Smart Inverter for Voltage Regulation: Physical and Market Implementation

Chenye Wu, *Member, IEEE*, Gabriela Hug, *Senior Member, IEEE*, and Soummya Kar, *Member, IEEE*

***Abstract*—Lack of efficient coordination schemes between photovoltaic (PV) panels may result in voltage stability issues. In this paper, we exploit the power control potential enabled by the PV inverters for voltage regulation. There are considerable obstacles to design a viable coordination scheme. Physically, the power flow equations lead to a highly nonconvex constraint set. With respect to the market implementation, voltage regulation is predominately conducted by the system operator, and voltage regulation related products rarely exist. We cast the voltage regulation as a nonconvex optimization problem, and devise an analytical framework to show that based on a linearized model, one can design a gradient descent-based distributed scheme, which, when implemented in the nonconvex branch flow model, will converge to a local minimum exponentially fast. Additionally, we design a compensation scheme which incentivizes the PV panel owners to provide voltage regulation. The compensation naturally leads to a game between all the PV panel owners. By design, the equilibria coincide with the global minimizers to the social planner problem. Simulation results confirm the convergence rate of the control actions in practice.**

***Index Terms*—Distributed algorithms, optimization methods, voltage control.**

I. Introduction

INVESTMENT in renewables today is mostly in utility-scale solar and wind plants, as well as small-scale distributed rooftop photovoltaics (PV). To date, the role of solar power is rather passive, not fully utilizing the intelligence available in the PV inverter - often referred to as the smart inverter. In this paper, we investigate the opportunities brought by these new components to the power grid.

A. Opportunities in Smart Inverter

The nomenclature of smart inverter highlights its capabilities, namely it has a digital architecture, bidirectional communication capability, and robust software and computation infrastructures. We imagine there are three possibilities to make better use of this currently under-utilized intelligence in the future grid. It can work with smart meters and conduct demand side management; it could enable the necessary infrastructure to support sharing the PV generation; it can also help tackle voltage stability issue. Our work focuses on this last opportunity and intends to provide more theoretical understanding on its potential.

More precisely, the smart inverter's intelligence in controlling the active and reactive power gives rise to its potential in conducting voltage regulation. This provides a means to tackle the voltage stability issue in distribution grids [1] imposed by the newly installed rooftop PV panels. To date, many distribution grid codes already require the PV inverters to conduct voltage regulation by specifying the minimum control response time for voltage control, power factor control, and reactive power control [2]. However, the distribution grid codes have not fully utilized the communication capabilities of the smart inverters. In this paper, we focus on utilizing these capabilities to enhance the voltage regulation performance. Also, to ensure the cooperation of PV panel owners, they need to be compensated for helping enhance the voltage stability. Many utilities have already conducted pilot projects to better exploit the voltage regulation potential in the smart inverter.

B. Challenges to Utilize the Inverter

A principal difficulty with voltage regulation lies in the inefficiency of reactive power transport: transmitting reactive power over long distance causes high active power losses. Therefore, voltage regulation and reactive power support are often conducted locally within the substation. However, even within the substation, the conventional distribution grid lacks voltage control and local measurement units to conduct accurate regulation. This situation has been changed dramatically over the past decades with the blossom of smart grid initiatives.

Distributed generation units with voltage control potential and phasor measurement units (PMUs) are increasingly being deployed in the distribution grids across the world. Yet, it is still a delicate task to utilize these new components due to two challenges: a) the lack of an efficient coordination scheme to utilize the potential of the new components; and b) the missing economic incentives for the PV owners to conduct voltage regulation. Hence, when devising a coordination approach, the economic design must incentivize the PV owners to actually follow the desired coordination scheme.

Manuscript received August 1, 2017; revised December 18, 2017 and April 25, 2017; accepted June 15, 2018. Date of publication July 11, 2018; date of current version October 18, 2018. Paper no. TPWRS-01162-2017. *(Corresponding author: Chenye Wu.)*

C. Wu is with the Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China (e-mail: chenyewu@tsinghua.edu.cn).

G. Hug is with the Power Systems Laboratory, ETH Zurich, 8092 Zürich, Switzerland (e-mail: ghug@ethz.ch).

S. Kar is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: soummyak@andrew.cmu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2018.2854903

C. Our Research Contributions

We study the specific problem of utilizing the PV inverters to conduct voltage regulation in the distribution grids. The principal contributions of this paper are:

- *Distributed Control Design:* We utilize the capability of the smart inverters to provide flexible active and reactive power for voltage regulation. Our focus is on coordinating all the PV panels to achieve the desired voltage profile in the distribution grid. To avoid the need for a centralized control entity, a distributed control scheme is proposed.
- *Role of Communication Network:* We examine the role of the communication network in designing the distributed control scheme. In particular, we propose the depth-k decomposition to investigate how to enhance the basic communication network (the one that has the same structure as the power network) to facilitate parallel control.
- *Convergence in Non-convex Model:* We devise an analytical framework to show based on a linearized model, one can design a gradient descent based distributed scheme, which, when implemented in the non-convex branch flow model, will converge to a local minimum exponentially fast. To the best of our knowledge, our work is the first attempt to analyze gradient type algorithms derived from the linearized model for the *non-convex* model.
- *Incentive Design:* To provide adequate incentives for PV owners to conduct voltage regulation, we discuss the necessary compensation structure, which leads to a game among the PV owners. We show that the equilibria coincide with the global minimizers of the social planner problem based on which the distributed coordination scheme is devised.

D. Related Works

One major body of related literature on distributed voltage regulation schemes investigates the resource coordination for the purpose of steady state scheduling. The power flow equations in the distribution network render the optimization or the control problem non-convex. Most works rely on linearizing the power flow equations around the operating point and then conduct the voltage regulation, including [3]–[7]. Other works convexify the problem by semidefinite programming (SDP) relaxation [8] or second-order cone programming (SOCP) relaxation [9], [10]. For example, in [11], Zhang *et al.* give sufficient conditions for the SDP relaxation to be tight for conducting voltage regulation in distribution grids.

Another research line focuses on the dynamics of voltage control (typically together with frequency control) in microgrids. For example, in [12], Schiffer *et al.* prove that a consensus-based distributed voltage control can uniquely determine the equilibrium point of the closed-loop voltage and reactive power dynamics. For more related works in this vein, see [13] for an excellent review.

Our work falls into the first category. While previous works illuminate the structure of the problem and the performance of distributed schemes in the convexified network model, the performance analysis in the non-convex network model has not been fully addressed. Farivar, Zhou, Chen *et al.* pioneered the line of research laying out the analytical distributed voltage regulation framework for linearized branch flow model [14]–[16]. However, different from this line of research, our work develops the analytical framework for the *non-convex* branch flow model. Hauswirth *et al.* investigate the convergence of gradient descent algorithm for the constrained optimization problem with a view to solving the optimal AC power flow in an online fashion in [17]. Different from using projection methods, we use the soft constraints to enforce the constraints in the original optimization problem. Recently, in [18], Jafarian *et al.* propose a price based distributed optimization approach to bridging the gap between the physical distribution grid and the electricity market. Different from this work, we use a potential game formulation to implement the market, which aims at encouraging PV panel owners to provide voltage regulation services. In this paper, we further the results in our recent work [19], where we study the *pure* reactive power control for voltage regulation in the branch flow model. Specifically, we discuss the joint control of active and reactive power and propose the necessary compensation structure to support the market implementation of the distributed control scheme.

The remainder of this paper is organized as follows. In Section II, we briefly review the branch flow model and its simplifications (linearization). Then, we investigate the distributed control schemes to conduct voltage regulation using the linearized branch flow model in Section III. In this section, we also highlight the role of communication network on the design. To show the performance of our proposed scheme for the branch flow model, we first analyze the performance of the scheme for the linearized model in Section IV. Based on this analysis, we then show the convergence guarantee for the non-convex model in Section V. We discuss the necessary compensation structure to incentivize PV owners in Section VI. We verify the performance of the scheme by simulation in Section VII. Finally, the concluding remarks and future directions are given in Section VIII. Fig. 1 visualizes the organization of the paper.

II. Network Model and Its Simplifications

A. Branch Flow Model

Consider a radial distribution system, represented by a graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$. Here, $\mathcal{N}$ is the set of nodes, indexed by $i = 0, \cdots, n$, where node 0 is the substation, and $\mathcal{E}$ stands for the set of distribution lines in the system. The radial structure allows us to conveniently define the direction of distribution lines. For each $(i, j) \in \mathcal{E}$, node j is the parent of node i. Note that each node i has only one parent, which allows for dropping j from the indexing.

For each line (i, j), denote its impedance by $z_i = r_i + \mathbf{i}x_i$; let I_i and $S_i = P_i + \mathbf{i}Q_i$ be the complex current and the complex power flowing from node i to j, respectively. At each node $i \in \mathcal{N}$, we denote its complex voltage by V_i, its complex power consumption and generation by $s_i^d = p_i^d + \mathbf{i}q_i^d$, and $s_i^g = p_i^g + \mathbf{i}q_i^g$, respectively.

The branch flow model [20] assumes a given and fixed voltage V_0 at the substation [21]. For notational simplicity, we define $l_i := |I_i|^2$, $v_i := |V_i|^2$. We also denote the child set of node i

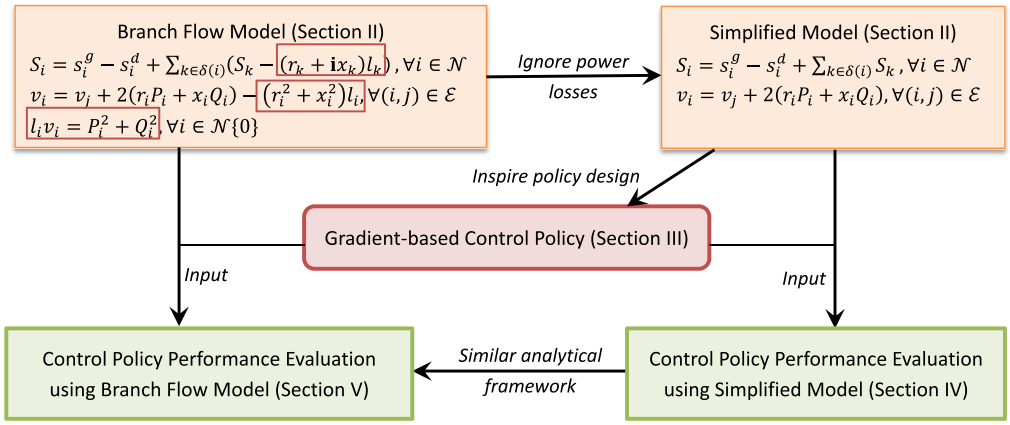


Fig. 1. Flowchart to visualize the organization of the paper.

by $\delta(i)$. These yield [20]:

$$S_i = s_i^g - s_i^d + \sum_{k\in\delta(i)} (S_k - (r_k + \mathbf{i}x_k)l_k), \forall i \in \mathcal{N}, \quad (1)$$

$$v_i = v_j + 2(r_iP_i + x_iQ_i) - (r_i^2 + x_i^2)l_i, \forall (i,j) \in \mathcal{E}, \quad (2)$$

$$l_iv_i = P_i^2 + Q_i^2, \forall i \in \mathcal{N}\backslash\{0\}. \quad (3)$$

Setting $S_0 = 0 + \mathbf{i}0$ enforces the power balance constraint at the substation, i.e., $s_0^g - s_0^d$ is the total net power injection into the distribution network from the main grid.

B. Simplified Branch Flow Model

In this paper, we first study the voltage regulation problem in a simplified network model to highlight the structure of the problem. In Section V, we will show the performance of the proposed control scheme for the non-convex branch flow model. We follow [22] and assume that the real and reactive power losses are much smaller than the power flows. This simplifies the branch flow model resulting in

$$S_i = s_i^g - s_i^d + \sum_{k\in\delta(i)} S_k, \forall i \in \mathcal{N}, \quad (4)$$

$$v_i = v_j + 2(r_iP_i + x_iQ_i), \forall (i,j) \in \mathcal{E}. \quad (5)$$

Remark: This simplified model inspires the following distributed control scheme design, provides us the structural insights into the control scheme and allows us to capture the convergence rate of our proposed control scheme in the subsequent analysis. Nonetheless, this simplification introduces minor approximation error, at the order of 1% [23].

III. Distributed Scheme Design

A. Problem Formulation

We conduct the voltage regulation by adjusting the active and reactive power through the PV inverters. We assume the voltage control is performed in real time in our framework. Hence, for each control interval, we assume the PV generation is known and can be measured directly. One natural objective function is to minimize the cost for the local system operator to maintain voltage stability. We assume the cost function is quadratic in the deviations of voltage profiles (v_i's) from their reference values. Another objective function is to maximize the PV owners' profit (denoted by $\theta_i(p_i^g)$ for PV owner at node i) by participating in demand side management or net-metering programs. For example, $\theta_i(p_i^g)$ could take the following form for net metering:

$$\theta_i(p_i^g) = \pi_{nm}(p_i^g - p_i^d). \quad (6)$$

This assumes that the electricity price and the net-metering price (π_{nm}) are the same. When $p_i^g \le p_i^d$, PV owner i still needs to pay for electricity. Otherwise, it will receive a net payment as a result from net metering. In this paper, we assume $\theta_i(p_i^g)$ takes this form.

Thus, the optimization problem may be cast as follows:

$$\begin{aligned} \min_{p_i^g, q_i^g} \quad & \sum_{i=1}^n \kappa \left(v_i - v_i^{ref}\right)^2 - \sum_{i=1}^n \theta_i(p_i^g) \\ \text{s.t.} \quad & (p_i^g)^2 + (q_i^g)^2 \le s_i^2, \forall i \in \mathcal{N}\backslash\{0\}, \\ & S_i = s_i^g - s_i^d + \sum_{k\in\delta(i)} S_k, \forall i \in \mathcal{N}, \\ & v_i = v_j + 2(r_iP_i + x_iQ_i), \forall (i,j) \in \mathcal{E}, \end{aligned} \quad (7)$$

where v_i^{ref} is the reference voltage at bus i. The first term in the objective function is the cost to maintain voltage stability and the second one stands for the total profits from participating in net-metering programs. The first constraint denotes the maximal apparent power constraint, and s_i is the maximal apparent power for PV panel i[1]. The last two constraints are the simplified power flow equations.

Remark: The parameters can be obtained from historical operation data. For example, the historical data will allow us to estimate κ, which captures the voltage deviation's amortized cost[2] on the distribution grid. Also, there could be additional cost for utilizing reactive power, although in practice, such cost is relatively small compared to the cost of using active power. As

[1] If there is no PV panel at bus i, then we can simply set $s_i = 0$.

[2] Here, we assume the distribution grid operator will keep track of the voltage deviations in the system. This is already possible thanks to the large scale deployment of PMUs.

long as the cost for utilizing reactive power can be modeled as a convex function, all of our subsequent conclusions will follow.

To further simplify the design of the control scheme, we introduce the following soft constraint to replace the maximal apparent power constraint:

$$f_i(p_i^g, q_i^g) = \begin{cases} +\infty, & \text{if } (p_i^g)^2 + (q_i^g)^2 > s_i^2, \\ \tan\left(\frac{(p_i^g)^2+(q_i^g)^2}{s_i^2} \cdot \frac{\pi}{2}\right), & \text{otherwise.} \end{cases} \tag{8}$$

Together with a small positive parameter β, we can reformulate the optimization problem (7) as follows:

$$\begin{aligned} \min_{p_i^g, q_i^g} \quad & \sum_{i=1}^{n} \kappa \left(v_i - v_i^{ref}\right)^2 - \sum_{i=1}^{n} \theta_i(p_i^g) \\ & + \beta \sum_{i=1}^{n} f_i(p_i^g, q_i^g) \\ \text{s.t.} \quad & S_i = s_i^g - s_i^d + \sum_{k \in \delta(i)} S_k, \forall i \in \mathcal{N}, \\ & v_i = v_j + 2(r_i P_i + x_i Q_i), \forall (i,j) \in \mathcal{E}. \end{aligned} \tag{9}$$

Note that, when $(p_i^g)^2 + (q_i^g)^2$ approaches s_i^2, the value of $f_i(p_i^g, q_i^g)$ will approach infinity. These soft constraints ensure that any feasible solution to (9) is a feasible solution to the original problem (7). As long as the selected parameter β is small enough, the solution to problem (9) can be very close to that to the original optimization.[3]

Remark: For simplicity, in this paper we intentionally choose the maximal apparent power constraint to represent the active and reactive power coupling. In practice, we need to consider more general constraints to capture the power factor requirement. These constraints can also be handled by including more soft constraints.

We will design the distributed control scheme based on problem (9) in the next section. Thereby we distinguish between three different communication network structures: 1) tree communication network, 2) fully connected network, and 3) general communication network structure.

B. Design With Tree Communication Network

Denote the unique path from node 0 to node i by $Path_i$. Define $\mathcal{P}_i$ to be $Path_i \backslash \{0\}$. Let w_i be the Lagrangian multiplier associated with the ith voltage constraint in (9). Denoting the Lagrangian function by $\mathcal{L}$, the Karush-Kuhn-Tucker (KKT) condition [25] for each voltage v_i is

$$\frac{\partial \mathcal{L}}{\partial v_i} = 2\kappa(v_i - v_i^{ref}) - w_i + \sum_{j \in \delta(i)} w_j = 0. \tag{10}$$

Using (10) and other KKT conditions, we can design the primal-dual subgradient method [25] as follows:

Scheme with Tree Communication Network

At each iteration t, $t = 1, 2, \cdots$, do the following:

Phase 1: from leaves to the root, sequentially compute the Lagrangian multipliers:

For node i, after receiving all its child(ren) information, it can compute its own multiplier

$$w_i^t = 2\kappa(v_i^t - v_i^{ref}) + \sum_{j \in \delta(i)} w_j^t.$$

It will send its parent its own multiplier as well as all information about its child(ren).

Phase 2: from root to leaves, sequentially exchange the multipliers and conduct the control: for node i, after receiving all its ancestor information, it can conduct the voltage regulation control for the next round

$$p_i^{g,t+1} = p_i^{g,t} - \alpha \left(\beta \frac{\partial f_i}{\partial p_i^g} - \frac{\partial \theta_i}{\partial p_i^g} + 2 \sum_{j \in \mathcal{P}_i} w_j^t r_j \right), \tag{11}$$

$$q_i^{g,t+1} = q_i^{g,t} - \alpha \left(\beta \frac{\partial f_i}{\partial q_i^g} + 2 \sum_{j \in \mathcal{P}_i} w_j^t x_j \right), \tag{12}$$

where α is the step size. Then, it passes its Lagrangian multiplier to all the nodes in its child set.

Stopping criteria: given tolerance $\eta > 0$,

$$|v_i^{t+1} - v_i^t| < \eta, \ \forall i \in \mathcal{N}.$$

Remark: Note that after the first phase, no nodes know their parent's information (multiplier). This is why we require two sequential processes in the distributed control. Since the optimization problem is strictly convex, this simple iterative distributed scheme is guaranteed to converge as long as a suitable α is selected [25]. Note that the soft constraints, $f(\cdot)$'s, retain feasibility at all iterations.

C. Design With Complete Communication Network

The distributed control scheme discussed above requires a tree communication network, which has exactly the same structure as the radial distribution network. If we have a different communication network, we may be able to devise a more efficient distributed control scheme.

For example, suppose we have a complete communication network. We can combine the two constraints in (9):

$$v_k = v_0 + \sum_{i=1}^{n} R_{ki}(p_i^g - p_i^d) + \sum_{i=1}^{n} X_{ki}(q_i^g - q_i^d), \tag{13}$$

where

$$R_{ki} := 2 \sum_{h \in \mathcal{P}_k \cap \mathcal{P}_i} r_h, \text{ and } X_{ki} := 2 \sum_{h \in \mathcal{P}_k \cap \mathcal{P}_i} x_h.$$

The two constraints in (9) show that to obtain v_k, it suffices to know the voltage profile of its parent and the flow between node k and its parent, while eq. (13) highlights the fact that it is also possible to obtain v_k directly from the flow information, if one has global information about the infeed and withdrawal at all the nodes. This inspires us to design the following distributed control scheme, which can be implemented in parallel.

[3]For more detailed discussion on the soft constraint implementation, please refer to [24].

Using (13), optimization problem (9) then becomes

$$\min_{p_i^g, q_i^g} \sum_{i=1}^{n} \kappa \left(v_i - v_i^{ref}\right)^2 - \sum_{i=1}^{n} \theta_i(p_i^g) + \beta \sum_{i=1}^{n} f_i(p_i^g, q_i^g)$$

$$\text{s.t. } v_k = v_0 + \sum_{i=1}^{n} R_{ki}(p_i^g - p_i^d) + \sum_{i=1}^{n} X_{ki}(q_i^g - q_i^d). \quad (14)$$

Note that problem (14) is not really a constrained optimization problem. The constraints simply define how v_k's are functions of the control actions $\mathbf{a} = (p_1^g, \cdots, p_n^g, q_1^g, \cdots, q_n^g)^T$, i.e., $v_i(\mathbf{a})$. In essence, problem (14) is an unconstrained optimization problem, for which the objective function is defined as

$$h(\mathbf{a}) = \sum_{i=1}^{n} \kappa \left(v_i(\mathbf{a}) - v_i^{ref}\right)^2 - \sum_{i=1}^{n} \theta_i(p_i^g) + \beta \sum_{i=1}^{n} f_i(p_i^g, q_i^g). \quad (15)$$

We want to emphasize that all the physical network information is hidden in R_{ki}'s and X_{ki}'s, and hence is hidden in the function $v_i(\mathbf{a})$. This means that problem (14) still captures the physical network information. As long as $h(\mathbf{a})$ is convex, the original optimization problem can be solved by the gradient descent algorithm. This is just a special case of the primal-dual subgradient algorithm. It leads to the following distributed control scheme:

Scheme with Complete Communication Network

At each iteration t, $t = 1, 2, \cdots$, do the following:

Phase 1: each node i measures the local voltage and computes

$$w_i^t = 2\kappa(v_i^t - v_i^{ref}).$$

Then, each node broadcasts the signal to all the other nodes.

Phase 2: each node will update its own reactive power control signal in a gradient descent way:

$$p_i^{g,t+1} = p_i^{g,t} - \alpha \left(\beta \frac{\partial f_i}{\partial p_i^g} - \frac{\partial \theta_i}{\partial p_i^g} + \sum_{k=1}^{n} w_k^t R_{ki}\right), \quad (16)$$

$$q_i^{g,t+1} = q_i^{g,t} - \alpha \left(\beta \frac{\partial f_i}{\partial q_i^g} + \sum_{k=1}^{n} w_k^t X_{ki}\right), \quad (17)$$

where α is the step size.

Stopping criteria: given tolerance $\eta > 0$,

$$|v_i^{t+1} - v_i^t| < \eta, \ \forall i \in \mathcal{N}.$$

Remark: This is a more efficient distributed control scheme compared to the scheme with the tree structure communication network. In this case, we do not need two rounds of sequential updates. All the nodes can compute the local Lagrangian multiplier by local voltage measurements. By exchanging the Lagrangian multipliers, all the nodes can perform the voltage regulation control, which will iteratively solve the problem.

D. Design With General Communication Network

As we have shown in the two extreme cases, the design of the control scheme relies on the manipulation of the two constraints in (9). This is also the case with a general communication structure. Define the depth d_k of a node k by the number of edges between itself and the root (node 0). Clearly, the depth of the root, d_0, is 0. Denote the set of nodes with depth at most k by D_k. That is, $D_k = \{i | d_i \leq k, i \in \mathcal{N}\}$. Thus, we can conduct the **depth-k decomposition** of the two constraints in (9) as

$$v_i = v_j + 2(r_i P_i + x_i Q_i), \ \forall i \in D_k, (i,j) \in \mathcal{E}, \quad (18)$$

$$S_i = s_i^g - s_i^d + \sum_{k \in \delta(i)} S_k, \ \forall i \in D_k, \quad (19)$$

$$v_j = v_{k(j)} + \sum_{i \in \bar{D}_k} R_{ij}^k (p_i^g - p_i^d) + \sum_{i \in \bar{D}_k} X_{ij}^k (q_i^g - q_i^d), \forall i \in \bar{D}_k, \quad (20)$$

where $k(j)$ is the single element in the set $\{i | i \in P_j, d_i = k\}$. That is, $k(j)$ is node j's unique ancestor of depth k. $\bar{D}_k$ is the complement set of D_k, i.e., $\bar{D}_k = \mathcal{N} \backslash D_k$. R_{ij}^k and X_{ij}^k are defined as follows:

$$R_{ij}^k := 2 \sum_{h \in \mathcal{P}_i \cap \mathcal{P}_j \cap \bar{D}_k} r_h, \quad (21)$$

$$X_{ij}^k := 2 \sum_{h \in \mathcal{P}_i \cap \mathcal{P}_j \cap \bar{D}_k} x_h. \quad (22)$$

Based on depth-k decomposition, when designing the primal-dual subgradient algorithm, the Lagrangian multipliers associated with the voltage constraints (denoted by w_i) will be updated according to the following KKT conditions:

$$w_i = 2\kappa(v_i - v_i^{ref}) + \sum_{j \in \delta(i)} w_j, \forall i \in D_{k-1}, \quad (23)$$

$$w_i = 2\kappa(v_i - v_i^{ref}) + \sum_{j \in \sigma(i)} w_j, \forall i \in D_k \cap \bar{D}_{k-1}, \quad (24)$$

$$w_i = 2\kappa(v_i - v_i^{ref}), \forall i \in \bar{D}_k, \quad (25)$$

where $\sigma(i) = \{j | j \in \mathcal{P}_i\}$. That is, $\sigma(i)$ is the set of all the successors to node i.

We want to exploit the necessary communication structure to enable the *sequential* updating rule in the upper level (from root to the nodes of depth k), and the *parallel* updating rule in the lower level (all the other nodes). For the purpose of calculating the Lagrangian multipliers, it suffices for the upper level to have the exact same network structure as the physical network while it requires additional communication links between i and $k(i)$ for the nodes in the lower level.

Next, we discuss how to update the control actions. Take the reactive power control signal as an example, the primal-dual subgradient algorithm requires:

$$q_i^{g,t+1} = q_i^{g,t} - \alpha \left(\beta \frac{\partial f_i}{\partial q_i^g} + 2 \sum_{j \in \mathcal{P}_i} w_j^t x_j\right), \forall i \in D_k, \quad (26)$$

$$q_i^{g,t+1} = q_i^{g,t} - \alpha \left(\beta \frac{\partial f_i}{\partial q_i^g} + \sum_{j \in \bar{D}_k} w_j^t X_{ij}^k\right), \forall i \in \bar{D}_k. \quad (27)$$

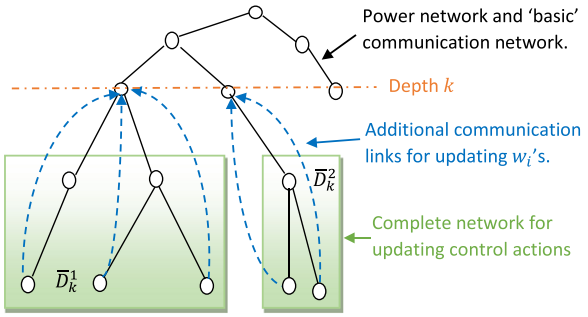


Fig. 2. Communication network requirement based on depth-k decomposition. 'Basic' communication network refers to the one exactly the same as the radial power network.

Again, as for updating the control signals, it suffices for nodes in the upper level to have the same communication network as the power network. However, it requires a complete communication network of several subsets in the lower level. Suppose the power network has μ nodes of depth k. Then, $\bar{D}_k$ can be divided into μ groups: $\bar{D}_k^1, \cdots, \bar{D}_k^\mu$. For each $\bar{D}_k^i$, all the nodes in the set share the same ancestor of depth k. The control signal updating process requires that the communication network within each $\bar{D}_k^i$ is complete. We visualize this design requirement in Fig. 2.

Remark: Fig. 3 summarizes the key differences among the three communication networks. We close this section by pointing out an interesting observation. The three distributed control schemes are different implementations of the same gradient descent algorithm. Each of them has its own advantage in utilizing a specific communication network. In Appendix A, we prove this observation. It also allows us to focus on the *complete communication network* case in the subsequent analysis.

IV. Performance in Linearized Model

The optimization problem (14) is convex due to the convex objective function and the linear constraint set. This implies convergence of the gradient descent algorithm. However, when it comes to the non-convex branch flow model, this conclusion no longer holds. In this section, we prove that the optimization problem (14) is strongly convex, which implies the exponentially fast convergence of the gradient descent algorithm. This serves as the basis of the analysis for the non-convex branch flow model.

A. Strong Convexity

Definition 1: A differentiable function f is called strongly convex with parameter $m > 0$ if the following inequality holds for all points x, y in its domain:

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq m\|x - y\|_2^2, \tag{28}$$

where $\|x\|_2$ is the $2-$ norm of vector x.

The last two terms in $h(\mathbf{a})$ are convex due to the convexity of linear functions and tangent functions. To show the strong convexity of $h(\mathbf{a})$, it suffices to show that its first term (for notational simplicity, denoted by $\lambda(\mathbf{a})$) is strongly convex. Standard mathematical manipulation yields (see Appendix B for the detailed derivation)

$$\begin{aligned}&(\nabla\lambda(\mathbf{a}^1) - \nabla\lambda(\mathbf{a}^2))^T(\mathbf{a}^1 - \mathbf{a}^2)\\ &= 2\kappa \sum_{k=1}^n \left(\sum_{i=1}^n R_{ki}(p_i^1 - p_i^2) + \sum_{i=1}^n X_{ki}(q_i^1 - q_i^2) \right)^2\\ &\geq \gamma \, \|\mathbf{a}^1 - \mathbf{a}^2\|_2^2, \end{aligned} \tag{29}$$

where

$$\gamma := 2\kappa \sum_{k=1}^n \min \left\{ \min_i R_{ki}, \min_i X_{ki} \right\}. \tag{30}$$

Note that, in the branch flow model, we assume that the voltage level at the substation is given and fixed. Therefore, if the root 0 has multiple children, they are naturally decoupled by this assumption. This allows us to focus on the case where root 0 has only *one* child. In this case, for all k, $\min_i R_{ki} \geq r_1 > 0$, $\min_i X_{ki} \geq x_1 > 0$. Thus, $\gamma \geq 2\kappa n \min\{r_1, x_1\} > 0$. Together with the convexity of the other two terms in $h(\mathbf{a})$, we prove the strong convexity of problem (14), i.e.,

$$(\nabla h(\mathbf{a}^1) - \nabla h(\mathbf{a}^2))^T (\mathbf{a}^1 - \mathbf{a}^2) \geq \gamma \|\mathbf{a}^1 - \mathbf{a}^2\|_2^2.$$

B. Convergence Rate

Theorem 2: The update rule $\mathbf{x}^{t+1} = \mathbf{x}^t - \alpha \mathbf{g}^t$ in the gradient descent algorithm (with step size α) to find the optimum $\mathbf{x}^*$ of strongly convex function $f(\mathbf{x})$ with parameter m, satisfies for any $t = 1, \cdots, T$,

$$\|\mathbf{x}^t - \mathbf{x}^*\|_2^2 \leq (1 - 2\alpha m)^{t-1} \|\mathbf{x}^1 - \mathbf{x}^*\|_2^2 + \frac{\alpha}{2m} \max_{1 \leq k \leq t} \|\mathbf{g}^k\|_2^2.$$

In particular, $\mathbf{x}^t$ converges to $\mathbf{x}^*$ exponentially fast with systematic error $\frac{\alpha}{2m} \max_{1 \leq k \leq t} \|\mathbf{g}^k\|_2^2$.

The exponential convergence rate sheds light on providing the performance guarantee in the non-convex branch flow model.

V. Performance Using Branch Flow Model

In Section IV, we have shown the convergence guarantee of the proposed gradient descent based control scheme when implemented in the linearized model. In this section, on the other hand, we will prove the convergence guarantee of our proposed distributed control scheme when implemented in the non-convex branch flow model.

Fig. 4 highlights the key connections between the implementation of the proposed scheme on the linearized model and that on the non-linear branch flow model. First, the distributed control scheme requires measuring the local voltage profiles from the real physical system. The measurement data will be used to compute the control signals according to the gradient descent approach, which will be again applied to the real physical system. Then, the physical system (and hence the voltage levels) will change according to the physical laws. The new measurement will reflect these physical dynamics, which have not been captured by the linearized model. Hence, the involvement of the physical system will challenge the convergence guarantee of our proposed scheme, when implemented in practice. In this section, we show that our distributed scheme will converge to some local optimum with bounded error.

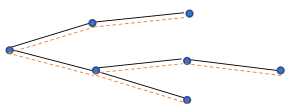
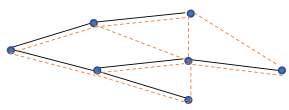
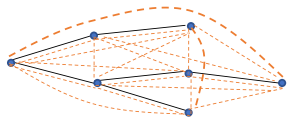
	Tree Network	General Network	Complete Network
Type of Communication Network			
Capability of Parallel Control	None	Partial Parallel	Fully Parallel
# of Communication in each iteration	$2\|\mathcal{E}\|$	$2\|D_k\| + \sum_{i=1}^{u} \|\overline{D}_k^{\mu}\|^2$	$\|\mathcal{E}\|^2$

Fig. 3. Features of three communication networks. Solid lines stand for the physical distribution lines and the dashed lines stand for the communication links.

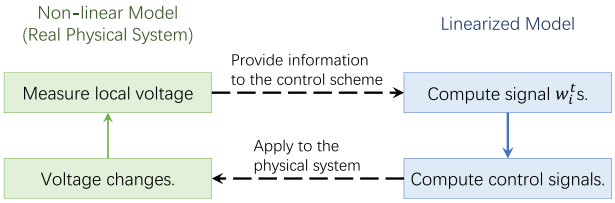


Fig. 4. Visualization of implementing the distributed control on real systems.

Note that the branch flow model defines a non-convex constraint set, which may lead to many local optima. In many cases, finding the global optimum can be very hard. Hence, one will turn to approach some local optimum. The major contribution of this paper is to theoretically prove the convergence rate as well as the bounded error between the convergent point and some local optimum. We will also compare the performance of the convergent point and that of the global optimum by simulations.

The most critical challenge in the convergence analysis is that in the branch flow model, the gradient is more complex than what we have used in the distributed control scheme in the linearized model. More importantly, in practice, even the branch flow model is an approximation. That is, if we use the measured voltage to compute the gradient, it is not the actual gradient for the optimization problem (14).

Denote the gradient for problem (14) by $\mathbf{g}^t$, and the actually implemented gradient by $\hat{\mathbf{g}}_i^t$, we have

$$\hat{\mathbf{g}}^t = \mathbf{g}^t + \xi^t, \tag{31}$$

where ξ^t takes into account all the disturbances and can be either positive or negative. With these, we can prove the following theorem:

Theorem 3: The update rule $\mathbf{a}^{t+1} = \mathbf{a}^t - \alpha\hat{\mathbf{g}}^t$ satisfies for any $t = 1, \ldots, T$,

$$\begin{aligned}&\|\mathbf{a}^t - \mathbf{a}^*\|_2^2\\ &\le (1-2\alpha\gamma)^{t-1}\|\mathbf{a}^1 - \mathbf{a}^*\|_2^2 + \frac{1}{\gamma}\max_{1\le k\le t}\|\xi^k\|\|\mathbf{a}^k - \mathbf{a}^*\|\\ &\quad + \frac{\alpha}{2\gamma}\max_{1\le k\le t}\|\mathbf{g}^k + \xi^k\|_2^2,\end{aligned}$$

where $\mathbf{a}^*$ is the unique global optimum to problem (14). In particular, $\mathbf{a}^t$ converges to $\mathbf{a}^*$ exponentially fast with systematic error $\frac{1}{\gamma}\max_k\|\xi^k\|\|\mathbf{a}^k - \mathbf{a}^*\| + \frac{\alpha}{2\gamma}\max_k\|\mathbf{g}^k + \xi^k\|_2^2$.

Remark: The systematic error can be very small. Note that $\gamma \ge 2\kappa n \min\{r_1, x_1\}$, which grows linearly in the number of nodes. Another important factor in determining this error is $\|\xi^t\|_2$. It has been shown in [6], [23] that under relatively flat voltage profile, the approximation error introduced by neglecting the power losses is about 0.25% (1%) if there is a 5% (10%) deviation in voltage magnitude. Bolognani *et al.* in [26] illustrate a way to bound the approximation errors for different power system linearizations, e.g., the approximation error between the LinDistFlow model [22] and the actual nonlinear physical equations. It is possible to use the same power flow manifold approach to bound ξ^t. However, a detailed analysis is beyond the scope of our paper.

The convergent point could be a local minimizer, a local maximizer, or a saddle point. In practice, one could select different starting points to achieve the desired local minimizer. In the simulation, we show that as long as the step size is not too small (no smaller than 10^{-5}), our scheme does converge to a local minimizer.

VI. Market Implementation

The success of any control scheme relies on a good business model. Central in the business model is the incentive design. In this section, we devise the necessary compensation structure to support our control scheme. The main idea is to conduct reverse engineering on problem (7) and design the compensation structure which will lead to a potential game [27]. One way to achieve this goal is to design the game such that the objective function in (7) is one potential function of the game.

Following this route, we design the following compensation: For each PV owner i, by conducting voltage regulation, it will receive:

$$\phi_i(p_i^g, q_i^g; p_{-i}^g, q_{-i}^g) = C_i - \kappa\sum_{j=1}^{n}(v_j - v_i^{ref})^2, \tag{32}$$

where C_i is the design parameter, p_{-i}^g and q_{-i}^g are the active and reactive power control actions of PV owners other than i, respectively; and the second term in (32) is the total cost for the local system operator to conduct voltage regulation. Note that, the simplified power flow equations define how v_k's are functions of the control action $\mathbf{a}$. This implies that the compensation relies not only on owner i's action, but also on all the other

players' actions. Together with the profits from net-metering, this compensation leads to the following game:

Voltage Regulation Game (VRG)

- *Players:* The set of $\mathcal{N}$.
- *Strategy Space:* For each player i, its strategy space A_i contains the maximal apparent power constraint, i.e., $A_i = \{(p_i^g, q_i^g) | (p_i^g)^2 + (q_i^g)^2 \leq s_i^2\}$.
- *Payoff Function:* For each player i, its payoff function is the sum of the voltage regulation compensation and the profit from net-metering, i.e.,

$$u_i(p_i^g, q_i^g; p_{-i}^g, q_{-i}^g) = \phi_i(p_i^g, q_i^g; p_{-i}^g, q_{-i}^g) + \theta_i(p_i^g).$$

To characterize the Nash equilibrium of the game, we first examine the potential function of the VRG. Define $A = \prod_{i\in\mathcal{N}} A_i$, and $A_{-i} = \prod_{j\in\mathcal{N}, j\neq i} A_j$.

Definition 4: A game is an ordinal potential game if there is a function $\Psi : A \to \mathbb{R}$, such that $\forall a_{-i} \in A_{-i}$, $\forall a_i^\dagger, a_i^\ddagger \in A_i$,

$$u_i(a_i^\dagger, a_{-i}) - u_i(a_i^\ddagger, a_{-i}) > 0 \Leftrightarrow \Psi(a_i^\dagger, a_{-i}) - \Psi(a_i^\ddagger, a_{-i}) > 0.$$

And the function Ψ is called a potential function of the game.

Intuitively, this definition ensures that in a potential game, when every player tries to maximize its own pay-off function, equivalently, it maximizes the potential function. In other words, although this is a non-cooperative game, the following interesting phenomenon happens: when one player is better off, everyone else is also better off. In our setting, this guarantees that every PV owner will try to maintain the voltage profile across the grid as close to the reference profile as possible. Formally, we can prove the following results.

Lemma 5: Define

$$\Psi = \sum_{i=1}^{n} \theta_i(p_i^g) - \sum_{i=1}^{n} \kappa \left(v_i - v_i^{ref}\right)^2.$$

Then, Ψ is a potential function of VRG.

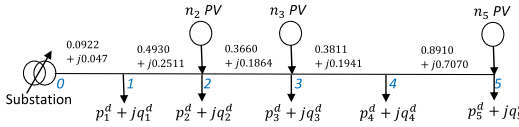
This is an immediate result from verifying the definition of potential function. The Lemma implies that the VRG is a potential game and enjoys all the nice properties of the general potential game.

Theorem 6: The VRG is an ordinal potential game. Any global minimizer to (7) is a pure Nash equilibrium of the VRG.

VII. Simulation Results

In this section, we use our proposed distributed control scheme for voltage control in various test systems. In particular, we show how the performance (e.g., convergence rate, voltage deviation, etc.) of the proposed scheme can be affected by the design parameters, the variation of demand due to PV penetration level, and electric vehicle (EV) penetration level.

In our simulations, we consider a prototype 6-bus test system, the IEEE 33-bus, and 123-bus distribution test systems. In each case, we select a subset of buses to install PV panels. We look at a variety of PV locations and PV penetration levels (number of installed PV panel at each bus) to understand the impact of PV generation on the voltage profiles. We assume that the maximal apparent power for each PV inverter is 7.2 kVA. To emulate the non-convex branch flow model for the distribution grid, we use the AC power flow solver in matpower [28]. We assume that



Bus Number		1	2	3	4	5
Scenario 1	p_i^d kW	100	90	120	60	60
	q_i^d kVA	60	10	80	30	20
	n_i # of PV	0	5	5	0	5
Scenario 2	p_i^d kW	300	270	360	180	180
	q_i^d kVA	180	30	240	90	60
	n_i # of PV	0	5	5	0	5
Scenario 3	p_i^d kW	50	45	-120	15	30
	q_i^d kVA	30	5	-80	7.5	10
	n_i # of PV	0	5	5	0	5
Scenario 4	p_i^d kW	100	90	-360	30	60
	q_i^d kVA	60	10	-240	15	20
	n_i # of PV	0	5	5	0	5

Fig. 5. Prototype 6-bus system, impedances are given in p.u.

we are given a complete communication network, and conduct all the case studies using Matlab 2017b on a Thinkpad TD 350 server (CPU: E5-3630V4 $\times$ 2, 64 GB RAM). We choose 1×10^{-5} for the soft constraint parameter β, 10 cents/kWh for the net metering price π_{nm}, and \$10/p.u. for voltage deviation cost κ. The stopping criteria η for the first small test system is set to be 10^{-9} while we set η to be 10^{-8} for the other relatively larger systems. Each PV generation is 5 kW.

A. Prototype 6-Bus Test System

This prototype 6-bus system uses the information of the first 6 buses in the IEEE 33-bus feeder system [29] (line characteristics shown in Fig. 5). This prototype system could allow us to better illustrate the impact of different factors on the voltage profiles. We assume the PV panels are located at buses 2, 3, and 5. Five PV panels are installed at each of the three selected buses. We are interested in four scenarios:

1) high PV penetration in modest loaded distribution grid: This is the standard situation. We follow the information of the first 6 buses in the IEEE 34-bus feeder system.
2) high PV penetration in high loaded distribution grid: We triple the load at each bus compared with scenario 1. This is possible due to the EV charging.
3) high PV penetration in light loaded distribution grid with modest counter flow: We modify the load parameters such that we create the counter flow in the distribution system while maintaining the whole distribution system is a net load. This describes the case when the EVs are discharging for demand side management (e.g., peak shaving for the whole grid).
4) high PV penetration in modest loaded distribution grid with strong counter flow: In this case, the whole distribution system is injecting power back to the grid.

These four scenarios try to capture the possible cases that the future distribution grid may evolve to due to the popularity of PV panels and electric vehicles. The performance of our proposed control scheme is demonstrated by Table I and Table II. It is not surprising to find that in the first two scenarios, the optimal control policy is to inject all the generated active power back to the grid and then inject as much reactive power as possible to the system to support the voltage profiles. In the third scenario, the voltage profiles are already close to 1. To

TABLE I
OPTIMAL CONTROL ACTION (KW OR KVA)

Action	p_2^g	q_2^g	p_3^g	q_3^g	p_5^g	q_5^g
Scenario 1	5	5.2	5	5.2	5	5.2
Scenario 2	5	5.2	5	5.2	5	5.2
Scenario 3	5	5.2	5	2	5	-1.1
Scenario 4	5	-5.2	0	-7.2	0	-7.2

TABLE II
OPTIMAL VOLTAGE PROFILE

Bus	1	2	3	4	5
Scenario 1	0.9998	0.9987	0.9981	0.9978	0.9974
Scenario 2	0.9991	0.9955	0.9935	0.9925	0.9913
Scenario 3	1.0000	1.0000	1.0001	0.9999	0.9998
Scenario 4	1.0000	1.0006	1.0013	1.0010	1.0006

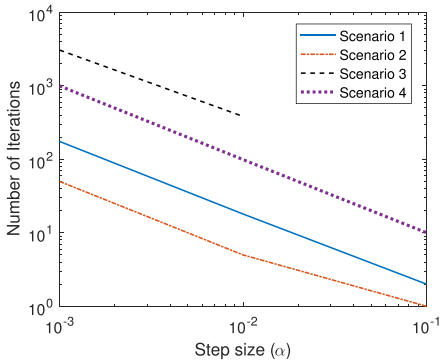


Fig. 6. Impact of step size on the convergence speed. Step size 0.1 does not lead to a convergent point in Scenario 3.

prevent from pushing them too high, the optimal control policy will not inject too much reactive power. Instead, at bus 5, the optimal control policy requires absorbing certain amount of reactive power from the system. In the last scenario, the optimal control policy will even stop injecting all the generated solar power back to the system. Instead, it asks the PV panels at bus 3 and bus 5 to absorb as much reactive power as possible. The optimality of the convergent point is guaranteed by observing the gradient error, which vanishes to zero after the initial several iterations.

The convergence rates for the four scenarios are visualized in Fig. 6. In the log-log plot, the number of iterations decreases linearly in the step size α. Our distributed control scheme converges very fast. For example, with a step size of 0.01, in conventional conditions (scenario 1 and 2), our scheme converges within 50 iterations; and in future conditions with counter flow (scenario 3 and 4), our scheme converges within 136 iterations.

B. IEEE 33-Bus Test System

We assume the PV panels are located at buses 2, 3, 4, 5, 6, 13, 19, 20, 25, 30. Some of these buses (e.g., bus 3 and 6) are critical points connecting two radial networks. Some of them (e.g., bus 25) are at the end of a radial network. Ten PV panels are installed at each of these buses. Following the four-scenario analysis for the 6-bus system, we scale the load parameters of two sets of buses to establish the four scenarios for this test system, as shown in Table III. We observe similar patterns in

TABLE III
ESTABLISH THE 4 SCENARIOS FOR IEEE 33-BUS SYSTEM

	scale for buses $\{8, 14, 19, 30\}$	scale for other buses
Scenario 1	1	1
Scenario 2	1.25	1.25
Scenario 3	-1	0.5
Scenario 4	-3	1

TABLE IV
CONVERGENCE RATE FOR IEEE 33-BUS SYSTEM

Step size α	0.01	0.001	0.0001
Scenario 1	2	18	177
Scenario 2	2	14	135
Scenario 3	12	114	1136
Scenario 4	14	136	1357

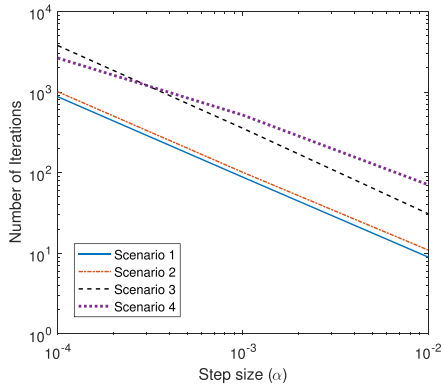


Fig. 7. Impact of step size on the convergence speed.

the optimal control policy for the four scenarios. Hence, we will only show the convergence rate results to demonstrate the scalability of our scheme. Table IV confirms in the log-log scale, the number of iterations decreases linearly in the step size. It is worth noting that selecting a step size of 0.01, our scheme converges faster in the 33-bus system than in the 6-bus system. This observation partially verifies our remark of Theorem 3, the systematic error vanishes as the system scales up.

C. IEEE 123-Bus Test System

To further verify the scalability of the proposed control scheme, we conduct the case study on the IEEE 123-bus distribution system. Note that this system is not a balanced three phase system. For simplicity, we use the first phase information for this study. We randomly choose 30 out of the total 123 buses, and assume 10 PV panels are installed at each of the 30 selected buses, which allows these buses to conduct voltage control. We again consider the four scenarios in this case, and use exactly the same scaling parameters as shown in Table III. Figure 7 plots the relationship between the convergent rate and the step size for this system. With a step size of 0.01, our control scheme converges within 80 steps for all the four scenarios (9 iterations for scenario 1; 11 for scenarios 2; 31 for scenario 3; and 71 for scenario 4). Hence, suppose the sampling

rate of the voltage profiles is 60 Hz[4] (the communication delay in a local network is often rather small), our control scheme can converge in around 1 second. This justifies our assumption that within the control interval, the PV generation is known and remains constant.

VIII. Conclusions

We examined the physical and the market implementation of utilizing the smart inverters for distributed voltage regulation. In particular, we highlight the theoretical convergence guarantee of our scheme in the non-convex branch flow model.

Additional work is necessary to enable a practical application of the proposed approaches, e.g., a systematic way to select the parameter in the compensation structure design. In fact, many other non-linear pricing schemes may also be able to achieve the same performance. Note that in this paper, we have implicitly assumed that within each time slot, the load remains constant. This assumption ignores the uncertainties in the load, as well as the uncertainties brought by the stochastic nature of the PV generation. We plan to employ the multi time step stochastic optimization to tackle these challenges. It is also very interesting to consider other components in the system with the same voltage control capabilities, e.g., EVs. This will allow the proposed distributed control scheme to handle more complicated cases (e.g., when there is not enough solar radiation).

Appendix

A. Proof for the Equivalence Observation

It suffices to show that the updating rules for the tree communication network and that for the complete communication network are the same. In particular, we will show that (11) and (16) are the same. Similarly, one can prove that (12) and (17) are the same.

We denote the Lagrangian multipliers for the tree network by w_j^T, and that for the complete network by w_j^C. Comparing (11) and (16), we only need to show that the following condition holds:

$$2\sum_{j\in\mathcal{P}_i} w_j^T r_j = \sum_{k=1}^{n} w_k^C R_{ki}, \forall i \in \mathcal{N}. \tag{33}$$

Based on the definition of w_j^T, we have

$$\begin{aligned} 2\sum_{j\in\mathcal{P}_i} w_j^T r_j &= 2\sum_{j\in\mathcal{P}_i} r_j \left(2\kappa\left(v_j - v_j^{ref}\right) + \sum_{u\in\delta(j)} w_u^T\right) \\ &= 2\sum_{j\in\mathcal{P}_i} 2\kappa r_j \sum_{u\in\Omega(j)} \left(v_u - v_u^{ref}\right), \end{aligned} \tag{34}$$

where $\Omega(j) = \{j\} \cup \delta(j)$.

Next, we try to evaluate the coefficient before each $v_u - v_u^{ref}$. As shown in Fig. 8, there are only three cases: 1) $u \in \mathcal{P}_i$; 2) $u \in \Omega(i)$; 3) all the rest, i.e., $u \in \mathcal{N}\backslash\mathcal{P}_i\backslash\Omega(i)$. We analyze the corresponding coefficient case by case.

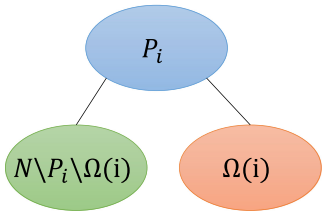


Fig. 8. Visualization of three cases in proving the equivalence observation.

Case 1: In the first case, the coefficient should be $2\sum_{j\in\mathcal{P}_i} 2\kappa r_j$. Note that, since in this case, $u \in \Omega(i)$, this implies that $\mathcal{P}_i \cap \mathcal{P}_u = \mathcal{P}_i$. Hence, the coefficient can be written as $2\sum_{j\in\mathcal{P}_i\cap\mathcal{P}_u} 2\kappa r_j$.

Case 2: In the second case, the coefficient should be $2\sum_{j\in\mathcal{P}_u} 2\kappa r_j$. Recall that, in this case, $u \in \mathcal{P}_i$. We have $\mathcal{P}_i \cap \mathcal{P}_u = \mathcal{P}_u$. Hence, the coefficient can be again written as $2\sum_{j\in\mathcal{P}_i\cap\mathcal{P}_u} 2\kappa r_j$.

Case 3: The last case can be analyzed similarly. The coefficient here should be $2\sum_{j\in\mathcal{P}_i} 2\kappa r_j$. Note that, in this case $\mathcal{P}_i \cap \mathcal{P}_u = \mathcal{P}_i$. The coefficient can also be written as $2\sum_{j\in\mathcal{P}_i\cap\mathcal{P}_u} 2\kappa r_j$.

Combining all the three cases lead to the following conclusion:

$$\begin{aligned} 2\sum_{j\in\mathcal{P}_i} w_j^T r_j &= 2\sum_{j\in\mathcal{P}_i} 2\kappa r_j \sum_{u\in\Omega(j)} \left(v_u - v_u^{ref}\right) \\ &= \sum_{k=1}^{n} \left(2\sum_{u\in\mathcal{P}_k\cap\mathcal{P}_i} r_u\right) 2\kappa\left(v_k - v_k^{ref}\right). \end{aligned} \tag{35}$$

Recall that, we have defined

$$R_{ki} = 2\sum_{u\in\mathcal{P}_k\cap\mathcal{P}_i} r_u.$$

Also, the Lagrangian multiplier for the complete network is defined as follows:

$$w_k^C = 2\kappa(v_k - v_k^{ref}).$$

Together with (35), we can obtain the desired conclusion that

$$2\sum_{j\in\mathcal{P}_i} w_j^T r_j = \sum_{k=1}^{n} w_k^C R_{ki}, \forall i \in \mathcal{N}. \tag{36}$$

B. Proof for Equation (29)

Note that

$$\frac{\partial\lambda}{\partial p_i^g} = \sum_{k=1}^{n} 2\kappa\left(v_k - v_k^{ref}\right)\frac{\partial v_k}{\partial p_i^g} = 2\kappa\sum_{k=1}^{n}\left(v_k - v_k^{ref}\right)R_{ki},$$

$$\frac{\partial\lambda}{\partial q_i^g} = \sum_{k=1}^{n} 2\kappa\left(v_k - v_k^{ref}\right)\frac{\partial v_k}{\partial q_i^g} = 2\kappa\sum_{k=1}^{n}\left(v_k - v_k^{ref}\right)X_{ki}.$$

[4]Here, we assume that the voltage magnitude can be measured accurately with a high sampling rate. There are still certain technical challenges to achieve this goal. However, a detailed discussion is beyond the scope of this paper.

Hence, we have

$$\begin{aligned}
&\frac{1}{2\kappa}(\nabla\lambda(\mathbf{a}^1)-\nabla\lambda(\mathbf{a}^2))^T(\mathbf{a}^1-\mathbf{a}^2)\\
&=\sum_{i=1}^n\sum_{k=1}^n(v_k(\mathbf{a}^1)-v_k(\mathbf{a}^2))R_{ki}(p_i^1-p_i^2)\\
&\quad+\sum_{i=1}^n\sum_{k=1}^n(v_k(\mathbf{a}^1)-v_k(\mathbf{a}^2))X_{ki}(q_i^1-q_i^2)\\
&=\sum_{i=1}^n\sum_{k=1}^n\sum_{j=1}^n R_{kj}R_{ki}(p_j^1-p_j^2)(p_i^1-p_i^2)\\
&\quad+2\sum_{i=1}^n\sum_{k=1}^n\sum_{j=1}^n R_{kj}X_{ki}(p_j^1-p_j^2)(q_i^1-q_i^2)\\
&\quad+\sum_{i=1}^n\sum_{k=1}^n\sum_{j=1}^n X_{kj}X_{ki}(q_j^1-q_j^2)(q_i^1-q_i^2).
\end{aligned}$$

After standard manipulation, we can show that

$$\begin{aligned}
&\frac{1}{2\kappa}(\nabla\lambda(\mathbf{a}^1)-\nabla\lambda(\mathbf{a}^2))^T(\mathbf{a}^1-\mathbf{a}^2)\\
&=\sum_{k=1}^n\left(\sum_{i=1}^n R_{ki}(p_i^1-p_i^2)\right)^2+\sum_{k=1}^n\left(\sum_{i=1}^n X_{ki}(q_i^1-q_i^2)\right)^2\\
&\quad+2\sum_{i=1}^n\sum_{k=1}^n\sum_{j=1}^n R_{kj}X_{ki}(p_j^1-p_j^2)(q_i^1-q_i^2)\\
&=\sum_{k=1}^n\left(\sum_{i=1}^n R_{ki}(p_i^1-p_i^2)+\sum_{i=1}^n X_{ki}(q_i^1-q_i^2)\right)^2\\
&\geq\left(\sum_{k=1}^n\min\{\min_i R_{ki},\min_i X_{ki}\}\right)\|\mathbf{a}^1-\mathbf{a}^2\|_2^2,
\end{aligned}$$

which yields equation (29).

C. Proof for Theorem 2

$$\begin{aligned}
&\|\mathbf{x}^{t+1}-\mathbf{x}^*\|_2^2=\|\mathbf{x}^t-\alpha\mathbf{g}^t-\mathbf{x}^*\|_2^2\\
&=\|\mathbf{x}^t-\mathbf{x}^*\|_2^2-2\alpha(\mathbf{g}^t)^T(\mathbf{x}^t-\mathbf{x}^*)+\alpha^2\|\mathbf{x}^t\|_2^2\\
&\leq\|\mathbf{x}^t-\mathbf{x}^*\|_2^2-2\alpha m\|\mathbf{x}^t-\mathbf{x}^*\|_2^2+\alpha^2\|\mathbf{g}^t\|_2^2\\
&\leq(1-2\alpha m)^t\|\mathbf{x}^1-\mathbf{x}^*\|_2^2\\
&\quad+\alpha(1-(1-2\alpha m)^t)/2m\max_{1\leq k\leq t}\|\mathbf{g}^k\|_2^2\\
&\leq(1-2\alpha m)^t\|\mathbf{x}^1-\mathbf{x}^*\|_2^2+\alpha/2m\max_{1\leq k\leq t}\|\mathbf{g}^k\|_2^2. \qquad (37)
\end{aligned}$$

The first inequality utilizes the fact that $f(\mathbf{x})$ is strongly convex, with parameter m. The last two inequalities are standard manipulations.

D. Proof for Theorem 3

This proof aligns with that of Theorem 2

$$\begin{aligned}
&\|\mathbf{a}^{t+1}-\mathbf{a}^*\|_2^2=\|\mathbf{a}^t-\alpha\hat{\mathbf{g}}^t-\mathbf{a}^*\|_2^2\\
&=\|\mathbf{a}^t-\alpha(\mathbf{g}^t+\xi^t)-\mathbf{a}^*\|_2^2\\
&=\|\mathbf{a}^t-\mathbf{a}^*\|_2^2-2\alpha(\mathbf{g}^t+\xi^t)^T(\mathbf{a}^t-\mathbf{a}^*)+\alpha^2\|\mathbf{g}^t+\xi^t\|_2^2\\
&\leq\|\mathbf{a}^t-\mathbf{a}^*\|_2^2-2\alpha\gamma\|\mathbf{a}^t-\mathbf{a}^*\|_2^2\\
&\quad-2\alpha(\xi^t)^T(\mathbf{a}^t-\mathbf{a}^*)+\alpha^2\|\mathbf{g}^t+\xi^t\|_2^2\\
&\leq(1-2\gamma\alpha)^t\|\mathbf{a}^1-\mathbf{a}^*\|_2^2\\
&\quad+(1-(1-2\gamma\alpha)^t)/\gamma\max_{1\leq k\leq t}|(\xi^k)^T(\mathbf{a}^k-\mathbf{a}^*)|\\
&\quad+\alpha(1-(1-2\gamma\alpha)^t)/2\gamma\max_{1\leq k\leq t}\|\mathbf{g}^k+\xi^k\|_2^2\\
&\leq(1-2\gamma\alpha)^t\|\mathbf{a}^1-\mathbf{a}^*\|_2^2+1/\gamma\max_{1\leq k\leq t}\|\xi^k\|\|\mathbf{a}^k-\mathbf{a}^*\|\\
&\quad+\alpha/2\gamma\max_{1\leq k\leq t}\|\mathbf{g}^k+\xi^k\|_2^2. \qquad (38)
\end{aligned}$$

Again, the first inequality utilizes the fact that $h(\mathbf{a})$ is strongly convex. The last inequality uses the Cauchy-Schwarz inequality. Note that, due to the strong convexity of problem (14), it admits a unique global minimizer $\mathbf{a}^*$.

For the second part, we propose the sufficient condition, under which the convergence is guaranteed. We have

$$\begin{aligned}
&\|\mathbf{a}^{t+1}-\mathbf{a}^*\|_2^2=\|\mathbf{a}^t-\alpha\hat{\mathbf{g}}^t-\mathbf{a}^*\|_2^2\\
&\leq\|\mathbf{a}^t-\mathbf{a}^*\|_2^2-2\alpha\gamma\|\mathbf{a}^t-\mathbf{a}^*\|_2^2\\
&\quad-2\alpha(\xi^t)^T(\mathbf{a}^t-\mathbf{a}^*)+\alpha^2\|\mathbf{g}^t+\xi^t\|_2^2\\
&\leq\|\mathbf{a}^t-\mathbf{a}^*\|_2^2-2\alpha\gamma\|\mathbf{a}^t-\mathbf{a}^*\|_2^2\\
&\quad+2\alpha\|\xi^t\|\|\mathbf{a}^t-\mathbf{a}^*\|+\alpha^2\|\mathbf{g}^t+\xi^t\|_2^2\\
&\leq\|\mathbf{a}^t-\mathbf{a}^*\|_2^2-\alpha\gamma\|\mathbf{a}^t-\mathbf{a}^*\|_2^2\\
&\leq(1-\gamma\alpha)^t\|\mathbf{a}^1-\mathbf{a}^*\|_2^2. \qquad (39)
\end{aligned}$$

In the third inequality, we use the assumption that for all $t=1,\cdots,T$,

$$2\|\xi^t\|\|\mathbf{a}^t-\mathbf{a}^*\|+\alpha\|\mathbf{g}^t+\xi^t\|_2^2\leq\gamma\|\mathbf{a}^1-\mathbf{a}^*\|_2^2. \qquad (40)$$

This completes the convergence proof.

E. Sketch Proof for Lemma 5

It suffices to verify

$$u_i(a_i^\dagger,a_{-i})-u_i(a_i^\ddagger,a_{-i})=\Psi(a_i^\dagger,a_{-i})-\Psi(a_i^\ddagger,a_{-i}). \qquad (41)$$

The lemma immediate follows.

F. Sketch Proof for Theorem 6

This is one of the nice properties of potential games [27]. Intuitively, at the global minimizers to problem (9), every player's profit has automatically been maximized due to the potential function Ψ.

REFERENCES

[1] R. Yan and T. K. Saha, "Investigation of voltage stability for residential customers due to high photovoltaic penetrations," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 651–662, May 2012.

[2] B.-I. Crăciun, T. Kerekes, D. Séra, and R. Teodorescu, "Overview of recent grid codes for PV power integration," in *Proc. 13th Int. Conf. Optim. Elect. Electron. Equip.*, 2012, pp. 959–965.

[3] M. E. Baran and I. M. El-Markabi, "A multiagent-based dispatching scheme for distributed generators for voltage support on distribution feeders," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 52–59, Feb. 2007.

[4] K. Turitsyn, P. Ulc, S. Backhaus, and M. Chertkov, "Distributed control of reactive power flow in a radial distribution circuit with high photovoltaic penetration," in *Proc. IEEE PES General Meeting*, Jul. 2010, pp. 1–6.

[5] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, "Distributed reactive power feedback control for voltage regulation and loss minimization," *IEEE Trans. Autom. Control*, vol. 60, no. 4, pp. 966–981, Apr. 2015.

[6] 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.

[7] H. J. Liu, W. Shi, and H. Zhu, "Decentralized dynamic optimization for power network voltage control," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 3, pp. 568–579, Sep. 2017.

[8] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[9] N. Li, L. Chen, and S. H. Low, "Exact convex relaxation of OPF for radial networks using branch flow model," in *Proc. IEEE 3rd Int. Conf. Smart Grid Commun.*, Nov. 2012, pp. 7–12.

[10] L. Gan, N. Li, U. Topcu, and S. H. Low, "Exact convex relaxation of optimal power flow in radial networks," *IEEE Trans. Autom. Control*, vol. 60, no. 1, pp. 72–87, Jan. 2015.

[11] B. Zhang, A. Y. S. Lam, A. D. Domnguez-Garca, and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1714–1726, Jul. 2015.

[12] J. Schiffer, T. Seel, J. Raisch, and T. Sezi, "Voltage stability and reactive power sharing in inverter-based microgrids with consensus-based distributed voltage control," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 1, pp. 96–109, Jan. 2016.

[13] J. Schiffer, D. Zonetti, R. Ortega, A. M. Stankovic, T. Sezi, and J. Raisch, "Modeling of microgrids—From fundamental physics to phasors and voltage sources," *Automatica*, vol. 74, pp. 135–150, 2016.

[14] M. Farivar, X. Zhou, and L. Chen, "Local voltage control in distribution systems: An incremental control algorithm," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Nov. 2015, pp. 732–737.

[15] X. Zhou, M. Farivar, and L. Chen, "Pseudo-gradient based local voltage control in distribution networks," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2015, pp. 173–180.

[16] X. Zhou, Z. Liu, M. Farivar, L. Chen, and S. Low, "Reverse and forward engineering of local voltage control in distribution networks," arXiv:1801.02015, Jan. 2018.

[17] A. Hauswirth, S. Bolognani, F. Dorfler, and G. Hug, "Projected gradient descent on Riemannian manifolds with applications to online power system optimization," in *Proc. 54th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2016, pp. 225–232.

[18] M. Jafarian, J. Scherpen, and M. Aiello, "A price-based approach for voltage regulation and power loss minimization in power distribution networks," in *Proc. IEEE 55th Conf. Decision Control*, Dec. 2016, pp. 680–685.

[19] C. Wu, G. Hug, and S. Kar, "Distributed voltage regulation in distribution power grids: Utilizing the photovoltaics inverters," in *Proc. IEEE Amer. Control Conf.*, May 2017, pp. 2725–2731.

[20] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 735–743, Jan. 1989.

[21] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification—Part I," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2554–2564, Aug. 2013.

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

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

[24] R. Fletcher, *Practical Methods of Optimization*. Hoboken, NJ, USA: Wiley, 2013.

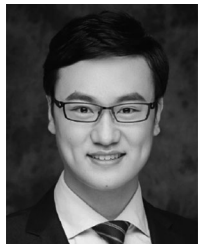
[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[26] S. Bolognani and F. Drfler, "Fast power system analysis via implicit linearization of the power flow manifold," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2015, pp. 402–409.

[27] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, 1996.

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

[29] W. Kersting, "Radial test feeders," 1991. [Online]. Available: http://ewh.ieee.org/soc/pes/dsacom/testfeeders.html



Chenye Wu (S'11–M'14) received the Ph.D. degree from the Institute for Interdisciplinary Information Sciences (IIIS), Tsinghua University, Beijing, China, 2013. He is currently an Assistant Professor with the IIIS, Tsinghua University. He worked at ETH Zurich as a wiss. Mitarbeiter (Research Scientist), working with Professor G. Hug, in 2016. Before that, Prof. K. Poolla and Prof. P. Varaiya hosted Dr. Wu as a Postdoctoral Researcher at UC Berkeley for two years. From 2013 to 2014, he spent one year at Carnegie Mellon University as a postdoctoral fellow, hosted by Prof. G. Hug and Prof. S. Kar.

His Ph.D. Advisor is Prof. A. Yao, the laureate of A.M. Turing Award in the year of 2000. He was the best paper co-recipients of IEEE SmartGridComm 2012, and IEEE PES General Meeting 2013. He is currently working on economic analysis, optimal control and operation of power systems.

Gabriela Hug (S'05–M'08–SM'14) was born in Baden, Switzerland. She received the M.Sc. degree in electrical engineering and the Ph.D. degree from the Swiss Federal Institute of Technology, Zurich, Switzerland, in 2004 and 2008, respectively.

After Ph.D. degree, she was with the Special Studies Group of Hydro One in Toronto, Canada, and from 2009 to 2015, she was an Assistant Professor with Carnegie Mellon University, Pittsburgh, PA, USA. She is currently an Associate Professor with the Power Systems Laboratory, ETH Zurich. Her research interest includes control and optimization of electric power systems.

Soummya Kar (S'05–M'10) received the B.Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India, in May 2005 and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2010. From June 2010 to May 2011, he was with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, as a Postdoctoral Research Associate.

He is currently an Associate Professor with Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. His research interests include decision-making in large-scale networked systems, stochastic systems, multiagent systems and data science, with applications to cyber-physical systems and smart energy systems.