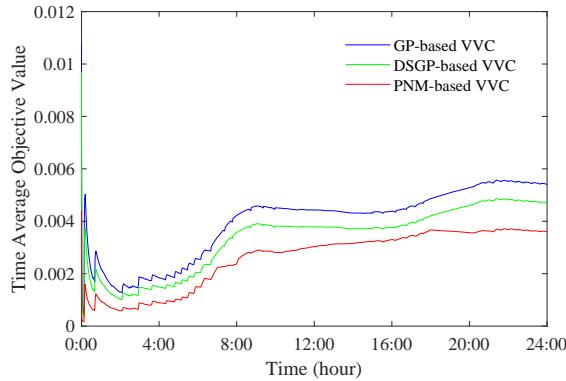


Fig. 8. Time average objective values with different VVC strategies under the dynamic scenario.

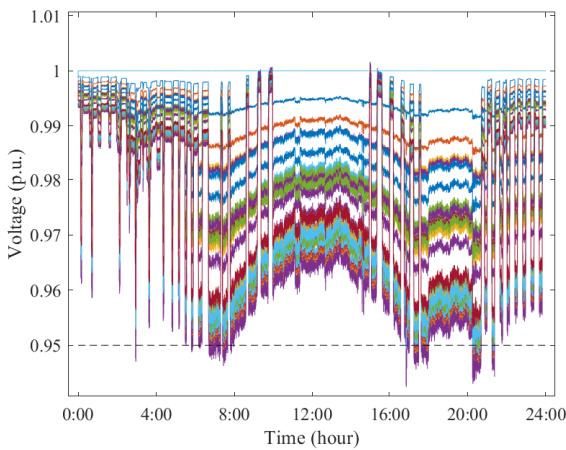


Fig. 9. Voltage profiles without control in the phase-a network (each curve depicts the voltage magnitude fluctuation for each phase-a node of bus) under the dynamic scenario.

In the dynamic scenario, three different VVC strategies, including GP-based, DSGP-based, PNM-based VVC, are taken into account⁷. The control period of those VVC strategies is set as 2s, i.e., the VAr outputs of PV inverters are updated every 2 seconds. In the dynamic scenario, the capacities of PV inverters are set as 50 kVA. And the VAr limits $\underline{q}^g, \bar{q}^g$ are updated online based on the given inverter capacities and the instantaneous real power of PV generators. Time average objective values across time steps are chose as the index to test the performance of those different VVC strategies in the dynamic scenario. Fig.8 shows the time average objective values across time steps computed by different VVC strategies. It can be observed from Fig.8 that the PNM-based VVC exhibits a better performance compared with GP-based VVC and DSGP-based VVC. Due to the fast convergence of PNM, the online PNM-based VVC has a great tracking capability to follow the time-varying changes in the system, thus leading to the better performance of PNM-based VVC compared with GP-based and DSGP-based VVC.

Taking phase a as an example, Fig.9 and Fig.10 show volt-

⁷Considering the offline OPF can only be implemented after convergence, it might not be suitable to apply the offline OPF to the dynamic scenario with a fast change rate. Thus, it is not carried out in the dynamic scenario.

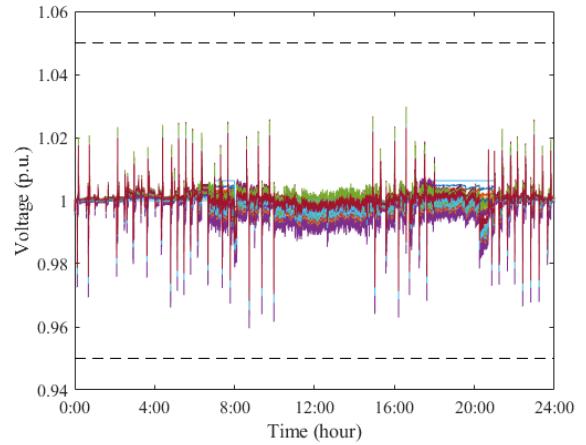


Fig. 10. Voltage profiles with the online PNM-based VVC in the phase-a network (each curve depicts the voltage magnitude fluctuation for each phase-a node of bus) under the dynamic scenario.

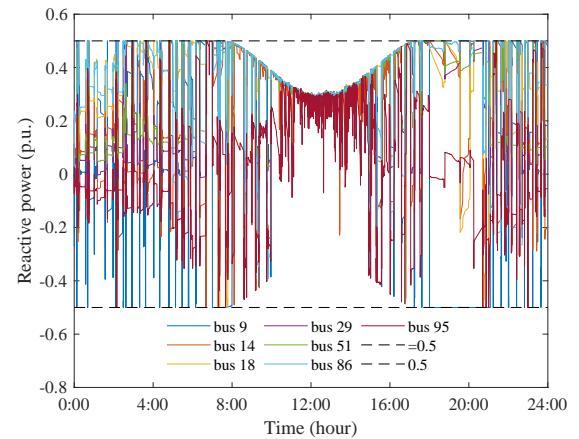


Fig. 11. VAr outputs of buses 9, 14, 18, 29, 51, 86, 95 in the phase-a network for the online PNM-based VVC under the dynamic scenario.

age profiles without control and with the online PNM-based VVC in the phase-a network, respectively. In addition, Fig.11 exhibits the VAr outputs of buses 9, 14, 18, 29, 51, 86, 95 under the dynamic scenario. It can be seen that there are under-voltage violations (below 0.95 p.u.) around 8:00, 16:00 and 20:00 for the distribution network without control due to the high load demands but low PV generation around those periods. In contrast, there is no any voltage violation across one day as the online PNM-based VVC is applied to the distribution network. It means the proposed online PNM-based VVC can effectively eliminate voltage violations in the dynamic scenario.

D. Algorithm Scalability: IEEE 8500-node Test Feeder

The IEEE 8500-node test feeder [39] is used to test the scalability of the online PNM-based VVC strategy. The test feeder is modified by adding several PV generators, as shown in Fig.12, where the capacities of those PV inverters are set as 100 kVA. The base voltage for the IEEE 8500-node test feeder is 7.2 kV and the base power is 100 kVA. The aggregate load

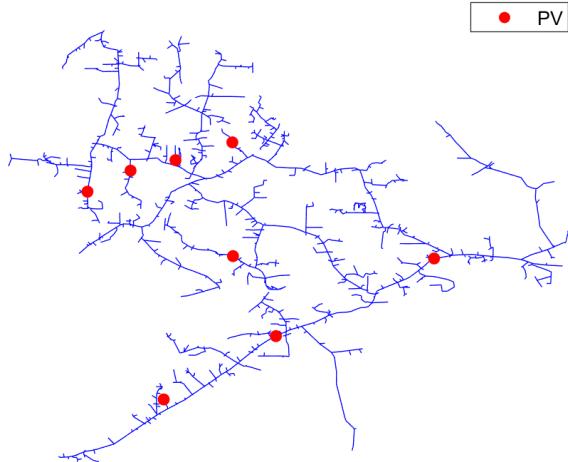


Fig. 12. IEEE 8500-node test feeder.

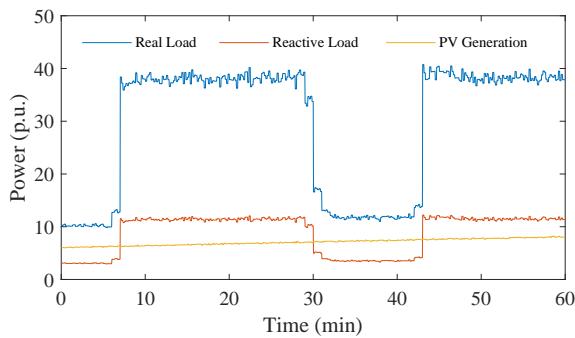


Fig. 13. Aggregate load and PV generation across the IEEE 8500-node test feeder.

and PV generation across the IEEE 8500-node test feeder are shown in Fig.13.

Fig. 14 reports minimum voltage profiles without control in phase-a, phase-b and phase-c networks for each time step. As seen in Fig. 14 , voltage violations occur in phase-a networks for this distribution network without control. Next, suppose one change is made: the proposed online PNM-based VVC is applied to manage the VAr outputs of PV generators. In this case, the maximum voltage magnitude across this IEEE 8500-node test feeder over time steps is 1.036 when implementing the online PNM-based VVC. In addition, Fig. 15 reports minimum voltage profiles with the online PNM-based VVC in phase-a, phase-b and phase-c networks for each time step. As seen in Fig. 15, no voltage violation occurs in phase-a, phase-b, and phase-c networks. These results illustrate the online PNM-based VVC is still effective to protect against voltage violations for this large distribution network, validating the algorithm scalability of our proposed online PNM-based VVC.

VI. CONCLUSION

This paper proposes an online voltage control method for unbalanced distribution networks where PNM is leveraged to improve the convergence performance. Through the feedback of voltage measurements from local bus agents, the central

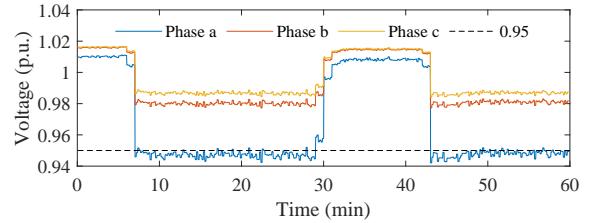


Fig. 14. Minimum voltage profiles without control in phase-a, phase-b and phase-c networks for each time step.

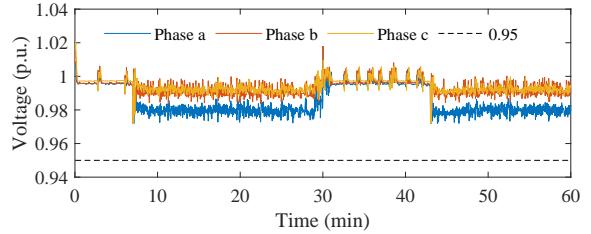


Fig. 15. Minimum voltage profiles with the online PNM-based VVC in phase-a, phase-b and phase-c networks for each time step.

agent can effectively coordinate the VAr outputs of DERs in real time to cope with fast system variations. The theoretical analysis and proof of this work are established on unbalanced distribution networks.

Numerical case studies are conducted in both static and dynamic scenarios to test the performance of the proposed online PNM-based VVC. Through numerical case studies, it is found that the online voltage control with feedback exhibits a better performance than offline OPF without feedback since the closed-loop nature of online feedback can asymptotically compensate the model errors via the feedback mechanism. Our proposed online PNM-based VVC can effectively eliminate the voltage violations in unbalanced distribution networks, and shows a better control performance compared with other methods (e.g., GP and DSGP), validating our theoretical analysis.

This work is the model-based VVC strategy, relying on the knowledge of distribution networks. However, it is promising to reduce the requirement for the knowledge of distribution networks due to the increasing deployment of advanced monitoring and metering infrastructures and available data in the distribution network level. Our future work will investigate data-driven model-free VVC strategies based on advanced machine learning techniques.

APPENDIX A

Proof of Proposition 4: Making use of [22, Prop. 2.1.2], it follows that \mathbf{q}^{g*} is the optimal solution of (11), if and only if:

$$(\mathbf{q}^g(t) - \mathbf{q}^{g*})^T \nabla h(\mathbf{q}^{g*}) \geq 0, \forall \mathbf{q}^g(t) \in [\underline{\mathbf{q}}^g, \bar{\mathbf{q}}^g] \quad (30)$$

Now, we show the sufficient and necessary conditions of (30) are:

$$\frac{\partial h(\mathbf{q}^{g*})}{\partial [\mathbf{q}^g]_i} \geq 0 \text{ if } [\mathbf{q}^{g*}]_i = [\underline{\mathbf{q}}^g]_i \quad (31a)$$

$$\frac{\partial h(\mathbf{q}^{g*})}{\partial [\mathbf{q}^g]_i} \leq 0 \text{ if } [\mathbf{q}^{g*}]_i = [\bar{\mathbf{q}}^g]_i \quad (31b)$$

$$\frac{\partial h(\mathbf{q}^{g*})}{\partial [\mathbf{q}^g]_i} = 0, \text{ if } [\underline{\mathbf{q}}^g]_i < [\mathbf{q}^{g*}]_i < [\bar{\mathbf{q}}^g]_i \quad (31c)$$

Sufficiency: As the condition (31) holds, it is straightforward that (30) holds.

Necessity: It will be established using proof by contrapositive. Suppose there exists a $[\mathbf{q}^{g*}]_i$ such that the condition (31) fails to hold, there always exists a $\mathbf{q}^g(t) \in [\underline{\mathbf{q}}^g, \bar{\mathbf{q}}^g]$ such that the condition (30) fails to hold. For example, assume (31a) does not hold, i.e., $\frac{\partial h(\mathbf{q}^{g*})}{\partial [\mathbf{q}^g]_i} < 0$ as $[\mathbf{q}^{g*}]_i = [\underline{\mathbf{q}}^g]_i$. Then, let all the elements in $\mathbf{q}^g(t)$ be the same as \mathbf{q}^{g*} except $[\mathbf{q}^g(t)]_i > [\underline{\mathbf{q}}^g]_i = [\mathbf{q}^{g*}]_i$. In this case, such a $\mathbf{q}^g(t)$ does not satisfy (30). Thus, we conclude that (31) is the necessary condition of (30).

From the above analysis, we can know that if $\mathbf{q}^g(t)$ is the optimal solution \mathbf{q}^{g*} , then (31) holds. It follows from (31) that [P4.a] holds.

We then prove [P4.b] holds. Note that with respect to $i \in \mathbf{I}(t)$, it follows from the definition of $\mathbf{I}(t)$ and Proposition 3 that:

$$[\mathbf{u}(t)]_i > 0 \text{ if } [\underline{\mathbf{q}}^g]_i \leq [\mathbf{q}^g(t)]_i \leq [\mathbf{q}^g]_i + [\varepsilon(t)]_i, i \in \mathbf{I}(t) \quad (32a)$$

$$[\mathbf{u}(t)]_i < 0 \text{ if } [\bar{\mathbf{q}}^g]_i - [\varepsilon(t)]_i \leq [\mathbf{q}^g(t)]_i \leq [\bar{\mathbf{q}}^g]_i, i \in \mathbf{I}(t) \quad (32b)$$

Consider the following sets of indices:

$$\begin{aligned} \mathbf{I}_1 = & \{i | [\bar{\mathbf{q}}^g]_i > [\mathbf{q}^g(t)]_i > [\underline{\mathbf{q}}^g]_i \text{ or } [\mathbf{q}^g(t)]_i = [\underline{\mathbf{q}}^g]_i, [\mathbf{u}(t)]_i < 0 \\ & \text{or } [\mathbf{q}^g(t)]_i = [\bar{\mathbf{q}}^g]_i, [\mathbf{u}(t)]_i > 0 \text{ for } i \in \{1, \dots, m\} \setminus \mathbf{I}(t)\} \end{aligned} \quad (33a)$$

$$\begin{aligned} \mathbf{I}_2 = & \{i | [\mathbf{q}^g(t)]_i = [\underline{\mathbf{q}}^g]_i, [\mathbf{u}(t)]_i \geq 0 \text{ or } [\mathbf{q}^g(t)]_i = [\bar{\mathbf{q}}^g]_i, \\ & [\mathbf{u}(t)]_i \leq 0 \text{ for } i \in \{1, \dots, m\} \setminus \mathbf{I}(t)\} \end{aligned} \quad (33b)$$

$$\mathbf{I}_3 = \{i | [\mathbf{q}^g(t)]_i = [\underline{\mathbf{q}}^g]_i \text{ or } [\mathbf{q}^g(t)]_i = [\bar{\mathbf{q}}^g]_i, \text{ for } i \in \mathbf{I}(t)\} \quad (33c)$$

$$\begin{aligned} \mathbf{I}_4 = & \{i | [\underline{\mathbf{q}}^g]_i < [\mathbf{q}^g(t)]_i \leq [\mathbf{q}^g]_i + [\varepsilon(t)]_i \\ & \text{or } [\bar{\mathbf{q}}^g]_i - [\varepsilon(t)]_i \leq [\mathbf{q}^g(t)]_i < [\bar{\mathbf{q}}^g]_i, \text{ for } i \in \mathbf{I}(t)\} \end{aligned} \quad (33d)$$

And set $\tilde{\alpha}$:

$$\tilde{\alpha} = \sup\{\alpha(k) | [\bar{\mathbf{q}}^g]_i \geq [\mathbf{q}^g(t)]_i - \alpha(t)[\mathbf{u}(t)]_i \geq [\underline{\mathbf{q}}^g]_i, \forall i \in \mathbf{I}_1 \cup \mathbf{I}_4\} \quad (34)$$

In view of the definitions of \mathbf{I}_1 and \mathbf{I}_4 as well as (32), we can know $\tilde{\alpha} > 0$. Define the vector $\bar{\mathbf{u}}(t)$ as follows:

$$[\bar{\mathbf{u}}(t)]_i = \begin{cases} [\mathbf{u}(t)]_i, i \in \mathbf{I}_1 \cup \mathbf{I}_4 \\ 0, i \in \mathbf{I}_2 \cup \mathbf{I}_3 \end{cases} \quad (35)$$

In view of (17) and (32)-(35), we have:

$$\begin{aligned} \mathbf{q}^g(t+1) &= [\mathbf{q}^g(t) - \alpha(t)\mathbf{u}(t)]_{\underline{\mathbf{q}}^g}^{\bar{\mathbf{q}}^g} \\ &= \mathbf{q}^g(t) - \alpha(t)\bar{\mathbf{u}}(t), \forall \alpha(t) \in (0, \tilde{\alpha}) \end{aligned} \quad (36)$$

In view of the definitions of \mathbf{I}_2 and $\mathbf{I}(t)$, it follows that:

$$\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} \leq 0, \text{ if } i \in \mathbf{I}_2, [\mathbf{q}^g(t)]_i = [\underline{\mathbf{q}}^g]_i \quad (37a)$$

$$\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} \geq 0, \text{ if } i \in \mathbf{I}_2, [\mathbf{q}^g(t)]_i = [\bar{\mathbf{q}}^g]_i \quad (37b)$$

Combining (37b) and (37), it follows that:

$$\sum_{i \in \mathbf{I}_2} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [\mathbf{u}(t)]_i \leq 0 \quad (38)$$

Using (35) and (38), it follows that:

$$\begin{aligned} \nabla h(\mathbf{q}^g(t))^T \bar{\mathbf{u}}(t) &= \sum_{i \in \mathbf{I}_1 \cup \mathbf{I}_4} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [\mathbf{u}(t)]_i \\ &\geq \sum_{i \in \mathbf{I}_1 \cup \mathbf{I}_2} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [\mathbf{u}(t)]_i + \sum_{i \in \mathbf{I}_4} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [\mathbf{u}(t)]_i \end{aligned} \quad (39)$$

Using (18) and the structure of $\mathbf{D}(t)$, and rearranging terms, we have:

$$\begin{aligned} &\sum_{i \in \mathbf{I}_1 \cup \mathbf{I}_2} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [\mathbf{u}(t)]_i + \sum_{i \in \mathbf{I}_4} \frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} [\mathbf{u}(t)]_i \\ &= \mathbf{y}(t)^T \mathbf{B}(t)^{-1} \mathbf{y}(t) + \sum_{i \in \mathbf{I}_4} |[\mathbf{H}]_{ii}|^{-1} \left[\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} \right]^2 \geq 0 \end{aligned} \quad (40)$$

where $\mathbf{B}(t)$ is a positive definite matrix, defined in (26), and $\mathbf{y}(t)$ is a column vector formed by rearranging $\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i}$ for $i \in \mathbf{I}_1 \cup \mathbf{I}_2$ in the order consistent with $\mathbf{B}(t)$.

From the definitions of \mathbf{I}_3 and $\mathbf{I}(t)$, we can know that for $i \in \mathbf{I}_3$, $\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i}$ satisfies (31). We next prove: as $\mathbf{q}^g(t)$ is not the optimal solution, there must exist a $\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} \neq 0$ for $i \in \{\mathbf{I}_1 \cup \mathbf{I}_2 \cup \mathbf{I}_4\}$, which can be proved by contradiction. Suppose $\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} = 0$ for $i \in \{\mathbf{I}_1 \cup \mathbf{I}_2 \cup \mathbf{I}_4\}$, then it follows that for $i \in \{\mathbf{I}_1 \cup \mathbf{I}_2 \cup \mathbf{I}_4\}$, $\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i}$ satisfies (31). Plus that for $\forall i \in \mathbf{I}_3$, $\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i}$ satisfies (31), it follows that $\mathbf{q}^g(t)$ is the optimal solution, resulting in the contradiction. Thus, we can know there must exist a $\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} \neq 0$ for $i \in \{\mathbf{I}_1 \cup \mathbf{I}_2 \cup \mathbf{I}_4\}$. Combing it with (40), it implies that:

$$\mathbf{y}(t)^T \mathbf{B}(t)^{-1} \mathbf{y}(t) + \sum_{i \in \mathbf{I}_4} |[\mathbf{H}]_{ii}|^{-1} \left[\frac{\partial h(\mathbf{q}^g(t))}{\partial [\mathbf{q}^g]_i} \right]^2 > 0 \quad (41)$$

It then follows that:

$$\nabla h(\mathbf{q}^g(t))^T \bar{\mathbf{u}}(t) > 0 \quad (42)$$

Combing this relation with (36) and the fact $\tilde{\alpha} > 0$, it follows that $\bar{\mathbf{u}}(t)$ is a feasible descent direction at $\mathbf{q}^g(t)$ and there exists a scalar $\bar{\alpha}(t)$ for which the desired relation [P4. b] is satisfied. Q.E.D.

APPENDIX B

The determination of the conventional devices in the upper layer

The voltage control devices in distribution networks exhibit different temporal characteristics, which should be addressed in the voltage control system design. The conventional discrete devices with slow and discrete nature should be operated in the slow timescale. Frequent adjustments of those discrete devices are not beneficial for the economic operation. However, the inverter-based DERs, e.g., PV, with fast and continuous nature could effectively respond to voltage issues with a fast timescale. Consequently, those voltage regulation devices should be operated in different timescales.

The two-layer VVC framework is depicted in Fig.16. In the upper layer, a model predictive control (MPC)-based control scheme is proposed to optimally coordinate the operation of conventional discrete devices, including on-load tap changer

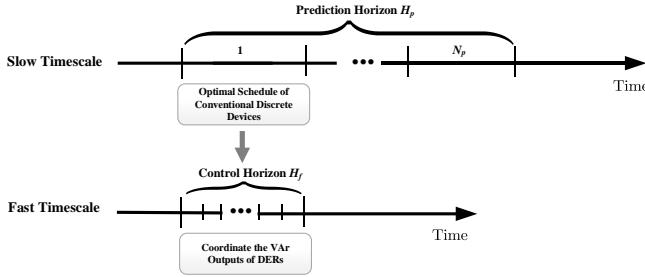


Fig. 16. Two-layer Volt/VAr control framework.

(OLTC) and capacitor banks (CBs), over slow timescale. Although the VAr outputs of DERs are also considered in the upper layer, only conventional discrete devices are scheduled in this upper layer.

MPC is widely employed to make optimal operation decisions in over a given time horizon. Optimal operation decisions over this given time horizon are generated, but only the decision for the first time step is implemented in practice. In the MPC, let H_p and $\mathcal{K}_p = \{1, \dots, N_p\}$ denote the prediction horizon and the prediction control steps, where T_p is the duration of each control period in the slow timescale and $N_p = H_p/T_p$.

Define the parameters and variables in the column vector form: let $\mathbf{n}_{tap}(t) = [n_{tap}^\phi(t)]_{\phi \in \Phi_0}, \Delta\mathbf{n}_{tap}(t) = [\Delta n_{tap}^\phi(t)]_{\phi \in \Phi_0} \in \mathbb{Z}^3$ denote the tap position and change of OLTC, $\mathbf{n}_{cb}(t) = [\mathbf{n}_{i,cb}(t)]_{i \in \mathcal{N}}, \Delta\mathbf{n}_{tap}(t) = [\Delta\mathbf{n}_{i,cb}(t)]_{i \in \mathcal{N}} \in \mathbb{Z}^m$ denote the number and switching times of CBs, $\mathbf{q}^l(t) = [\mathbf{q}_i^l(t)]_{i \in \mathcal{N}}, \mathbf{q}^{cb}(t) = [\mathbf{q}_i^{cb}(t)]_{i \in \mathcal{N}} \in \mathbb{R}^m$ denote the reactive load consumption power, reactive power generated by CBs, respectively. Let $\Delta\mathbf{q}^{cb} = [\Delta q_i^{cb}]_{i \in \mathcal{N}} \in \mathbb{R}^m$ denote capacitor bank unit reactive power output, and $\Delta tap = [\Delta^\phi tap]_{\phi \in \Phi_0} \in \mathbb{R}^3$ denote the tap step size of OLTC, where $\Delta^\phi tap$ is the tap step size of OLTC in phase ϕ . Then $v_0^\phi(t)$ can be expressed as follows:

$$v_0^\phi(t) = (1 + n_{tap}^\phi(t)\Delta^\phi tap)^2, \forall t \in \mathcal{K}_p \quad (43)$$

However, (43) could lead to a nonlinear and nonconvex mixed-integer optimization problem, making the optimization problem hard to solve. To this end, a linear approximation of (43) is derived as follows:

$$\begin{aligned} v_0^\phi(t) &= 1 + 2n_{tap}^\phi(t)\Delta^\phi tap + (n_{tap}^\phi(t)\Delta^\phi tap)^2 \\ &\approx 1 + 2n_{tap}^\phi(t)\Delta^\phi tap, \forall t \in \mathcal{K}_p \end{aligned} \quad (44)$$

Such an approximation is believed to hold since the term $(n_{tap}^\phi(t)\Delta^\phi tap)^2 \ll 1$, which can be written in a compact form:

$$\mathbf{v}_0(t) = \mathbf{1}_3 + 2\text{diag}(\mathbf{n}_{tap}(t))\Delta tap, \forall t \in \mathcal{K}_p \quad (45)$$

where $\mathbf{1}_3$ is a 3×1 column vector of ones.

Finally, in the upper layer, the MPC problem can be formulated as follows:

$$\begin{aligned} \min_{t \in \mathcal{K}_p} \sum_{t \in \mathcal{K}_p} \frac{1}{2} & \left(\|\mathbf{v}(t) - \mathbf{v}_r\|_{C_v}^2 + \|\mathbf{n}_{tap}(t) - \mathbf{n}_{tap}(t-1)\|_{C_{tap}}^2 \right. \\ & \left. + \|\mathbf{n}_{cb}(t) - \mathbf{n}_{cb}(t-1)\|_{C_{cb}}^2 \right) \end{aligned} \quad (46a)$$

$$\text{subject to: (5) and (45), } \forall t \in \mathcal{K}_p \quad (46b)$$

$$\mathbf{q}(t) = \mathbf{q}^c(t) - \mathbf{q}^g(t), \forall t \in \mathcal{K}_p \quad (46c)$$

$$\mathbf{q}^c(t) = \mathbf{q}^l(t) - \mathbf{q}^{cb}(t), \forall t \in \mathcal{K}_p \quad (46d)$$

$$\underline{\mathbf{q}}^g(t) \leq \mathbf{q}^g(t) \leq \bar{\mathbf{q}}^g(t), \forall t \in \mathcal{K}_p \quad (46e)$$

$$\mathbf{n}_{tap} \leq \mathbf{n}_{tap}(t) \leq \bar{\mathbf{n}}_{tap}, \forall t \in \mathcal{K}_p \quad (46f)$$

$$\Delta\mathbf{n}_{tap} \leq \Delta\mathbf{n}_{tap}(t) \leq \Delta\bar{\mathbf{n}}_{tap}, \forall t \in \mathcal{K}_p \quad (46g)$$

$$\mathbf{0} \leq \mathbf{n}_{cb}(t) \leq \bar{\mathbf{n}}_{cb}, \forall t \in \mathcal{K}_p \quad (46h)$$

$$\Delta\mathbf{n}_{cb} \leq \Delta\mathbf{n}_{cb}(t) \leq \Delta\bar{\mathbf{n}}_{cb}, \forall t \in \mathcal{K}_p \quad (46i)$$

$$\mathbf{q}^{cb}(t) = \text{diag}(\mathbf{n}_{cb}(t))\Delta\mathbf{q}^{cb}, \forall t \in \mathcal{K}_p \quad (46j)$$

In (46a), \mathbf{v}_r is the reference of squared voltage magnitude, C_v , C_{tap} , and C_{cb} are the weighting matrices corresponding to the three terms. (46b) denotes the power flow constraints and the constraints of head bus 0; (46c)-(46d) are the reactive power relationships; (46e) is the VAr limits for DERs; (46f) denotes the OLTC tap position limits; (46g) constrains the OLTC tap change; (46h) and (46i) constrain the maximum number and switching times of CBs; (46j) obtains the reactive power generated by CBs based on $\mathbf{n}_{cb}(t)$. The above problem (46) is a mixed-integer convex programming (MICP) problem and can thus be efficiently solved by MICP solvers. Only the optimal solution of OLTC and CBs in the first time step will be implemented.

REFERENCES

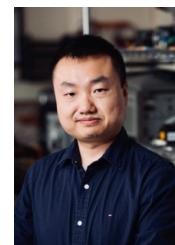
- [1] K. Turitsyn, P. Sulc, S. Backhaus and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proc. IEEE*, vol. 99, no. 6, pp. 1063-1073, June 2011.
- [2] M. Farivar, R. Neal, C. Clarke and S. Low, "Optimal inverter VAR control in distribution systems with high PV penetration," in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1-7.
- [3] S. Deshmukh, B. Natarajan and A. Pahwa, "Voltage/VAR control in distribution networks via reactive power injection through distributed generators," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1226-1234, Sept. 2012.
- [4] X. Zhou and S. Zou, P. Wang and Z. Ma, "Voltage regulation in constrained distribution networks by coordinating electric vehicle charging based on hierarchical ADMM", *IET Generation Transmission & Distribution*, vol. 14, pp. 3444-3457, 2020.
- [5] Y. Guo, H. Gao, H. Xing, Q. Wu and Z. Lin, "Decentralized coordinated voltage control for VSC-HVDC connected wind farms based on ADMM," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 800-810, April 2019.
- [6] E. Dall'Anese, S.V. Dhopde, B.B. Johnson and G.B. Giannakis, "Decentralized Optimal Dispatch of Photovoltaic Inverters in Residential Distribution Systems," *IEEE Transactions on Energy Conversion*, vol. 29, no. 4, pp. 957-967, Dec. 2014.
- [7] B. Zhang, A.Y.S. Lam, A.D. Domínguez-García and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1714-1726, July 2015.
- [8] B.A. Robbins and A.D. Domínguez-García, "Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2903-2913, Jul. 2016.
- [9] P. Šulc, S. Backhaus, and M. Chertkov, "Optimal distributed control of reactive power via the alternating direction method of multipliers," *IEEE Trans. Energy Convers.*, vol. 29, no. 4, pp. 968-977, Dec. 2014.
- [10] Q. Zhang, K. Dehghanpour and Z. Wang, "Distributed CVR in unbalanced distribution systems with PV penetration," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5308-5319, Sept. 2019.
- [11] W. Zheng, W. Wu, B. Zhang, H. Sun and Y. Liu, "A fully distributed reactive power optimization and control method for active distribution networks," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1021-1033, March 2016.
- [12] D. K. Molzahn, F. Dörfler, H. Sandberg, et al, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2941-2962, Nov. 2017.

- [13] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *Proc. 52nd IEEE Conf. Decis. Control*, 2013, pp. 4329–4334.
- [14] P. Jahangiri and D.C. Aliprantis, "Distributed Volt/VAR control by PV inverters," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3429–3439, Aug. 2013.
- [15] *IEEE standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces*, IEEE Standard 1547-2018, Feb. 15, 2018.
- [16] N. Li, G. Qu and M. Dahleh, "Real-time decentralized voltage control in distribution networks," in *Proc. 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2014, pp. 582-588.
- [17] S. Bolognani, R. Carli, G. Cavarro and S. Zampieri, "Distributed reactive power feedback control for voltage regulation and loss minimization," *IEEE Transactions on Automatic Control*, vol. 60, no. 4, pp. 966-981, April 2015.
- [18] J. Li, Z. Xu, J. Zhao and C. Zhang, "Distributed online voltage control in active distribution networks considering PV curtailment," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5519-5530, Oct. 2019.
- [19] G. Qu and N. Li, "Optimal distributed feedback voltage control under limited reactive power," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 315-331, Jan. 2020.
- [20] Y. Guo, H. Gao, D. Wang and Q. Wu, "Online optimal feedback voltage control of wind farms: decentralized and asynchronous implementations," *IEEE Trans. Sustain. Energy*, vol. 12, no. 2, pp. 1489-1492, April 2021.
- [21] H.J. Liu, W. Shi, and H. Zhu, "Distributed voltage control in distribution networks: online and robust implementations," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6106–6117, Nov. 2018.
- [22] D. P. Bertsekas, *Nonlinear Programming*, 1999, Second Edition, Athena Scientific, Belmont, MA, USA.
- [23] Z. Tang, D.J. Hill and T. Liu, "Fast distributed reactive power control for voltage regulation in distribution networks," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 802-805, Jan. 2019.
- [24] H. Zhu and H.J. Liu, "Fast local voltage control under limited reactive power: optimality and stability analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3794–3803, Sep. 2016.
- [25] W. Jiao, J. Chen, Q. Wu, S. Huang, C. Li and B. Zhou, "Distributed coordinated voltage control for distribution networks with DG and OLTC based on MPC and gradient projection," *IEEE Trans. Power Syst.*, in press, 2021.
- [26] D. P. Bertsekas, "On the Goldstein-Levitin-Polyak gradient projection method," *IEEE Transactions on Automatic Control*, vol. 21, no. 2, pp. 174-184, April 1976.
- [27] Y. Tang, K. Dvijotham and S.H. Low, "Real-time optimal power flow," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2963-2973, Nov. 2017.
- [28] D. P. Bertsekas, "Projected Newton methods for optimization problems with simple constraints," *SIAM J. Control Optim.*, vol. 20, no. 2, pp. 221–246, Mar. 1982.
- [29] L. Gan and S.H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial network," in *18th Power Systems Computation Conference (PSCC)*, 2014.
- [30] E. Dall'Anese, H. Zhu and G.B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464-1475, Sept. 2013.
- [31] M.E. Baran and F.F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Del.* 4(1), 725-734, Jan. 1989.
- [32] R. Cheng, Z. Wang, Y. Guo, "An online feedback-based linearized power flow model for unbalanced distribution networks," *IEEE Trans. Power Syst.*, in press, 2021.
- [33] R. C. Dugan and T. E. McDermott, "An open source platform for collaborating on smart grid research," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2011, pp. 1-7.
- [34] E. Polak, *Computational Methods in Optimization: A Unified Approach*, Academic press, 1971.
- [35] J. C. Dunn, "Newton's method and the Goldstein step-length rule for Constrained minimization problems," *SIAM Journal on Control and Optimization*, vol. 6, pp. 659-674, 1980.
- [36] U. M. Garcaí-Palomares, O. L. Mangasarian, "Superlinearly convergent quasi-newton algorithms for nonlinearly constrained optimization problems," *Mathematical Programming*, vol. 11, pp. 1-13, 1976.
- [37] W.H. Kersting, "Radial distribution test feeders," in *Proc. IEEE Power Eng. Soc. Win. Meeting*, 2001, pp. 908-912.
- [38] C. Bliek, P. Bonami, and A. Lodi, "Solving mixed-integer quadratic programming problems with IBM-CPLEX: a progress report," in *Proc. 26th RAMP Symposium*, Tokyo, Japan, Oct. 2014.
- [39] R.F. Arritt and R.C. Dugan, "The IEEE 8500-node test feeder," *IEEE PES T&D 2010*, 2010, pp. 1-6.



Rui Cheng (Graduate Student Member, IEEE) is currently a Ph.D. student in the Department of Electrical & Computer Engineering at Iowa State University. He received the B.S. degree in electrical engineering from Hangzhou Dianzi University in 2015 and the M.S. degree in electrical engineering from North China Electric Power University in 2018. From 2018 to 2019, he was an Electrical Engineer with State Grid Corporation of China, Hangzhou, China.

His research interests include power distribution systems, voltage/var control, transactive energy markets, and applications of optimization and machine learning methods to power systems.



Zhaoyu Wang (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiaotong University, and the M.S. and Ph.D. degrees in electrical and computer engineering from Georgia Institute of Technology. He is the Northrop Grumman Endowed Associate Professor with Iowa State University. His research interests include optimization and data analytics in power distribution systems and microgrids. He was the recipient of the National Science Foundation CAREER Award, the Society-Level Outstanding Young Engineer Award from IEEE Power and Energy Society (PES), the Northrop Grumman Endowment, College of Engineering's Early Achievement in Research Award, and the Harpole-Pentair Young Faculty Award Endowment. He is the Principal Investigator for a multitude of projects funded by the National Science Foundation, the Department of Energy, National Laboratories, PSERC, and Iowa Economic Development Authority. He is the Chair of IEEE PES PSOPE Award Subcommittee, the Co-Vice Chair of PES Distribution System Operation and Planning Subcommittee, and the Vice Chair of PES Task Force on Advances in Natural Disaster Mitigation Methods. He is an Associate Editor of IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE TRANSACTIONS ON SMART GRID, IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY, IEEE POWER ENGINEERING LETTERS, and IET Smart Grid.



Yifei Guo (M'19) received the B.E. and Ph. D. degrees in electrical engineering from Shandong University, Jinan, China, in 2014 and 2019, respectively. Currently, he is a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. He was a visiting student with the Department of Electrical Engineering, Technical University of Denmark, Lyngby, Denmark, in 2017–2018.

His research interests include voltage/var control, renewable energy integration, wind farm control, distribution system optimization and control, and power system protection.



Qianzhi Zhang (Graduate Student Member, IEEE) received the M.S. degree in electrical engineering from Arizona State University in 2015. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. He was a research engineer with Huadian Electric Power Research Institute from 2015 to 2016. He was a recipient of the Outstanding Reviewer Award from IEEE Transaction on Smart Grid, IEEE Transaction on Power Systems, and International Journal of Electrical Power and Energy Systems. His research interests include voltage/var control, power/energy management, system resilience enhancement, application of advanced optimization and machine learning techniques in power system operation and control.