Distributed Controllers Seeking AC Optimal Power Flow Solutions Using ADMM

Yijian Zhang, Mingyi Hong, *Member, IEEE*, Emiliano Dall'Anese, *Member, IEEE*, Sairaj Dhople, *Member, IEEE*, and Zi Xu

***Abstract*—This paper focuses on power distribution systems with inverter-interfaced renewable energy sources (RESs), and develops a distributed control framework to steer the RES output powers to solutions of AC optimal power flow (OPF) problems. The design of the distributed control algorithm is based on suitable linear approximation of the AC power-flow equations, and leverages the so-called alternating direction method of multipliers (ADMM). Convergence of the RES-inverter output powers to solutions of the approximate AC OPF problem is established under suitable conditions on the mismatches between the commanded setpoints and actual RES output powers. Overall, since the proposed scheme can be cast as an ADMM with inexact primal and dual updates, the convergence results can be applied to more general distributed optimization settings.**

***Index Terms*—ADMM, distribution systems, distributed algorithms, optimal power flow, renewable energy sources.**

I. Introduction

This paper focuses on power distribution systems with high integration of inverter-interfaced renewable energy sources (RESs) and develops distributed control algorithms that steer the RES output powers to solutions of AC optimal power flow (OPF) problems. The overarching objective is to leverage the flexibility offered by power-electronics-interfaced RESs to address reliability and power-quality concerns that emerge from reverse power flows and renewable generation volatility [1], [2]. Similar to e.g., [3]–[6], the general control strategy involves a continuous update of the RES setpoints based on current output powers and given OPF objectives (e.g., ensuring voltage regulation, minimization of power losses, as well as maximization of economic benefits to utility and end users).

Prior works in context include [7], wherein feedback control architectures that seek Karush-Kuhn-Tucker (KKT) optimality conditions for economic dispatch in transmission systems are developed, and [8], where a heuristic comprising continuous-time dual ascent and discrete-time reference-signal updates is proposed; local stability of the resultant closed-loop system is also established in [8]. A feedback control algorithm for a finite-horizon economic dispatch problem for distributed energy resources is also considered in [9]. Focusing on AC OPF models, a continuous-time saddle-point-flow method is utilized in [10]; however, stability analysis is available only for specific optimization settings. A reactive power control strategy is proposed in [11] for single-phase distribution systems with a tree topology based on the so-called extremum-seeking control method. Stochastic dual-subgradient solvers are developed in [12] to achieve the solutions of ergodic OPF formulations, based on exact and approximate grid models. An online AC OPF algorithm is proposed in [4] for distribution systems with a tree topology. A controller for a number of resources in general microgrid and distribution-system settings is developed in [5], [6], based on gradient-steering algorithms; the algorithm in [5], [6] is composable in the sense that subsystems can be aggregated into virtual devices that hide their internal complexity, it accounts for errors in the implementable power setpoints, and the average setpoints are provably convergent (on average) to the minimum of the considered control objective. A dual-subgradient method is utilized in [3] to develop feedback controllers that drive the RES output powers to solutions of convex surrogates of the AC OPF; convergence results are available for diminishing stepsize rules in the dual subgradient. A feedback control strategy is proposed in [13] to track solutions of time-varying OPF solutions based on primal-dual methods applied to a modified Lagrangian function.

A key contribution of the present paper consists in leveraging the so-called Alternating Direction Method of Multipliers (ADMM) [14] to develop distributed controllers that pursue solutions of the AC OPF problem. The choice of ADMM is motivated by its favorable scalability with respect to the system size as well as the superior convergence properties compared to subgradient methods [15], [16]. For instance, while convergence results are available for the control scheme in [3] only for for diminishing stepsize rules, the ADMM-based framework proposed here allows one to utilize a constant stepsize, which is desirable for practical implementations. Q-linear convergence is achieved in the gradient-based method proposed in [13], but at the cost of perturbing the optimal solution of the underlying AC OPF. To facilitate the design of computationally affordable ADMM-based controllers, the paper leverages appropriate linear approximations of the AC power-flow equations [17]–[21]. Based on this linear approximation, two distinct control strategies are developed to trade

Y. Zhang and M. Hong are with department of IMSE, Iowa State University, Ames, USA; emails: yijian@iastate.edu, mingyi@iastate.edu. E. Dall'Anese is with the National Renewable Energy Laboratory, Golden, USA; email: emiliano.dallanese@nrel.gov. S. Dhople is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, USA; email: sdhople@umn.edu. Z. Xu is with the Department of Mathematics, Shanghai University, Shanghai, China; email: xuzi@i.shu.edu.cn.

This work was supported by the U.S. Department of Energy under Contract No. DE-AC36-08GO28308 with the National Renewable Energy Laboratory. The work of E. Dall'Anese, Y. Zhang, and M. Hong was supported by the Laboratory Directed Research and Development Program at the National Renewable Energy Laboratory. S. V. Dhople was supported in part by the National Science Foundation under the CAREER award, ECCS-1453921. Zi Xu was supported by National Nature Science Foundation of China (11571221).

off convergence speed for computational complexity: in the first strategy, the update of the optimization variables that are proxies for voltage magnitudes is performed by solving a linearly-constrained quadratic program, whereas a simpler projected gradient step is involved in the second setting. In both cases, convergence of the RES-inverter output powers is established under suitable conditions on the stepsize and responsiveness of the RES inverters to power commands. The algorithms afford a distributed solution where both the distribution system operator (DSO) and RES-owners pursue given performance objectives, while ensuring that system operational constraints are observed.

The resultant control framework is close in spirit to the feedback-control strategies proposed in, e.g., [3], [4], [6], where RES setpoints are continuously updated based on current output powers, given OPF objectives, as well as relevant voltage constraints; however, compared to [6] and [4], the proposed framework does not resort to relaxations (e.g., barrier functions) to enforce voltage limits. Further, while [4] is applicable to single-phase radial systems, the method proposed here is applicable to multi-phase settings. Compared to [3], the proposed method requires less stringent assumptions on the mismatch between commanded setpoints and current system outputs and offers improved convergence properties.

Overall, the paper offers the following contributions:

- Online algorithms that pursue solutions of AC OPF problems are designed by leveraging (and suitably adapting) the ADMM;
- Two different algorithmic solutions are proposed to trade-off convergence for computational complexity; and,
- Convergence of ADMM with inexact primal and dual updates is established. To the best of our knowledge, this is a unique contribution in the broader optimization literature.

Some preliminary results were presented in [22].

II. Problem Formulation

A. Notation

Throughout the paper, $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary parts of a complex number, respectively; for given vector $\mathbf{x}$, $\mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix with diagonal entries composed of the components of $\mathbf{x}$; $j := \sqrt{-1}$. Notation $\|\mathbf{x}\|$ denotes the ℓ_2 norm of $\mathbf{x}$. For column vectors $\mathbf{x}, \mathbf{y}$, $[\mathbf{x}; \mathbf{y}] := [\mathbf{x}^\top, \mathbf{y}^\top]^\top$; For a given matrix $\mathbf{X}$, vector $\mathbf{X}(i)$ denotes the ith row of $\mathbf{X}$. For given matrix $\mathbf{D}$ and vector $\mathbf{z}$, $\|\mathbf{z}\|^2_{\mathbf{D}} = \mathbf{z}^\top \mathbf{D} \mathbf{z}$.

B. Problem setup

Consider modeling the dynamics of the output-powers of the RES inverters through the following general dynamical model [3], [23], [24]:

$$\dot{\mathbf{x}}_i(t) = \mathbf{f}_i(\mathbf{x}_i(t), \mathbf{u}_i(t)), \tag{1a}$$

$$\mathbf{y}_i(t) = \mathbf{r}_i(\mathbf{x}_i(t)), \tag{1b}$$

where:

- $\mathbf{x}_i(t) := [P_i(t), Q_i(t)]^\intercal$, with $P_i(t)$ and $Q_i(t)$ denoting the active and reactive output powers (averaged over one AC cycle) of the RES inverter i;
- $\mathbf{u}_i(t) := [\bar{P}_i(t), \bar{Q}_i(t)]^\intercal$ collects the *commanded* active and reactive powers (i.e., power setpoints);
- $\mathbf{f}_i : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ and $\mathbf{r}_i : \mathbb{R}^2 \to \mathbb{R}^2$ are arbitrary (non)linear functions; and,
- $\mathbf{y}_i(t)$ is a measurement of $\mathbf{x}_i(t)$ collected at time t.

These dynamics capture the behavior of primal-level controllers embedded into the RES inverters [24]. For a given power setpoint, the following is assumed regarding the regulation capabilities of the primal-level controllers [8], [23]:

Assumption 1. *For a given power setpoint $\mathbf{u}_i$, (1) is asymptotically stable and the equilibrium point $\mathbf{x}_i$ satisfies:*

$$\mathbf{0} = \mathbf{f}_i(\mathbf{x}_i, \mathbf{u}_i), \quad \mathbf{u}_i = \mathbf{r}_i(\mathbf{x}_i). \tag{2}$$

This assumption captures the operation of existing devices, where the primary-controllers are designed so that the output powers are regulated to the commanded powers $\mathbf{x}_i$, provided the commanded powers are feasible [24].

Regarding the electrical system, consider a distribution network with $N+1$ nodes collected in the set $\mathcal{N} = \{0\} \cup \mathcal{N}_D \cup \mathcal{N}_O$, where 0 denotes the secondary of the step-down transformer, and $\mathcal{N}_D, \mathcal{N}_O$ denote the set of locations with and without RESs, respectively. Let $\mathbf{Y}_{\text{net}} \in \mathbb{C}^{(N+1)\times(N+1)}$ denote the network admittance matrix, which is formed according to the system topology and π-equivalent model of the distribution lines. Define the vector $\mathbf{i} := [I_1, \ldots, I_N]^\top \in \mathbb{C}^N$, where I_n denotes the phasor of the current injected at node n. Let $\mathbf{v} := [V_1, \ldots, V_N]^\top \in \mathbb{C}^N$, where $V_i = |V_i| \angle \theta_i \in \mathbb{C}$ denotes the voltage phasor at node i, where $V_0 e^{j\theta_0}$ is the slack-bus voltage with V_0 denoting the voltage magnitude. Let $\bar{P}_i + j\bar{Q}_i$ denote the setpoints of RES $i \in \mathcal{N}_D$, and define $\mathbf{u}_i := [\bar{P}_i, \bar{Q}_i]^\top$ for brevity. Similarly, let $P_{l,i} + jQ_{l,i}$ denote the non-controllable complex load at node $i \in \mathcal{N}$ and $\mathbf{d}_i := [P_{l,i}, Q_{l,i}]^\top$. Based on Kirchhoff's Current Law and Ohm's Law, we can establish the following linear relationship:

$$\begin{bmatrix} I_0 \\ \mathbf{i} \end{bmatrix} = \underbrace{\begin{bmatrix} \tilde{y} & \bar{\mathbf{y}}^\top \\ \bar{\mathbf{y}} & \mathbf{Y} \end{bmatrix}}_{\mathbf{Y}_{\text{net}}} \begin{bmatrix} V_0 e^{j\theta_0} \\ \mathbf{v} \end{bmatrix}, \tag{3}$$

where $\mathbf{Y}_{\text{net}}$ is partitioned as $\bar{\mathbf{y}} \in \mathbb{C}^N$, $\mathbf{Y} \in \mathbb{C}^{N\times N}$, and $\tilde{y} \in \mathbb{C}\backslash\{0\}$.

Consider then the following prototypical OPF formulation to optimize the steady-state operation of the distribution network:

$$\min_{\mathbf{v},\mathbf{i},\mathbf{u}_i} H(\mathbf{v}) + \sum_{i\in\mathcal{N}_D} G_i(\mathbf{u}_i) \tag{OPF}$$

$$\text{s.t.} \quad \mathbf{i} = \mathbf{Y}\mathbf{v} + \bar{\mathbf{y}} V_0 e^{j\theta_0}, \tag{4a}$$

$$V_i I_i^* = \bar{P}_i - P_{l,i} + j(\bar{Q}_i - Q_{l,i}), \ \forall i \in \mathcal{N}_D \tag{4b}$$

$$V_n I_n^* = -P_{l,n} - jQ_{l,n}, \quad \forall n \in \mathcal{N}_O \tag{4c}$$

$$V^{\min} \le |V_i| \le V^{\max}, \quad \forall i \in \mathcal{N} \tag{4d}$$

$$\mathbf{u}_i = [\bar{P}_i, \bar{Q}_i]^\top \in \mathcal{Y}_i \quad \forall i \in \mathcal{N}, \tag{4e}$$

where $P_{l,i} + jQ_{l,i}$ denotes the loads at node i; (4b) and (4c) describe power-balance equations for nodes with and without

RES inverters, respectively; $V^{\min}$ and $V^{\max}$ are prescribed voltage limits (e.g., ANSI C84.1 limits); the function $H(\mathbf{v}) : \mathbb{C}^N \to \mathbb{R}$ captures network-oriented performance objectives; $G_i(\mathbf{u}_i) : \mathbb{R}^2 \to \mathbb{R}$ models optimization objectives at the RES side (e.g., minimization of real power curtailed and reactive power provisioning). Finally, the set $\mathcal{Y}_i \subset \mathbb{R}^2$ models hardware and operational constraints of the inverter i. For example, for photovoltaic (PV) systems, $\mathcal{Y}_i$ takes the following form:

$$\mathcal{Y}_i := \{(\bar{P}_i, \bar{Q}_i) : P_i^{\min} \leq \bar{P}_i \leq P_i^{\text{av}}, \bar{P}_i^2 + \bar{Q}_i^2 \leq S_i^2\} \quad (5)$$

where $P_i^{\text{av}} \geq 0$ denotes the available real power, and S_i is the inverter capacity.

Problem (4) defines the optimal operating setpoints $\mathbf{u}_i = [\bar{P}_i, \bar{Q}_i]^\intercal$ of the RES inverter i in terms of commanded inputs and, based on Assumption 1, of the steady-state output powers. However, problem (4) is a nonconvex and NP hard problem in general [25]. Recently, convex relaxation methods have been explored to solve the OPF with reduced computational burden, while possibly retaining globally optimal solutions [26]. In this paper, to facilitate the design of low-complexity controllers that can be implemented on microcontrollers that accompany power-electronics interfaces of gateways and inverters, this paper leverages suitable linear approximations of (4) [19]–[21]. In particular, a linearization approach proposed in [20] is utilized, which is briefly discussed in the next section.

Remark. For ease of exposition, the problem formulation is tailored to the case where one RES is connected at each of the nodes $\mathcal{N}_D$; however, the proposed algorithm can be utilized in settings where RES aggregations are present at (some of) the nodes.

C. Leveraging approximate linear models

By plugging (4a) into (4b)-(4c), the power-balance equations can be rewritten as:

$$\mathbf{s} = \text{diag}(\mathbf{v})\mathbf{i}^* = \text{diag}(\mathbf{v})(\mathbf{Y}^*\mathbf{v}^* + \bar{\mathbf{y}}^* V_0 e^{-j\theta_0}), \quad (6)$$

where $\mathbf{s}$ is a vector collecting the net complex power injections throughout the network. Consider then re-writing the voltages $\mathbf{v}$ satisfying the nonlinear power-balance equations (6) as $\mathbf{v} = \mathbf{v}_{\text{nom}} + \mathbf{v}_d$, where $\mathbf{v}_{\text{nom}} = |\mathbf{v}_{\text{nom}}| \angle \boldsymbol{\theta}_{\text{nom}} \in \mathbb{C}^N$ is a predefined nominal voltage profile and $\mathbf{v}_d$ captures deviations around $\mathbf{v}_{\text{nom}}$. Similar to [20], consider further setting $\mathbf{v}_{\text{nom}}$ as $\mathbf{v}_{\text{nom}} = -\mathbf{Y}^{-1}\bar{\mathbf{y}} V_0 e^{j\theta_0}$, which corresponds to the voltage across the network with zero current injections (however, other linearization points can be utilized). Then, by plugging $\mathbf{v}_{\text{nom}}$ into (6) and neglecting the second-order terms (in $\mathbf{v}_d$), we obtain the following expression:

$$\mathbf{v}_d = \mathbf{Y}^{-1}\text{diag}\left(\frac{1}{\mathbf{v}_{\text{nom}}^*}\right)\mathbf{s}^*. \quad (7)$$

After expanding (7), one can readily derive expressions for the real and the imaginary parts of $\mathbf{v}_d$ separately; however, the resulting expression will couple the components of $\mathbf{p}$ and $\mathbf{q}$, thus challenging the design of computationally-affordable distributed algorithms. To bypass this hurdle, consider rearranging terms to arrive at the following equivalent expression:

$$\text{diag}(\mathbf{v}_{\text{nom}}^*)\mathbf{Y}\mathbf{v}_d = \mathbf{s}^*. \quad (8)$$

Define $\mathbf{Y} := \mathbf{G} + j\mathbf{B}$, where $\mathbf{G} \in \mathbb{R}^{N\times N}$ is the conductance matrix and $\mathbf{B} \in \mathbb{R}^{N\times N}$ is the susceptance matrix. Further, let $\mathbf{M} := \text{diag}(|\mathbf{v}_{\text{nom}}| \cos\boldsymbol{\theta}_{\text{nom}})$ and $\mathbf{N} := \text{diag}(|\mathbf{v}_{\text{nom}}| \sin\boldsymbol{\theta}_{\text{nom}})$. By expanding (8), the following expressions can be obtained:

$$(\mathbf{MG} + \mathbf{NB})\text{Re}(\mathbf{v}_d) - (\mathbf{MB} - \mathbf{NG})\text{Im}(\mathbf{v}_d) = \mathbf{p} \quad (9a)$$

$$-(\mathbf{MG} + \mathbf{NB})\text{Im}(\mathbf{v}_d) - (\mathbf{MB} - \mathbf{NG})\text{Re}(\mathbf{v}_d) = \mathbf{q} \quad (9b)$$

where the components of vectors $\mathbf{p}$ and $\mathbf{q}$ are defined as: $p_i = \bar{P}_i - P_{l,i}$ and $q_i = \bar{Q}_i - Q_{l,i}$ for $i \in \mathcal{N}_D$; whereas, $p_i = -P_{l,i}$ and $q_i = -Q_{l,i}$ for $i \in \mathcal{N}_O$. Clearly the expression for $\mathbf{p}$ and $\mathbf{q}$ are decoupled. For notational simplicity, define the vector $\boldsymbol{\Delta} := [\text{Re}(\mathbf{v}_d); \text{Im}(\mathbf{v}_d)] \in \mathbb{R}^{2N}$.

Based on these definitions, and noticing that $|\mathbf{v}_{\text{nom}}| + \text{Re}\{\mathbf{v}_d\}$ serves as a first-order approximation to the voltage magnitudes across the distribution network whenever the entries of $\mathbf{v}_{\text{nom}}$ dominate $\mathbf{v}_d$, a convex surrogate of the OPF problem can be formulated as:

$$\min_{\boldsymbol{\Delta}, \mathbf{u}_i} H(\boldsymbol{\Delta}) + \sum_{i\in\mathcal{N}_D} G_i(\mathbf{u}_i) \quad \text{(OPF-2)}$$

$$\text{s.t.} \quad \mathbf{C}(i)\boldsymbol{\Delta} - \bar{P}_i + P_{l,i} = 0, \quad i \in \mathcal{N}\backslash\{0\} \quad (10a)$$

$$\mathbf{D}(i)\boldsymbol{\Delta} - \bar{Q}_i + Q_{l,i} = 0, \quad i \in \mathcal{N}\backslash\{0\} \quad (10b)$$

$$\boldsymbol{\Delta} \in \mathcal{V}, \ \mathbf{u}_i = [\bar{P}_i, \bar{Q}_i]^\top \in \mathcal{Y}_i.$$

where $\bar{P}_i = \bar{Q}_i = 0$ for nodes $i \in \mathcal{N}_O$ and:

$$\mathbf{C} := \big(\mathbf{MG} + \mathbf{NB}, -\mathbf{MB} + \mathbf{NG}\big) \in \mathbb{R}^{N\times 2N} \quad (11a)$$

$$\mathbf{D} := \big(-\mathbf{MB} + \mathbf{NG}, -\mathbf{MG} - \mathbf{NB}\big) \in \mathbb{R}^{N\times 2N}. \quad (11b)$$

The set $\mathcal{V}$ is designed to enforce voltage regulation as [19]:

$$\mathcal{V} := \{\boldsymbol{\Delta} \mid V^{\min} - |\mathbf{v}_{\text{nom,i}}| \leq \Delta_i \leq V^{\max} - |\mathbf{v}_{\text{nom,i}}|, \ i = 1, \ldots, N\}.$$

For notational simplicity, denote $\boldsymbol{\Phi}_i = [\mathbf{C}(i); \mathbf{D}(i)] \in \mathbb{R}^{2\times 2N}$, $\boldsymbol{\Phi} = [\boldsymbol{\Phi}_1; \cdots; \boldsymbol{\Phi}_N] \in \mathbb{R}^{2N\times 2N}$, and $\mathbf{d}_i = [P_{l,i}, Q_{l,i}]^\intercal$. Then, (10) can be rewritten in the following compact form:

$$\min_{\boldsymbol{\Delta}, \mathbf{u}_i} H(\boldsymbol{\Delta}) + \sum_{i\in\mathcal{N}_D} G_i(\mathbf{u}_i) \quad \text{(OPF-3)}$$

$$\text{s.t.} \quad \boldsymbol{\Phi}_i\boldsymbol{\Delta} - \mathbf{u}_i + \mathbf{d}_i = \mathbf{0}, i \in \mathcal{N}\backslash\{0\}, \quad (12a)$$

$$\boldsymbol{\Delta} \in \mathcal{V}, \ \mathbf{u}_i \in \mathcal{Y}_i. \quad (12b)$$

III. Design of Online OPF Solvers

The objective is to design a distributed control scheme that steers the RES-inverter setpoints $\{\mathbf{u}_i \in \mathcal{Y}_i\}_{i=1}^N$ (and, thus, the output powers $\{\mathbf{y}_i(t)\}_{i=1}^N$) to the solution of the OPF problem (12). A brief overview of ADMM-based algorithms is outlined next; the ADMM-based control architecture is then discussed in Section III-B.

A. Open-loop ADMM-based distributed optimization

Consider the following augmented Lagrangian function associated with problem (12):

$$\mathcal{L}(\boldsymbol{\Delta},\{\mathbf{u}_i\},\{\boldsymbol{\lambda}_i\}) := H(\boldsymbol{\Delta}) + \sum_{i\in\mathcal{N}_D} G_i(\mathbf{u}_i) + \frac{\rho}{2}\sum_{i\in\mathcal{N}\setminus\{0\}} \left\| \boldsymbol{\Phi}_i\boldsymbol{\Delta} - \mathbf{u}_i + \mathbf{d}_i + \frac{\boldsymbol{\lambda}_i}{\rho}\right\|^2, \quad (13)$$

where $\boldsymbol{\lambda}_i \in \mathbb{R}^N$ is the Lagrangian multiplier associated with the linear constraint (12a), $\rho > 0$ is a design parameter, and $\mathbf{u}_i = \mathbf{0}$ for $i \in \mathcal{N}_O$. ADMM involves an iterative procedure where the following steps are performed at each iteration k:

$$\boldsymbol{\Delta}^k = \arg\min_{\boldsymbol{\Delta}\in\mathcal{V}} H(\boldsymbol{\Delta}) + \sum_{i\in\mathcal{N}\setminus\{0\}} \frac{\rho}{2}\left\|\boldsymbol{\Phi}_i\boldsymbol{\Delta} - \mathbf{u}_i^{k-1} + \mathbf{d}_i - \boldsymbol{\lambda}_i^{k-1}/\rho\right\|^2, \quad (14a)$$

$$\boldsymbol{\lambda}_i^k = \boldsymbol{\lambda}_i^{k-1} - \rho(\boldsymbol{\Phi}_i\boldsymbol{\Delta}^k - \mathbf{u}_i^{k-1} + \mathbf{d}_i), \quad (14b)$$

$$\mathbf{u}_i^k = \arg\min_{\mathbf{u}_i\in\mathcal{Y}_i} G_i(\mathbf{u}_i) + \frac{\rho}{2}\left\|\boldsymbol{\Phi}_i\boldsymbol{\Delta}^k - \mathbf{u}_i + \mathbf{d}_i - \boldsymbol{\lambda}_i^k/\rho\right\|^2. \quad (14c)$$

One way to reduce the computational complexity associated with the update of the voltage-related vector $\boldsymbol{\Delta}^k$ is to consider solving the following quadratic approximation:

$$\boldsymbol{\Delta}^k = \arg\min_{\boldsymbol{\Delta}\in\mathcal{V}} \langle \mathbf{g}^{k-1}, \boldsymbol{\Delta}-\boldsymbol{\Delta}^{k-1}\rangle + \frac{L}{2}\left\|\boldsymbol{\Delta}-\boldsymbol{\Delta}^{k-1}\right\|^2, \quad (15)$$

where $L > 0$ is a design parameter, and $\mathbf{g}^{k-1}$ denotes the gradient of the augmented Lagrangian function with respect to $\boldsymbol{\Delta}$; particularly, $\mathbf{g}^{k-1}$ is given by:

$$\mathbf{g}^{k-1} = \nabla H(\boldsymbol{\Delta}^{k-1}) + \sum_{i\in\mathcal{N}_D} \boldsymbol{\Phi}_i^\top\left(\boldsymbol{\Phi}_i\boldsymbol{\Delta}^{k-1} - \mathbf{u}_i^{k-1} + \mathbf{d}_i + \boldsymbol{\lambda}_i^{k-1}/\rho\right). \quad (16)$$

It can be verified that the optimal solution of (15) amounts to a projected-gradient step in the following form:

$$\boldsymbol{\Delta}^k = \mathcal{P}_{\mathcal{V}}(\boldsymbol{\Delta}^{k-1} - \frac{1}{L}\mathbf{g}^{k-1}), \quad (17)$$

where $\mathcal{P}_{\mathcal{V}}$ denotes the projection operation onto the set $\mathcal{V}$.

The steps described above lead to a distributed procedure that is provably convergent to a solution of (12); the distributed algorithm is tabulated as Algorithm 1.

Algorithm 1 ADMM-based algorithm

1: Perform (14a) with two options:
- `Option 1`: Solve (14a).
- `Option 2`: Perform (17).

2: Perform (14b).
3: Perform (14c).

However, one drawback of Algorithm 1 is that the setpoints $\mathbf{u}_i$ can be commanded to the RES inverters only *upon convergence*. On the other hand, sending the setpoints to the RES inverters at each intermediate step k leads to an *open-loop* procedure where no actionable feedback from the electrical system is utilized; for instance, steps 2 and 3 of Algorithm 1 would utilize the commanded inputs $\{\mathbf{u}_i^k\}_{i\in\mathcal{N}_D}$, which may not necessarily coincide with the actual outputs powers of the RES inverters (commanded setpoints and output powers coincide only after a given settling time of the primary controllers of the inverters). To capture non-idealities of existing devices (which may not respond quickly to changes in the setpoints) as well as discrepancies between the input setpoints and the power outputs due to faulty estimations of the maximum available powers from the RESs, the next section will develop a control scheme that dynamically update the setpoints of the devices based on current system outputs and problem parameters. The setting is close in spirit to the feedback-control strategies proposed in e.g., [3], [4], [6]. Compared to [6] and [4], the proposed framework does not resort to barrier-type functions to enforce voltage limits and is applicable to multi-phase settings; the contribution over [3] consists in considering less stringent assumptions on the mismatch between the commanded setpoints and current system outputs, and improved convergence properties.

B. From open-loop optimization to feedback-control

Similar to, e.g., [3], [4], [6], consider updates performed at discrete time instants $t \in \{t_k, k \in \mathbb{N}\}$. At time t_k, let $\mathbf{u}^{t_k} = \{\mathbf{u}_i^{t_k}\}_{i\in\mathcal{N}\setminus\{0\}}$, $\boldsymbol{\Delta}^{t_k}$ and $\boldsymbol{\lambda}^{t_k} := \{\boldsymbol{\lambda}_i^{t_k}\}_{i\in\mathcal{N}\setminus\{0\}}$ denote the primal and dual variables, respectively. At time t_{k-1}, the RES outputs are sampled as [cf. Fig. 1]:

$$\mathbf{y}_i^{t_{k-1}} = \mathbf{r}_i(\mathbf{x}_i(t_{k-1}), \mathbf{d}_i), \forall\, i \in \mathcal{N}_D. \quad (18)$$

The measured output powers are then utilized to update the voltage-related vector $\boldsymbol{\Delta}$ and the dual variables as follows:

$$\texttt{Option 1}: \boldsymbol{\Delta}^{t_k} = \arg\min_{\boldsymbol{\Delta}\in\mathcal{V}} H(\boldsymbol{\Delta}) + \sum_{i\in\mathcal{N}\setminus\{0\}} \frac{\rho}{2}\left\|\boldsymbol{\Phi}_i\boldsymbol{\Delta} - \mathbf{y}_i^{t_{k-1}} + \mathbf{d}_i - \frac{\boldsymbol{\lambda}_i^{t_{k-1}}}{\rho}\right\|^2, \quad (19a)$$

$$\texttt{Option 2}: \boldsymbol{\Delta}^{t_k} = \mathcal{P}_{\mathcal{V}}\left(\boldsymbol{\Delta}^{t_{k-1}} - \frac{1}{L}\mathbf{g}^{t_{k-1}}\right), \quad (19b)$$

$$\boldsymbol{\lambda}_i^{t_k} = \boldsymbol{\lambda}_i^{t_{k-1}} - \rho(\boldsymbol{\Phi}_i\boldsymbol{\Delta}^{t_k} - \mathbf{y}_i^{t_{k-1}} + \mathbf{d}_i). \quad (19c)$$

$$\mathbf{u}_i^{t_k} = \arg\min_{\mathbf{u}_i\in\mathcal{Y}_i} G_i(\mathbf{u}_i) + \frac{\rho}{2}\left\|\boldsymbol{\Phi}_i\boldsymbol{\Delta}^{t_k} - \mathbf{u}_i + \mathbf{d}_i - \frac{\boldsymbol{\lambda}_i^{t_k}}{\rho}\right\|^2. \quad (19d)$$

Note that (19d) produces the commands to the RES inverter (1) [cf. Fig. 1]. Different from (17), in Option 2 the gradient vector $\mathbf{g}^{t_{k-1}}$ is evaluated at $\mathbf{y}_i^{t_{k-1}}$, i.e.,

$$\mathbf{g}^{t_{k-1}} = \nabla H(\boldsymbol{\Delta}^{t_{k-1}}) + \sum_{i\in\mathcal{N}_D} \boldsymbol{\Phi}_i^\top\left(\boldsymbol{\Phi}_i\boldsymbol{\Delta}^{t_{k-1}} - \mathbf{y}_i^{t_{k-1}} + \mathbf{d}_i + \boldsymbol{\lambda}_i^{k-1}/\rho\right).$$

In summary, the controllers perform the steps tabulated as Algorithm 2.

The algorithm (19) affords a distributed implementation. With reference to the illustrative diagram in Fig. 1, one possible distributed solution involves the following steps:

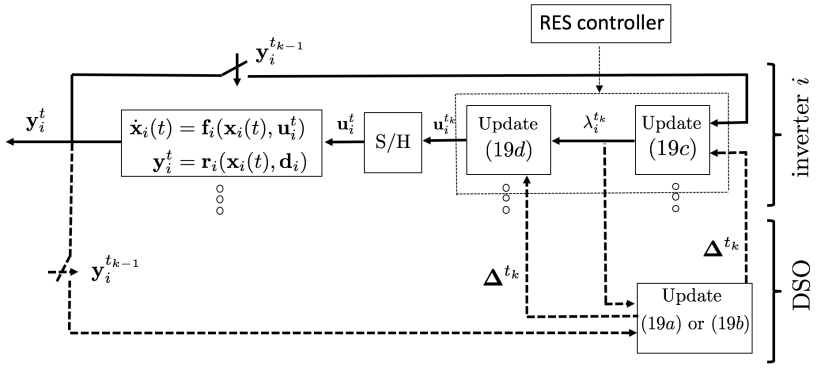


Figure 1: Illustration of the distributed framework (see also Algorithm 2). Steps. (19c) *and* (19d) *constitute the RES controller and they generate a discrete time signal* $\mathbf{u}_i^{t_k}$*, which is applied to the inverter by utilizing a sample-and-hold unit. The inverter output is then sampled and utilized to update the control signals.*

Algorithm 2 ADMM-based OPF Controllers

1: At time t_{k-1}, RES outputs are sampled as (18).
2: Perform (19a) or (19b).
3: Perform (19c) using the sampled RES outputs.
4: Perform (19d) and apply the resulting signal to (1a) during $(t_{k-1}, t_k]$, i.e. $\mathbf{u}_i(t) = \mathbf{u}^{t_k}, t \in (t_{k-1}, t_k]$.
5: When $t = t_k$ go to step 1.

i) update (19a) or (19b) can be performed at the DSO, after receiving $\mathbf{y}_i^{t_{k-1}}$ and $\boldsymbol{\lambda}_i^{t_{k-1}}$ from each RES i;

ii) the DSO subsequently broadcasts to the RESs the vector Δ^{t_k};

iii) updates (19c) and (19d) are then performed locally at each individual RES $i \in \mathcal{N}\backslash\{0\}$. These steps are computationally light and, when $G_i(\mathbf{u}_i)$ is linear or quadratic and $\mathcal{Y}_i$ is given as in (5), $\mathbf{u}_i^k$ admits a closed-form solution; see e.g., [13, Appendix B].

It is worth reiterating that the key differences compared to the open-loop optimization strategy (14) are: 1) the setpoints are commanded to the RES inverters at each time instant t_k (whereas Algorithm 1 produces setpoints only upon convergence of the ADMM); and, ii) measurements of the RES-inverter output-powers are used in the updates. The (continuous-time) reference signals $\{\mathbf{u}_i(t)\}_{i\in\mathcal{N}_D}$ produced by the controller have step changes at instants $\{t_k, k \in \mathbb{N}\}$ and are left-continuous functions that take the constant values $\{\mathbf{u}_i^{t_k}\}_{i\in\mathcal{N}_D}$ over the time interval $(t_{k-1}, t_k]$. When the interval $(t_{k-1}, t_k]$ is longer than the settling time of (1), then the RES output powers converge to the intermediate setpoints $\{\mathbf{u}_i^{t_k}\}_{i=1}^N$ at each iteration; that is, $\lim_{t\to t_k^-} \|\mathbf{y}_i^t - \mathbf{u}_i^{t_k}\| = 0$. Hence, (14) and (19) coincide, and the well-known convergence claims for the ADMM naturally apply to the present setup [14]. However, in case of slow-responding inverters, or, when the updates (19) can be performed faster than the systems' settling times, then one has that the inverter outputs may not coincide with the commanded setpoints; particularly, define the *error term* $\boldsymbol{\eta}_i^{t_k} = \mathbf{u}_i^{t_k} - \mathbf{y}_i^{t_k}$, $i \in \mathcal{N}_D$ to quantify this discrepancy. In the following, convergence of the RES output powers when $\boldsymbol{\eta}_i^{t_k} \neq \mathbf{0}$ is analyzed.

C. Convergence Analysis

Algorithm 2 can be interpreted as a variation of the ADMM with inexact primal and dual updates. To the best of our knowledge, convergence of the ADMM in this setting is not available in the prior literature. This paper considers the following two types of updates:

1) Exact minimization in the primal steps using RES output $\{\mathbf{y}_i^{t_k}\}$ [i.e., `option 1`];
2) Gradient steps are performed in the primal steps using RES output $\{\mathbf{y}_i^{t_k}\}$ [i.e., `option 2`].

In the remainder of this section, convergence of Option 2 is studied; in fact, Option 1 can be analyzed using similar techniques, but with considerably simpler steps. To simplify the notation, the superscript t_k is hereafter dropped.

Define $\eta^{t_k} := \|\mathbf{u}^{t_k} - \mathbf{y}^{t_k}\|$, and consider the following assumption.

Assumption 2. *The gradient stepsize* $\frac{1}{L} > 0$ *satisfies the following property:*

$$(L - \gamma)\mathbf{I}_{2N} - \rho\mathbf{\Phi}^\top\mathbf{\Phi} \succcurlyeq 0, \tag{20}$$

where $\mathbf{I}_{2N}$ *is the* $2N \times 2N$ *identity matrix, and* γ *denotes the Lipschitz constant of* $\nabla H(\mathbf{\Delta})$*, i.e.* $\|\nabla H(\mathbf{x}) - \nabla H(\mathbf{y})\| \geq \gamma\|\mathbf{x} - \mathbf{y}\|$*, for all* $\mathbf{x}, \mathbf{y} \in dom\ H$.

Further, assume that

$$\sum_{k=1}^{\infty} \eta^{t_k} < \infty.$$

Assumption 2 asserts that, for given loading and ambient conditions, the discrepancy between the commanded inputs and the output powers should diminish as the system reaches the AC OPF solution. From Assumption 2, it is clear that L has to be greater than or equal to the largest eigenvalue of the Hessian matrix of the augmented Lagrangian function (for a fixed ρ). In fact, from Assumption 2 it follows that:

$$L\mathbf{I}_{2N} \succcurlyeq \gamma\mathbf{I}_{2N} + \rho\mathbf{\Phi}^\top\mathbf{\Phi}.$$

Since γ is the Lipschitz constant of $\nabla H(\boldsymbol{\Delta})$, it follows that $\gamma \mathbf{I}_{2N} \succcurlyeq \nabla^2 H(\boldsymbol{\Delta})$, and hence:

$$L\mathbf{I}_{2N} \succcurlyeq \nabla^2 H(\boldsymbol{\Delta}) + \rho \boldsymbol{\Phi}^\top \boldsymbol{\Phi},$$

where the right-hand-side is the Hessian matrix of the augmented Lagrangian function.

To facilitate analysis, define the vectors $\mathbf{w}^{t_k} := [\mathbf{u}^{t_k}; \boldsymbol{\Delta}^{t_k}; \boldsymbol{\lambda}^{t_k}]$ and $\hat{\mathbf{w}}^{t_k} := [\hat{\mathbf{u}}^{t_k}; \hat{\boldsymbol{\Delta}}^{t_k}; \hat{\boldsymbol{\lambda}}^{t_k}]$, where $\mathbf{w}^{t_k}$ is the update generated by (19) (with possibly nonzero error terms $\boldsymbol{\eta}_i^{t_k} = \mathbf{u}_i^{t_k} - \mathbf{y}_i^{t_k}$), and $\hat{\mathbf{w}}^{t_k}$ is generated by the same iteration, but with zero error (i.e., $\mathbf{0} = \mathbf{u}_i^{t_k} - \mathbf{y}_i^{t_k}$). Henceforth, $\hat{\mathbf{w}}^{t_k}$ is referred to as the "error-free" iterates. Let W^* be the optimal set of (12), which is nonempty, closed, and convex.

The ultimate goal of this analysis is to prove the convergence of iterates $\mathbf{w}^{t_k}$. For the purpose of readibility, we first state the main result of this paper in the following theorem:

Theorem 1. *Suppose that Assumption 1 and 2 hold true. Then the sequence* $\mathbf{w}^{t_k}$ *generated by* (19b)-(19d) *converges to some* $\mathbf{w}^\infty \in W^*$*, where* $\mathbf{w}^\infty$ *is a cluster point of the sequence* $\{\mathbf{w}^{t_k}\}$.

To prove this result, we start from the optimality condition. Denote the set of optimality conditions of $\mathbf{w}$ as

$$D(\mathbf{w}) := \begin{pmatrix} \mathbf{u} - \mathcal{P}_{\mathcal{Y}}\{\mathbf{u} - [\nabla G(\mathbf{u}) + \boldsymbol{\lambda}]\} \\ \boldsymbol{\Delta} - \mathcal{P}_{\mathcal{V}}\{\boldsymbol{\Delta} - [\nabla H(\boldsymbol{\Delta}) - \boldsymbol{\Phi}^\top \boldsymbol{\lambda}]\} \\ \boldsymbol{\Phi}\boldsymbol{\Delta} - \mathbf{u} + \mathbf{d} \end{pmatrix},$$

$$\text{dist}(\mathbf{w}, W^*) := \min\{\|\mathbf{w} - \mathbf{z}^*\| \mid \mathbf{z}^* \in W^*\}.$$

Then it is easy to verify that the following holds:

$$\text{dist}(\mathbf{w}, W^*) = 0 \Leftrightarrow D(\mathbf{w}) = 0. \tag{21}$$

Lemma 1. *There exists a constant* $\tau > 0$*, such that*

$$\left\|D(\hat{\mathbf{w}}^{t_k})\right\|^2 \leq \tau \cdot \left\|\boldsymbol{\Phi}\boldsymbol{\Delta}^{t_{k-1}} - \hat{\mathbf{u}}^{t_k} + \mathbf{d}\right\|_{\rho\mathbf{I}}^2, \ \forall\ k \geq 1. \tag{22}$$

To show convergence, the right-hand-side of (22) needs to approach 0 when $k \to \infty$; this result is provided by the following two lemmas.

Lemma 2. *Let* $\mathbf{w}^* := [\mathbf{u}^*; \boldsymbol{\Delta}^*; \boldsymbol{\lambda}^*]$ *be an optimal solution of* (12)*. Then the following inequality holds*

$$\left\|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 \leq \left\|\mathbf{w}^{t_{k-1}} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\boldsymbol{\Phi}\boldsymbol{\Delta}^{t_{k-1}} - \hat{\mathbf{u}}^{t_k} + \mathbf{d}\right\|_{\rho\mathbf{I}}^2 - \left\|\hat{\boldsymbol{\Delta}}^{t_k} - \boldsymbol{\Delta}^{t_{k-1}}\right\|_{\boldsymbol{\Psi}}^2, \tag{23}$$

where $\tilde{\mathbf{H}} := \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & L\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{\rho}\mathbf{I} \end{pmatrix}$ *and* $\boldsymbol{\Psi} := (L - \gamma)\mathbf{I} - \rho\boldsymbol{\Phi}^\top\boldsymbol{\Phi}$.

Lemma 2 establishes a relationship between the exact and inexact updates in terms of the distance to an optimal solution.

Lemma 3. *Let* $\mathbf{w}^{t_k} = [\mathbf{u}^{t_k}; \boldsymbol{\Delta}^{t_k}; \boldsymbol{\lambda}^{t_k}]$ *and* $\hat{\mathbf{w}}^{t_k} = [\hat{\mathbf{u}}^{t_k}; \hat{\boldsymbol{\Delta}}^{t_k}; \hat{\boldsymbol{\lambda}}^{t_k}]$ *be the inexact and exact iterates, respectively. We have the following limit*

$$\lim_{k\to\infty} \left\|\boldsymbol{\Phi}\boldsymbol{\Delta}^{t_{k-1}} - \hat{\mathbf{u}}^{t_k} + \mathbf{d}\right\|_{\rho\mathbf{I}}^2 = 0. \tag{24}$$

The proofs of Theorem 1 and Lemma 2 are presented in the Appendix.

D. Relaxing the requirements on the RES outputs

In this section, Assumption 1 is relaxed to consider cases where the error sequence $\{\eta^{t_k}\}$ is no longer diminishing; this case captures scenarios where 1) the primary controllers have a steady-state regulation error (i.e., $\mathbf{y}^{t_k} \neq \mathbf{u}^{t_k}$) and/or 2) irradiance and load conditions are very fast changing.

To facilitate the design of distributed controllers in this setting, key is to consider a modified version of (12) where the subproblem solved to update $\boldsymbol{\Delta}$ is unconstrained (this particular problem structure will ensure convergence of the algorithm developed in this section). For notational simplicity, let $\text{LB}(j) \leq 0$ and $\text{UB}(j) \geq 0$ be the lower and upper bounds for Δ_j, respectively; that is:

$$\text{LB}(j) \leq \Delta_j \leq \text{UB}(j), \quad j = 1, \ldots, 2N.$$

From (12), it follows that $\boldsymbol{\Phi}\boldsymbol{\Delta} + \mathbf{d} = \mathbf{u}$, the following approximation can be utilized to express the voltage-regulation constraints as linear functions of $\mathbf{u}_i$:

$$\sum_{j=1}^{2N} |\boldsymbol{\Phi}_{i,j}| \cdot \text{LB}(j) + \mathbf{d}_i \leq \mathbf{u}_i \leq \sum_{j=1}^{2N} |\boldsymbol{\Phi}_{i,j}| \cdot \text{UB}(j) + \mathbf{d}_i. \tag{25}$$

We refer to the feasible set defined by (25) as $\mathcal{V}'$. Using (25), the approximate OPF problem becomes:

$$\min_{\boldsymbol{\Delta}, \mathbf{u}_i} \ H(\boldsymbol{\Delta}) + \sum_{i \in \mathcal{N}_D} G_i(\mathbf{u}_i) \tag{26a}$$

$$\text{s.t.} \quad \boldsymbol{\Phi}_i\boldsymbol{\Delta} - \mathbf{u}_i + \mathbf{d}_i = \mathbf{0}, i \in \mathcal{N}\backslash\{0\}, \tag{26b}$$

$$\mathbf{u}_i \in \mathcal{Y}_i \cap \mathcal{V}'. \tag{26c}$$

The algorithm (19) can be slightly modified to accommodate (26); particularly, the projection onto $\mathcal{V}$ in (19a) should be removed and a projection onto $\mathcal{V}'$ should be added in (19d). The resultant algorithm can be used to solve (26). We refer to this algorithm as Algorithm 3.

Consider then the following assumption.

Assumption 3. *Functions* $H(\boldsymbol{\Delta})$ *and* $G(\mathbf{u})$ *are strongly convex,* $\boldsymbol{\Phi} := [\boldsymbol{\Phi}_1; \cdots; \boldsymbol{\Phi}_N] \in \mathbb{R}^{2N \times 2N}$ *is full rank, and* ∇H *is Lipschitz continuous.*

Based on this assumption, one can leverage the results of [15, Theorem 3.4] to obtain the following.

Corollary 1. *If Assumption 3 holds, the iterates* $\mathbf{w}^{t_k}$ *generated by Algorithm 3 to solve the approximated problem* (26) *and "error-free" iterates* $\hat{\mathbf{w}}^{t_k}$ *satisfy the following inequality:*

$$\left\|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 \leq (1 - \delta)\left\|\mathbf{w}^{t_{k-1}} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^{t_{k-1}}\right\|_{\tilde{\mathbf{H}}}^2, \tag{27}$$

where $\delta \in (0, 1)$ *is some positive constant.*

Corollary 1 can be proved by following steps similar to [15, Theorem 3.4]. The main convergence results are established next.

Theorem 2. *Suppose that Assumption 3 holds and that there exists a constant* ϵ *such that*

$$\left\|\mathbf{y}^{t_k} - \mathbf{u}^{t_k}\right\| \leq \epsilon. \tag{28}$$

Then, the sequence $\mathbf{w}^{t_k}$ *generated by* (19) *to solve problem* (26) *satisfies* $\|\mathbf{w}^{t_k}-\mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2 \leq (\beta+\xi)^2$*, as* $k\to\infty$*, where*

$$\xi=\epsilon\sqrt{(1-\delta)(1+\theta)\times\frac{r}{1-r}},\quad \beta=\epsilon\sqrt{L\|\mathbf{\Phi}^\top\|^2+\rho},$$

are both constants and $\theta=\frac{1-\delta}{r-1+\delta}>0,\ 0<r<1$ *is some constant.*

The theorem asserts that, if $\|\mathbf{y}^{t_k}-\mathbf{u}^{t_k}\|\leq\epsilon$ for all k (which reflects the actual operation of some existing inverters), then the algorithm will converge to a ball centered around the optimal solution set of (26).

IV. Numerical experiment

Numerical results are provided to corroborate the analytical findings and demonstrate the efficacy of the proposed method. Consider the modified version of the IEEE 37-node test feeder taken from [3]; see also Fig. 2. In the OPF problem, the voltage limits are set to $V^{\min}=0.95$ pu and $V^{\max}=1.05$pu, whereas $V_0=1+\text{j}0$ pu. With reference to the node numbering utilized in [3], assume that there are 6 PV systems located at nodes $4,11,22,26,29$, and 32, and assume that the primary controllers of the PV systems are modeled as a first-order system [24]. The following ratings are assumed: $\{S_i\}_{i\in\mathcal{N}_D}=\{100,240,100,200,240,160\}$ kVA; Further, $\theta=\frac{\pi}{2}$, $P_i^{\min}=0$, and the objective functions are set to:

$$H(\mathbf{\Delta})=10\times\sum_{i=1}^{N}(\mathbf{\Delta}(i)-1)^2, \tag{29}$$

$$G_i(P_i,Q_i)=a_i(P_i^{\text{av}}-P_i)^2+b_i(P_i^{\text{av}}-P_i)+c_iQ_i^2+d_i|Q_i|, \tag{30}$$

where $H(\mathbf{\Delta})$ promotes a flat voltage profile, and $G_i(P_i,Q_i)$ penalizes real power curtailment and limits the amount of reactive power provided. As an example, the coefficients in (30) are chosen as $a_i=1,b_i=10,c_i=0.01,d_i=0.01$ for $i=1,\ldots,4$ and $a_i=1,b_i=10,c_i=0.03,d_i=0.03$ for $i=5,6$. For the ADMM-type methods, the following quantities are utilized to measure the optimality of the solutions [27]:

$$\|r_p^k\|=\|\mathbf{C}\mathbf{\Delta}^{t_k}-\mathbf{p}^{t_k}+\mathbf{p}_l\|,\ \|r_q^k\|=\|\mathbf{D}\mathbf{\Delta}^{t_k}-\mathbf{q}^{t_k}+\mathbf{q}_l\|$$
$$\|s_p^k\|=\|\mathbf{C}(\mathbf{\Delta}^{t_k}-\mathbf{\Delta}^{t_{k-1}})\|,\ \|s_q^k\|=\|\mathbf{D}(\mathbf{\Delta}^{t_k}-\mathbf{\Delta}^{t_{k-1}})\|.$$

and, for given load and ambient conditions, the algorithm terminates when quantities above are smaller than 5×10^{-4}; for the dual-subgradient method, only the first 300 iterations are plotted.

Fig. 3 shows that the ADMM-based algorithm (with either Option 1 or Option 2) converges to the optimal objective value. Dual-subgradient methods (e.g., [3]) are also convergent, but they require a significantly higher number of iterations. Each iterations of the dual-subgradient method and of the ADMM-Option 1 take a similar computational time since they both solve the $\mathbf{\Delta}$-subproblem exactly. Notice that compared to Option 1, Option 2 requires more iterations to converge; however, each iteration is computationally lighter for Option 2.

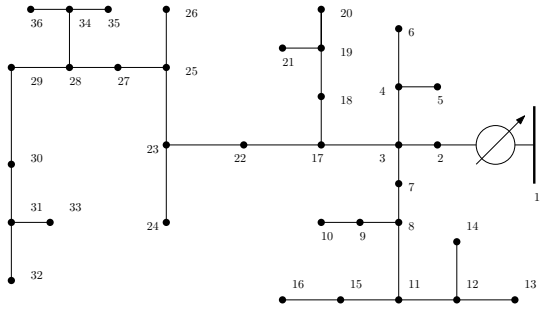


Figure 2: IEEE 37-node feeder.

This sets a natural trade-off between convergence and computational complexity. To further highlight this point, Figure 3(c) compares a few different scenarios in which *multiple* gradient steps (19b) are performed in each iteration. Clearly, the higher is the amount of gradient steps performed in each iteration, the fewer is the total iterations are required.

Next, adaptability of the proposed ADMM-based strategy to changing irradiance conditions is tested; particularly, assume the following changes in the available powers $\{P_i^{\text{av}}\}$ of the PV systems:

$$P^{\text{av}}(k)=[44,134,42,100,136,80]^\top\text{kW}, k\in[1,400]$$
$$P^{\text{av}}(k)=[50,160,48,110,170,90]^\top\text{kW}, k\in[400,600]$$
$$P^{\text{av}}(k)=[62,184,58,135,184,108]^\top\text{kW}, k\in[600,800]$$
$$P^{\text{av}}(k)=[52,168,50,114,172,94]^\top\text{kW}, k\in[800,1200].$$

Note that the changes are presupposed at iterations 400, 600 and 800. It can be seen from Fig. 4 that the inverter outputs $\mathbf{y}_i[t_k]=[P_i(t),Q_i(t)]^\top$ quickly converge to the new optimal setpoints within each interval.

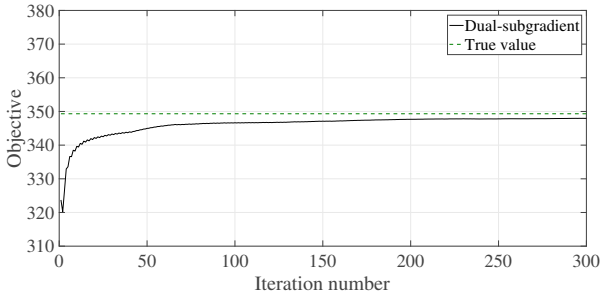
To assess whether the proposed scheme enforces voltage regulation, consider a setting where the PV capacities $\{S_i\}_{i\in\mathcal{N}_D}$ and available powers P^{av} are five time higher than the initial setting. In this case, the feeder would incur overvoltage conditions when the PV systems operate at the business-as-usual setpoint $(P^{\text{av}},0)$. Instead, Fig. 5 demonstrates that the ADMM-based controller maintains the voltage magnitude within the limits.

Finally, the ADMM-based Algorithm 3 is tested in the presence of constant error; particularly, it is assumed that $(\eta^k)^2=\|\mathbf{y}_i^{t_k}-\mathbf{u}_i^{t_k}\|^2=0.002$. The following objective functions are utilized:

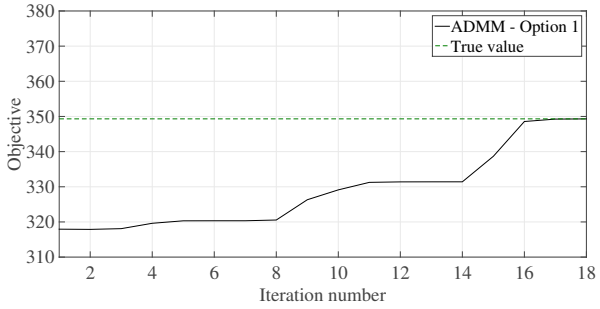
$$H(\mathbf{\Delta})=10\times\sum_{i=1}^{2N}(\mathbf{\Delta}(i)-1)^2, \tag{31}$$

$$G_i(P_i,Q_i)=a_i(P_i^{\text{av}}-P_i)^2+b_i(P_i^{\text{av}}-P_i)+c_iQ_i^2+d_i|Q_i|. \tag{32}$$

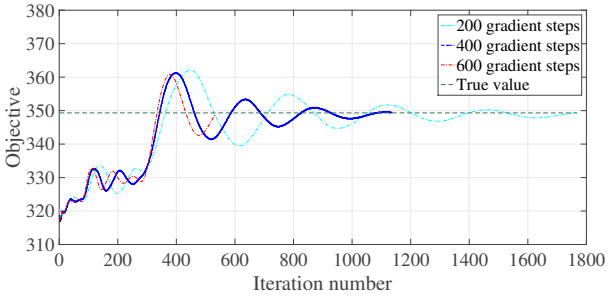
where the parameters a_i, b_i, c_i, and b_i are set as in the previous experiments. In Figure 6 it can be seen that the approximation error is negligible and, as established in Theorem 2, the algorithm converges to a neighborhood of the optimal value. In the figure, the trajectory "solving problem (12)" is used for comparison purposes, and it is generated by running (19) to solve the original problem (12).



(a) Dual-subgradient method

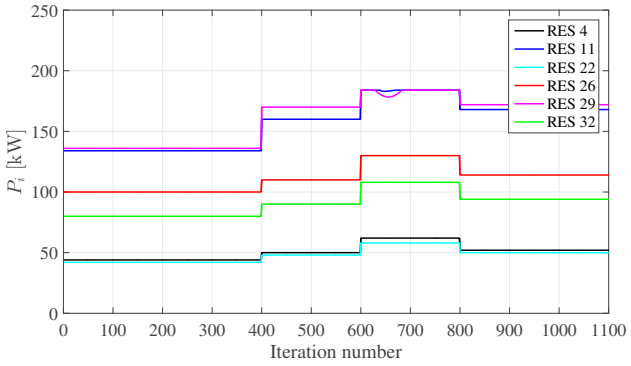


(b) ADMM with Option 1

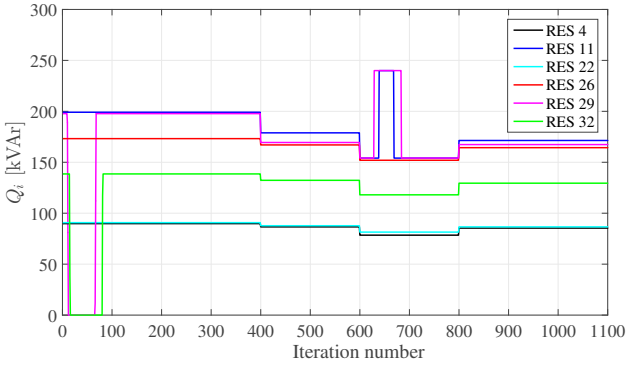


(c) ADMM with Option 2

Figure 3: Convergence of the dual-subgradient method as well as the ADMM-based algorithm. For the latter, both Option 1 and Option 2 are tested. A first-order system is used to emulate the behavior of the RES system. As a benchmark, CVX [28] is utilized to obtain the optimal solution of (12)*. Figure 3(c) illustrates the trade-off between the total number of iterations and the number of gradient steps used for each iteration*



(a) Active power



(b) Reactive power

Figure 4: Tracking performance of the proposed algorithm in case of changing operational conditions. Changes in the solar irradiance are presumed at $k = 400, 600,\ 800$.

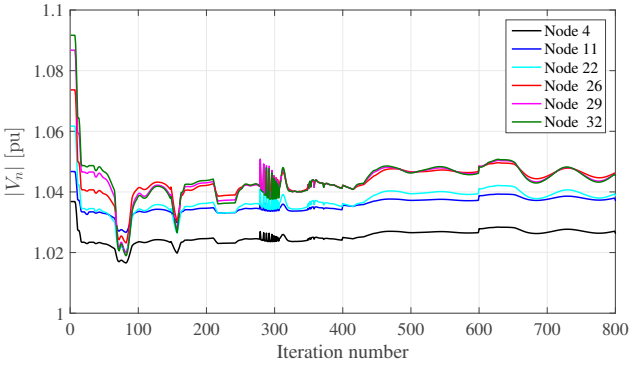


Figure 5: Example of trajectory for the voltage magnitude when the PV systems are controlled using the proposed ADMM-based algorithm.

The performance of ADMM-based algorithms depends on the tuning parameter ρ. For `Option` 1, we use the adaptive stepsize strategy explained in [27] to improve the convergence. In the `Option` 2), ρ is chosen empirically and it is set to 10^2.

The theoretical results outlined in the paper are applicable to the case where the non-controllable loads $\mathbf{d}_i$ are slow time-varying or constant. While extending the theoretical claims to the case of time-varying loads, constraints, and cost functions is the subject of future endeavors, in this section we provide some numerical results to show how the proposed ADMM-based algorithm can cope with time-varying problem parameters. To this end, we consider the simulation setting utilized in [29], where the the loads $\mathbf{d}_i$ and the maximum active powers available from the PV inverters are changing on a second basis; see [29] for a complete description of the dataset. Figure 7(a) reports the evolution of the voltage magnitude over time when CVX [28] is utilized to solve problem (12); to obtain reasonable simulation times, (12) was solved with CVX only every 1000 seconds. As for Algorithm 3, three iterations are performed every second. Figure 7(b) illustrates the trajectory of the voltage magnitudes; it can be seen that the ADMM-based method can successfully enforce voltage regulation and tracks the benchmark trajectories.

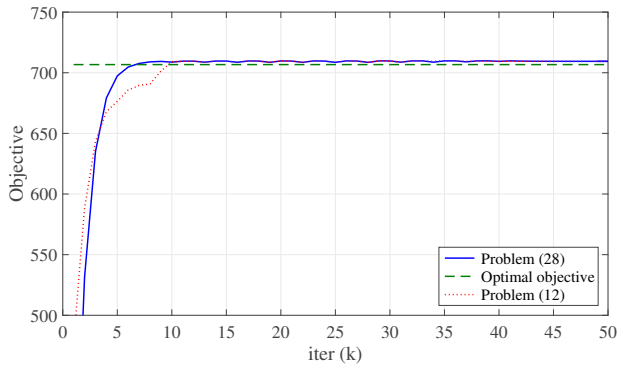
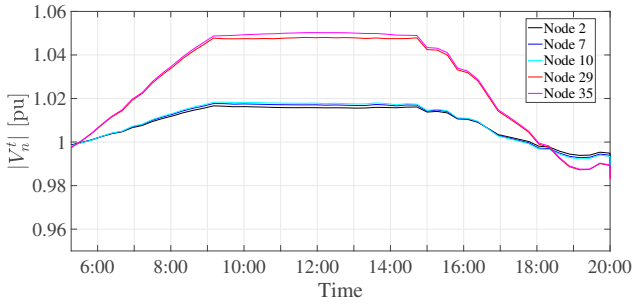
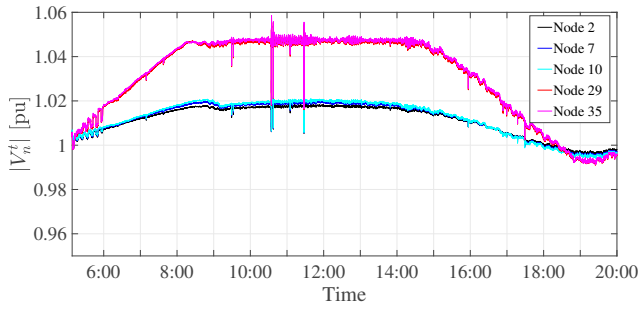


Figure 6: Convergence of ADMM with constant error in RES outputs. The two curves are generated by running (19) *to solve problem* (12)*, and by running Algorithm 3 to solve problem* (26)*, respectively.*



(a) True voltage magnitude (sampled every 1000 seconds).



(b) Voltage magnitude computed by ADMM

Figure 7: Voltage profiles when $\mathbf{d}_i$ is changing every second.

V. Conclusions and Future Work

A feedback controller for inverter-interfaced RES systems that drives the outputs of RESs to the optimal solution of convex surrogates of the AC OPF problem was developed in the paper. The design of the control strategy was based on a linear approximation of the AC power-flow equations, and the ADMM method. Convergence proprieties of the proposed ADMM-based control method are discussed under a variety of operational settings. Numerical experiments corroborated the analytical findings. Future research endeavors will look at the extension of the theoretical results to time-varying operational conditions.

Appendix A
Proof of Lemma 1

Proof: Define first the following:

$$D(\hat{\mathbf{w}}^k) = \begin{pmatrix} \hat{\mathbf{u}}^k - \mathcal{P}_{\mathcal{Y}}\{\hat{\mathbf{u}}^k - [\nabla G(\hat{\mathbf{u}}^k) + \hat{\boldsymbol{\lambda}}^k]\} \\ \hat{\boldsymbol{\Delta}}^k - \mathcal{P}_{\mathcal{V}}\{\hat{\boldsymbol{\Delta}}^k - [\nabla H(\hat{\boldsymbol{\Delta}}^k) - \boldsymbol{\Phi}^\top \hat{\boldsymbol{\lambda}}^k]\} \\ \boldsymbol{\Phi}\hat{\boldsymbol{\Delta}}^k - \hat{\mathbf{u}}^k + \mathbf{d} \end{pmatrix}$$

and notice that the optimality condition on $\hat{\boldsymbol{\Delta}}^k$ and $\hat{\mathbf{u}}^k$ imply the following:

$$\begin{aligned} \hat{\boldsymbol{\Delta}}^k &= \mathcal{P}_{\mathcal{V}}\{\hat{\boldsymbol{\Delta}}^k - [\nabla H(\hat{\boldsymbol{\Delta}}^k) - \boldsymbol{\Phi}^\top \hat{\boldsymbol{\lambda}}^{k-1} \\ &\quad + \rho \boldsymbol{\Phi}^\top (\boldsymbol{\Phi}\hat{\boldsymbol{\Delta}}^k - \hat{\mathbf{u}}^k + \mathbf{d})]\} \\ &= \mathcal{P}_{\mathcal{V}}\{\hat{\boldsymbol{\Delta}}^k - [\nabla H(\hat{\boldsymbol{\Delta}}^k) - \boldsymbol{\Phi}^\top \hat{\boldsymbol{\lambda}}^k]\} \\ \hat{\mathbf{u}}^k &= \mathcal{P}_{\mathcal{Y}}\{\hat{\mathbf{u}}^k - [\nabla G(\hat{\mathbf{u}}^k) + \hat{\boldsymbol{\lambda}}^k + \rho \boldsymbol{\Phi}(\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k)]\}. \end{aligned}$$

Combining the above equalities and using nonexpansive property of the projection operator, it follows that:

$$\begin{aligned} \left\|D(\hat{\mathbf{w}}^k)\right\| &\leq \left\| \begin{matrix} \rho\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}) \\ \mathbf{0} \\ \boldsymbol{\Phi}\boldsymbol{\Delta}^{k-1} - \hat{\mathbf{u}}^k + \mathbf{d} + \boldsymbol{\Phi}\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Phi}\boldsymbol{\Delta}^{k-1} \end{matrix} \right\| \\ &\leq (\rho+1)\left\|\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1})\right\| + \left\|\boldsymbol{\Phi}\boldsymbol{\Delta}^{k-1} - \hat{\mathbf{u}}^k + \mathbf{d}\right\| \end{aligned}$$

Thus, one can conclude that:

$$\left\|D(\hat{\mathbf{w}}^k)\right\|^2 \leq \tau \left\|\boldsymbol{\Phi}\boldsymbol{\Delta}^{k-1} - \hat{\mathbf{u}}^k + \mathbf{d}\right\|^2_{\rho\mathbf{I}},$$

where $\tau > 0$ is some constant. ■

Appendix B
Proof of Lemma 2

Proof: Based on the following equality

$$\|\mathbf{a}+\mathbf{b}\|^2 = \|\mathbf{a}\|^2 - \|\mathbf{b}\|^2 + 2(\mathbf{a}+\mathbf{b})^\top \mathbf{b} \tag{33}$$

we have that:

$$\begin{aligned} &\left\|\hat{\mathbf{w}}^k - \mathbf{w}^*\right\|^2_{\tilde{\mathbf{H}}} = \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*\right\|^2_{L\mathbf{I}} + \left\|\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^*\right\|^2_{\frac{1}{\rho}\mathbf{I}} \\ &= \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1} + \boldsymbol{\Delta}^{k-1} - \boldsymbol{\Delta}^*\right\|^2_{L\mathbf{I}} + \left\|\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1} + \boldsymbol{\lambda}^{k-1} - \boldsymbol{\lambda}^*\right\|^2_{\frac{1}{\rho}\mathbf{I}} \\ &\overset{(33)}{=} \left\|\boldsymbol{\Delta}^{k-1} - \boldsymbol{\Delta}^*\right\|^2_{L\mathbf{I}} + \left\|\boldsymbol{\lambda}^{k-1} - \boldsymbol{\lambda}^*\right\|^2_{\frac{1}{\rho}\mathbf{I}} - \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}\right\|^2_{L\mathbf{I}} \\ &\quad - \left\|\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1}\right\|^2_{\frac{1}{\rho}\mathbf{I}} + 2L(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*)^\top(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}) \\ &\quad + \frac{2}{\rho}(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^*)^\top(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1}). \end{aligned} \tag{34}$$

We then leverage the convergence results for the standard ADMM, and utilize the optimality condition for $\boldsymbol{\Delta}^*$ and $\hat{\boldsymbol{\Delta}}^k$ as well as the convexity of $H(\cdot)$ and $G(\cdot)$ to bound the cross term in (34). For $\forall \boldsymbol{\Delta} \in \mathcal{V}$ and $\forall \mathbf{u} \in \mathcal{Y}$:

$$\begin{aligned} &H(\boldsymbol{\Delta}) - H(\hat{\boldsymbol{\Delta}}^k) - (\boldsymbol{\Delta} - \hat{\boldsymbol{\Delta}}^k)^\top(\boldsymbol{\Phi}^\top\hat{\boldsymbol{\lambda}}^k + \rho\boldsymbol{\Phi}^\top(\boldsymbol{\Phi}\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Phi}\boldsymbol{\Delta}^{k-1})) \\ &\geq L(\boldsymbol{\Delta} - \hat{\boldsymbol{\Delta}}^k)^\top(\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k) - \frac{\gamma}{2}\left\|\boldsymbol{\Delta}^k - \hat{\boldsymbol{\Delta}}^k\right\|^2 \\ &H(\boldsymbol{\Delta}) - H(\boldsymbol{\Delta}^*) - (\boldsymbol{\Delta} - \boldsymbol{\Delta}^*)^\top(\boldsymbol{\Phi}^\top\boldsymbol{\lambda}^*) \geq 0 \\ &G(\mathbf{u}) - G(\hat{\mathbf{u}}^k) + (\mathbf{u} - \hat{\mathbf{u}}^k)^\top\left[\hat{\boldsymbol{\lambda}}^k - \rho(\boldsymbol{\Phi}\boldsymbol{\Delta}^{k-1} - \boldsymbol{\Phi}\hat{\boldsymbol{\Delta}}^k)\right] \geq 0 \\ &G(\mathbf{u}) - G(\mathbf{u}^*) + (\mathbf{u} - \mathbf{u}^*)^\top\boldsymbol{\lambda}^* \geq 0. \end{aligned}$$

Using the identification $\boldsymbol{\Delta} = \boldsymbol{\Delta}^*$ and $\boldsymbol{\Delta} = \hat{\boldsymbol{\Delta}}^k$ for the optimality condition of $\boldsymbol{\Delta}^*$ and $\hat{\boldsymbol{\Delta}}^k$, respectively, and doing the same for $\mathbf{u}^*$ and $\hat{\mathbf{u}}^k$, one can obtain the following inequalities:

$$\begin{aligned}&(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*)^\top(\boldsymbol{\Phi}^\top(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^*) + \rho\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))\\ \geq& L(\boldsymbol{\Delta}^* - \hat{\boldsymbol{\Delta}}^k)^\top(\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k) - \frac{\gamma}{2}\left\|\boldsymbol{\Delta}^k - \hat{\boldsymbol{\Delta}}^k\right\|^2 \quad (35)\\ &(\mathbf{u}^* - \hat{\mathbf{u}}^k)^\top(\boldsymbol{\lambda}^* - \hat{\boldsymbol{\lambda}}^k)\\ \leq& \rho(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\mathbf{u}^* - \hat{\mathbf{u}}^k). \quad (36)\end{aligned}$$

thus, adding up (35)-(36), one has that:

$$\begin{aligned}&(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^*)^\top(\boldsymbol{\Phi}\boldsymbol{\Delta}^* - \boldsymbol{\Phi}\hat{\boldsymbol{\Delta}}^k + \hat{\mathbf{u}}^k - \mathbf{u}^*)\\ &+ (\boldsymbol{\Delta}^* - \hat{\boldsymbol{\Delta}}^k)^\top(\rho\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))\\ \leq& \frac{\gamma}{2}\left\|\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k\right\|^2 - L(\boldsymbol{\Delta}^* - \hat{\boldsymbol{\Delta}}^k)^\top(\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k)\\ &+ \rho(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}))^\top(\mathbf{u}^* - \hat{\mathbf{u}}^k).\end{aligned}$$

Using the dual update of (19) in the above inequality, one obtains:

$$\begin{aligned}&\frac{1}{\rho}(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1})(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^*)\\ \leq& \frac{\gamma}{2}\left\|\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k\right\|^2 + \rho(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\mathbf{u}^* - \hat{\mathbf{u}}^{k+1})\\ &+ (\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*)^\top(L\mathbf{I} - \rho\boldsymbol{\Phi}^\top\boldsymbol{\Phi})(\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k).\end{aligned}$$

Now we can bound the cross term of (34) as follows:

$$\begin{aligned}&2L(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*)^\top(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}) + \frac{2}{\rho}(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^*)^\top(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1})\\ \leq& 2L(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*)^\top(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}) + \gamma\left\|\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k\right\|^2\\ &+ (\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*)^\top(2L\mathbf{I} - 2\rho\boldsymbol{\Phi}^\top\boldsymbol{\Phi})(\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k)\\ &+ 2\rho(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\mathbf{u}^* - \hat{\mathbf{u}}^k)\\ =& \gamma\left\|\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k\right\|^2 + (\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*)^\top(2\rho\boldsymbol{\Phi}^\top\boldsymbol{\Phi})(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1})\\ &+ 2\rho(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\mathbf{u}^* - \hat{\mathbf{u}}^k)\\ =& \gamma\left\|\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k\right\|^2 + 2\rho(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^*)\\ &+ \mathbf{u}^* - \hat{\mathbf{u}}^k)\\ =& \gamma\left\|\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k\right\|^2 + 2\rho(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\boldsymbol{\Phi}\hat{\boldsymbol{\Delta}}^k - \hat{\mathbf{u}}^k + \mathbf{d})\\ =& \gamma\left\|\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k\right\|^2 - 2(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1}). \quad (37)\end{aligned}$$

Combine (34)-(37) and define $\boldsymbol{\Psi} := (L-\gamma)\mathbf{I} - \rho\boldsymbol{\Phi}^\top\boldsymbol{\Phi}$. Then, $\left\|\hat{\mathbf{w}}^k - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2$ can be bounded as shown next:

$$\begin{aligned}\left\|\hat{\mathbf{w}}^k - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 \leq& \left\|\mathbf{w}^{k-1} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}\right\|_{L\mathbf{I}}^2\\ &- \left\|\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1}\right\|_{\frac{1}{\rho}\mathbf{I}}^2 + \gamma\left\|\boldsymbol{\Delta}^{k-1} - \hat{\boldsymbol{\Delta}}^k\right\|^2\\ &- 2(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1})\\ =& \left\|\mathbf{w}^{k-1} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}\right\|_{\rho\boldsymbol{\Phi}^\top\boldsymbol{\Phi}}^2 - \left\|\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1}\right\|_{\frac{1}{\rho}\mathbf{I}}^2\\ &- 2(\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}))^\top(\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1}) - \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}\right\|_{\boldsymbol{\Psi}}^2\\ =& \left\|\mathbf{w}^{k-1} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\hat{\boldsymbol{\lambda}}^k - \boldsymbol{\lambda}^{k-1} + \rho\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1})\right\|_{\frac{1}{\rho}\mathbf{I}}^2\\ &- \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}\right\|_{\boldsymbol{\Psi}}^2\\ =& \left\|\mathbf{w}^{k-1} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}\right\|_{\boldsymbol{\Psi}}^2\\ &- \left\|-\rho(\boldsymbol{\Phi}\hat{\boldsymbol{\Delta}}^k - \hat{\mathbf{u}}^k + \mathbf{d}) + \rho\boldsymbol{\Phi}(\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1})\right\|_{\frac{1}{\rho}\mathbf{I}}^2\\ =& \left\|\mathbf{w}^{k-1} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\boldsymbol{\Phi}\boldsymbol{\Delta}^{k-1} - \hat{\mathbf{u}}^k + \mathbf{d}\right\|_{\rho\mathbf{I}}^2\\ &- \left\|\hat{\boldsymbol{\Delta}}^k - \boldsymbol{\Delta}^{k-1}\right\|_{\boldsymbol{\Psi}}^2 \quad (38)\end{aligned}$$

■

APPENDIX C
PROOF OF LEMMA 3

Proof: It can be readily shown that

$$\begin{aligned}\|\mathbf{w}^{t_k} - \mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2 &= \|\mathbf{w}^{t_k} - \hat{\mathbf{w}}^{t_k} + \hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2\\ &= \|\mathbf{w}^{t_k} - \hat{\mathbf{w}}^{t_k}\|_{\tilde{\mathbf{H}}}^2 + \|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2\\ &\quad + 2\|\mathbf{w}^{t_k} - \hat{\mathbf{w}}^{t_k}\|_{\tilde{\mathbf{H}}} \cdot \|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\|_{\tilde{\mathbf{H}}}. \quad (39)\end{aligned}$$

On the other hand, from Lemma 2 it follows that

$$\begin{aligned}\left\|\mathbf{w}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}} &\leq \left\|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}} + \|\mathbf{w}^{t_k} - \hat{\mathbf{w}}^{t_k}\|_{\tilde{\mathbf{H}}}\\ &\leq \left\|\mathbf{w}^{t_{k-1}} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}} + \|\mathbf{w}^{t_k} - \hat{\mathbf{w}}^{t_k}\|_{\tilde{\mathbf{H}}}.\end{aligned}$$

According to Hölder's inequality and the fact that there are errors both in Δ and λ updates, we have

$$\begin{aligned}\|\mathbf{w}^{t_k} - \hat{\mathbf{w}}^{t_k}\|_{\tilde{\mathbf{H}}}^2 &= \left\|\begin{pmatrix}0\\ \boldsymbol{\Delta}^{t_k} - \hat{\boldsymbol{\Delta}}^{t_k}\\ \lambda^{t_k} - \hat{\lambda}^{t_k}\end{pmatrix}\right\|_{\tilde{\mathbf{H}}}^2 = \left\|\begin{pmatrix}0\\ \boldsymbol{\Phi}^\top(\mathbf{u}^{t_k} - \mathbf{y}^{t_k})\\ \rho(\mathbf{u}^{t_k} - \mathbf{y}^{t_k})\end{pmatrix}\right\|_{\tilde{\mathbf{H}}}^2\\ &\leq L(\eta^{t_k})^2\|\boldsymbol{\Phi}^\top\|^2 + \rho(\eta^{t_k})^2. \quad (40)\end{aligned}$$

Combining the above two inequalities we can derive

$$\left\|\mathbf{w}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}} \leq \left\|\mathbf{w}^{t_{k-1}} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}} + \eta^{t_k}\sqrt{L\|\boldsymbol{\Phi}^\top\|^2 + \rho}. \quad (41)$$

Summing both sides over k, we obtain

$$\left\|\mathbf{w}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}} \leq \sum_{i=1}^{k}\sigma\eta^{t_i} \quad (42)$$

where $\sigma := \sqrt{L\|\boldsymbol{\Phi}^\top\|^2 + \rho}$. The above inequality implies that if $\sum_{k=1}^{\infty}\eta^{t_k} < +\infty$, then $\|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\|_{\tilde{\mathbf{H}}} \leq c$, where c is some constant. Consequently, combining (39) and (42) one can obtain the following:

$$\left\|\mathbf{w}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 \leq \left\|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 + (\sigma\eta^{t_k})^2 + 2\sigma\eta^{t_k}c. \quad (43)$$

Combining (43) with Lemma 2 and Assumption 1, it follows that:

$$\begin{aligned}\left\|\mathbf{w}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 &\leq \left\|\hat{\mathbf{w}}^{t_k} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 + (\sigma\eta^k)^2 + 2\sigma\eta^{t_k}c\\ &\leq \left\|\mathbf{w}^{t_{k-1}} - \mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\boldsymbol{\Phi}\boldsymbol{\Delta}^{t_{k-1}} - \hat{\mathbf{u}}^{t_k} + \mathbf{d}\right\|_{\rho\mathbf{I}}^2\\ &\quad + (\sigma\eta^{t_k})^2 + 2\sigma\eta^{t_k}c. \quad (44)\end{aligned}$$

Summing (44) from 1 to k, we obtain:

$$\left\|\mathbf{w}^{t_k}-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 \le \left\|\mathbf{w}^0-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \sum_{i=1}^{k}\left\|\mathbf{\Phi}\mathbf{\Delta}^{t_{i-1}}-\hat{\mathbf{u}}^{t_i}+\mathbf{d}\right\|_{\rho\mathbf{I}}^2 + \sum_{i=1}^{k}(\sigma\eta^{t_i})^2 + 2\sum_{i=1}^{k}\sigma\eta^{t_i}c. \quad (45)$$

Further, letting $k\to\infty$ for (45), it is clear that $\|\mathbf{w}^{t_\infty}-\mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2$ and $\left\|\mathbf{w}^0-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2$ are finite. On the other hand, with Assumption 2 we know that $\sum_{i=1}^{\infty}(\sigma\eta^{t_i})^2$ and $2\sum_{i=1}^{\infty}\sigma\eta^{t_i}c$ are also finite. Then one can show that

$$\sum_{i=1}^{+\infty}\left\|\mathbf{\Phi}\mathbf{\Delta}^{t_{i-1}}-\hat{\mathbf{u}}^{t_i}+\mathbf{d}\right\|_{\rho\mathbf{I}}^2 < +\infty,$$

which leads to the following result:

$$\lim_{k\to\infty}\left\|\mathbf{\Phi}\mathbf{\Delta}^{t_{k-1}}-\hat{\mathbf{u}}^{t_k}+\mathbf{d}\right\|_{\rho\mathbf{I}}^2 = 0. \quad (46)$$

■

APPENDIX D
PROOF OF THEOREM 1

Proof: From Lemma 1 and (46), it follows that

$$\lim_{k\to\infty} D(\hat{\mathbf{w}}^k) = 0\,.$$

On the other hand, from Assumption 1 and (45), one has that $\{\mathbf{w}^k\}$ is bounded and so is $\{\hat{\mathbf{w}}^k\}$. Then, there exists a closed and bounded set S such that $\{\hat{\mathbf{w}}^k\}\subset S$, $\lim_{k\to\infty}\|\hat{\mathbf{w}}^k-\mathbf{w}^k\|=0$. It remains to show that $\lim_{k\to\infty}\text{dist}(\hat{\mathbf{w}}^k, W^*)=0$.

Suppose that $\lim_{k\to\infty}\text{dist}(\hat{\mathbf{w}}^k, W^*)\neq 0$. Then, there exists a $\delta>0$ such that $\limsup_{k\to\infty}\text{dist}(\hat{\mathbf{w}}^k, W^*)=\delta>0$. Further,

$$\{\hat{\mathbf{w}}^k\}\subset S\cap\{\mathbf{z}|\text{dist}(\mathbf{z},W^*)\ge\frac{\delta}{2}\}\triangleq S_1,$$

and, since $S_1\cap W^*\neq\emptyset$, then $D(\mathbf{z})\neq 0$ for many $\mathbf{z}\in S_1$; that is, $\min_{\mathbf{z}\in S}\{\|D(\mathbf{z})\|^2\}=\epsilon>0$. This contradicts the fact that $\{\hat{\mathbf{w}}^k\}\subset S$ and $\lim_{k\to\infty}\|D(\hat{\mathbf{w}}^k)\|=0$.

Since $\lim_{k\to\infty}\text{dist}(\mathbf{w}^k, W^*)=0$, every subsequence of $\{\mathbf{w}^k\}$ converges to an optimal solution. Without loss of generality, let $\mathbf{w}^\infty$ be a cluster point of $\{\mathbf{w}^k\}$, and $\{\mathbf{w}^{k_j}\}$ be the subsequence of $\{\mathbf{w}^k\}$, which converges to $\mathbf{w}^\infty$. For $\mathbf{w}^\infty\in W^*$ and for all $\epsilon>0$, there exists an integer l such that:

$$\left\|\mathbf{w}^{k_l}-\mathbf{w}^\infty\right\|_{\tilde{\mathbf{H}}}^2 < \frac{\epsilon^2}{3},$$
$$\sum_{i=k_l}^{\infty}\sigma\eta_i < \frac{\epsilon^2}{6c},\quad \sum_{i=k_l}^{\infty}(\sigma\eta_i)^2 < \frac{\epsilon^2}{3}.$$

By (44), we have that $\forall k\ge k_l+1$

$$\left\|\mathbf{w}^k-\mathbf{w}^\infty\right\|_{\tilde{\mathbf{H}}}^2 \le \left\|\mathbf{w}^{k_l}-\mathbf{w}^\infty\right\|_{\tilde{\mathbf{H}}}^2 + \sum_{i=k_l}^{k-1}(\sigma\eta_i)^2 + \sum_{i=k_l}^{k-1}2c\sigma\eta_i - \sum_{i=k_l}^{k-1}\left\|\mathbf{\Phi}\mathbf{\Delta}^{i-1}-\hat{\mathbf{u}}^i+\mathbf{d}\right\|_{\rho\mathbf{I}}^2 \le \frac{\epsilon^2}{3}+\frac{\epsilon^2}{3}+\frac{\epsilon^2}{3}=\epsilon^2$$

Hence, $\|\mathbf{w}^k-\mathbf{w}^\infty\|_{\tilde{\mathbf{H}}}\to 0$, i.e. $\mathbf{w}^k\to\mathbf{w}^\infty$. ■

APPENDIX E
PROOF OF THEOREM 2

Proof: Given underlying assumptions, one can always find a $\theta=\frac{1-\delta}{r-1+\delta}>0$ such that

$$(1-\delta)(1+\frac{1}{\theta})=r<1,$$

where r is some constant. Since $\left\|\mathbf{y}^k-\mathbf{u}^k\right\|\le\epsilon$, then it holds that

$$\left\|\mathbf{w}^k-\hat{\mathbf{w}}^k\right\|_{\tilde{\mathbf{H}}}\le\epsilon\sqrt{L\|\mathbf{\Phi}^\top\|^2+\rho}. \quad (47)$$

For simplicity, define $\beta=\epsilon\sqrt{L\|\mathbf{\Phi}^\top\|^2+\rho}$. From Corollary 1, we have the following inequality:

$$\left\|\hat{\mathbf{w}}^k-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 \le (1-\delta)(1+\theta)\left\|\mathbf{w}^{k-1}-\hat{\mathbf{w}}^{k-1}\right\|_{\tilde{\mathbf{H}}}^2 + (1-\delta)(1+\frac{1}{\theta})\left\|\hat{\mathbf{w}}^{k-1}-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\hat{\mathbf{w}}^k-\mathbf{w}^{k-1}\right\|_{\tilde{\mathbf{H}}}^2$$

and, by applying the same iteration to $\|\hat{\mathbf{w}}^{k-1}-\mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2$, we obtain:

$$\begin{aligned}\left\|\hat{\mathbf{w}}^k-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 &\le (1-\delta)(1+\theta)\beta^2 + r\left\|\hat{\mathbf{w}}^{k-1}-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\hat{\mathbf{w}}^k-\mathbf{w}^{k-1}\right\|_{\tilde{\mathbf{H}}}^2\\ &\le (1-\delta)(1+\theta)\beta^2 + r[(1-\delta)(1+\theta)\beta^2 + r\left\|\hat{\mathbf{w}}^{k-2}-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2 - \left\|\hat{\mathbf{w}}^{k-1}-\mathbf{w}^{k-2}\right\|_{\tilde{\mathbf{H}}}^2] - \left\|\hat{\mathbf{w}}^k-\mathbf{w}^{k-1}\right\|_{\tilde{\mathbf{H}}}^2\\ &\cdots\\ &\le (1-\delta)(1+\theta)\beta^2\sum_{i=0}^{k-1}r^i + r^k\left\|\hat{\mathbf{w}}^{k-2}-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2.\end{aligned}$$

Letting $k\to\infty$, one has that $\lim_{k\to\infty}\sum_{i=0}^{k-1}r^i=\frac{r}{1-r}$ is a constant and

$$r^k\left\|\hat{\mathbf{w}}^{k-2}-\mathbf{w}^*\right\|_{\tilde{\mathbf{H}}}^2\to 0\,.$$

It thus follow that:

$$\|\hat{\mathbf{w}}^\infty-\mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2\le(1-\delta)(1+\theta)\beta^2\times\frac{r}{1-r}$$

Let $\xi^2=(1-\delta)(1+\theta)\beta^2\times\frac{r}{1-r}$. Since $\|\hat{\mathbf{w}}^k-\mathbf{w}^k\|_{\tilde{\mathbf{H}}}^2\le\beta^2$ is a constant, it follows that

$$\begin{aligned}\|\mathbf{w}^\infty-\mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2 &= \|\mathbf{w}^\infty-\hat{\mathbf{w}}^\infty+\hat{\mathbf{w}}^\infty-\mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2\\ &= \|\mathbf{w}^\infty-\hat{\mathbf{w}}^\infty\|_{\tilde{\mathbf{H}}}^2+\|\hat{\mathbf{w}}^\infty-\mathbf{w}^*\|_{\tilde{\mathbf{H}}}^2 + 2\|\mathbf{w}^\infty-\hat{\mathbf{w}}^\infty\|_{\tilde{\mathbf{H}}}\cdot\|\hat{\mathbf{w}}^\infty-\mathbf{w}^*\|_{\tilde{\mathbf{H}}}\\ &\le \beta^2+\xi^2+2\beta\xi=(\beta+\xi)^2\end{aligned}$$

■

REFERENCES

[1] Y. Liu, J. Bebic, B. Kroposki, J. de Bedout, and W. Ren, "Distribution system voltage performance analysis for high-penetration PV," in *IEEE Energy 2030 Conf.*, Nov. 2008.

[2] A. Woyte, V. Van Thong, R. Belmans, and J. Nijs, "Voltage fluctuations on distribution level introduced by photovoltaic systems," *IEEE Trans. on Energy Conv.*, vol. 21, no. 1, pp. 202–209, 2006.

[3] E. Dall'Anese, S. V. Dhople, and G. B. Giannakis, "Photovoltaic inverter controller seeking AC optimal power flow solutions," *IEEE Trans. on Power Systems*, vol. 31, no. 4, pp. 2809–2823, 2016.

[4] L. Gan and S. H. Low, "An online gradient algorithm for optimal power flow in radial networks," *IEEE J. on Sel. Areas in Commun.*, 2016, to appear.

[5] A. Bernstein, L. Reyes-Chamorro, J. Le Boudec, and M. Paolone, "A composable method for real-time control of active distribution networks with explicit power setpoints. part I: Framework," *Electric Power Systems Research*, vol. 125, pp. 254 – 264, 2015.

[6] A. Bernstein, N. J. Bouman, and J.-Y. Le Boudec, "Design of resource agents with guaranteed tracking properties for real-time control of electrical grids," 2015, [Online] Available at: http://arxiv.org/abs/1511.08628.

[7] A. Jokić, M. Lazar, and P. Van den Bosch, "Real-time control of power systems using nodal prices," *Intl. J. of Electrical Power & Energy Systems*, vol. 31, no. 9, pp. 522–530, 2009.

[8] K. Hirata, J. P. Hespanha, and K. Uchida, "Real-time pricing leading to optimal operation under distributed decision makings," in *Proc. of American Control Conf.*, Portland, OR, June 2014.

[9] A. Cherukuri and J. Cortes, "Distributed coordination of ders with storage for dynamic economic dispatch," 2016, [Online] Available at: https://arxiv.org/abs/1605.00721.

[10] X. Ma and N. Elia, "A distributed continuous-time gradient dynamics approach for the active power loss minimizations," in *Proc. of 51st Annual Allerton Conf. on Commun., Contr., and Comp.*, UIUC, IL, USA, Oct. 2013.

[11] D. B. Arnold, M. Negrete-Pincetic, M. D. Sankur, D. M. Auslander, and D. S. Callaway, "Model-free optimal control of VAR resources in distribution systems: An extremum seeking approach," *IEEE Trans. on Power Systems*, 2015.

[12] G. Wang, V. Kekatos, A. J. Conejo, and G. B. Giannakis, "Ergodic energy management leveraging resource variability in distribution grids," *IEEE Transactions on Power Systems*, 2016.

[13] E. Dall'Anese and A. Simonetto, "Optimal power flow pursuit," *arXiv preprint arXiv:1601.07263*, 2016.

[14] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice hall Englewood Cliffs, NJ, 1989, vol. 23.

[15] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," *Journal of Scientific Computing*, pp. 1–28, 2012.

[16] M. Hong and Z.-Q. Luo, "On the linear convergence of the alternating direction method of multipliers," *arXiv preprint arXiv:1208.3922*, 2012.

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

[18] K. Christakou, J.-Y. Le Boudec, M. Paolone, and D.-C. Tomozei, "Efficient Computation of Sensitivity Coefficients of Node Voltages and Line Currents in Unbalanced Radial Electrical Distribution Networks," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 741–750, 2013.

[19] S. Guggilam, E. Dall'Anese, Y. Chen, S. Dhople, and G. B. Giannakis, "Scalable optimization methods for distribution networks with high PV integration," *IEEE Trans. on Smart Grid*, 2016, to appear.

[20] S. Dhople, S. Guggilam, and Y. Chen, "Linear approximations to AC power flow in rectangular coordinates," in *Proc. of 53rd Annual Allerton Conf. on Commun., Control, and Comp.*, Monticello, IL, Oct. 2015.

[21] S. Bolognani and F. Dörfler, "Fast power system analysis via implicit linearization of the power flow manifold," in *Proc. of 53rd Annual Allerton Conf. on Commun., Control, and Comp.*, Monticello, IL, Oct. 2015.

[22] Y. Zhang, M. Hong, E. Dall'Anese, S. Dhople, and X. Zu, "Regulation of renewable energy sources to optimal power flow solutions using ADMM," in *American control Conference*, May 2017.

[23] A. Jokić, M. Lazar, and P. P. Van den Bosch, "On constrained steady-state regulation: dynamic KKT controllers," *IEEE Trans. Auto. Contr.*, vol. 54, no. 9, pp. 2250–2254, Sep. 2009.

[24] A. Yazdani and R. Iravani, *Voltage-sourced converters in power systems: modeling, control, and applications*. John Wiley & Sons, 2010.

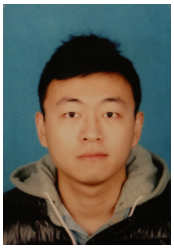
[25] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[26] S. H. Low, "Convex relaxation of optimal power flow part i: Formulations and equivalence," *IEEE Trans. on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, March 2014.

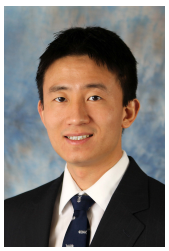
[27] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[28] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," 2008.

[29] E. Dall'Anese and S. Simonetto, "Optimal power flow pursuit," *IEEE Trans. on Smart Grid.*, 2016, [Online] Available at: http://arxiv.org/abs/1601.07263.



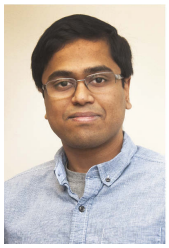
Yijian Zhang received his B.S. degree in Information and Computing Science and M.S. degree in Operations Research both from Beijing University of Technology, China, in 2012 and 2015, respectively. He is currently a Ph.D. student with Department of Industrial and Manufacturing Systems Engineering (IMSE), Iowa State University, Ames. His research interests include nonlinear programming, distributed algorithms design for large-scale optimization problems and their applications in machine learning, power systems etc.



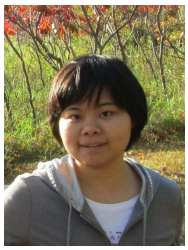
Mingyi Hong (M?11) received his B.E. degree in Communications Engineering from Zhejiang University, China, in 2005, his M.S. degree in Electrical Engineering from Stony Brook University in 2007, and Ph.D. degree in Systems Engineering from University of Virginia in 2011. From 2011 to 2014 he has been a Post Doctoral Fellow and a Research Assistant Professor with the Department of Electrical and Computer Engineering, University of Minnesota, Twin Cities. He is currently an Assistant Professor with the Department of Industrial and Manufacturing Systems Engineering (IMSE), Iowa State University, Ames. His research interests include signal processing, wireless communications, large-scale optimization and its applications in compressive sensing, complex networks and high-dimensional data analysis.

Emiliano Dall'Anese (S'08-M'11) received the Laurea Triennale (B.Sc Degree) and the Laurea Specialistica (M.Sc Degree) in Telecommunications Engineering from the University of Padova, Italy, in 2005 and 2007, respectively, and the Ph.D. in Information Engineering from the Department of Information Engineering, University of Padova, Italy, in 2011. From January 2009 to September 2010, he was a visiting scholar at the Department of Electrical and Computer Engineering, University of Minnesota, USA. From January 2011 to November 2014 he was a Postdoctoral Associate at the Department of Electrical and Computer Engineering and Digital Technology Center of the University of Minnesota, USA. Since December 2014 he has been a Senior Engineer at the National Renewable Energy Laboratory, Golden, CO, USA.

His research interests lie in the areas of optimization, signal processing, power systems, and communications. Current efforts focus on distributed optimization and control of power distribution systems with distributed (renewable) energy resources, and statistical inference for grid data analytics.



Sairaj V. Dhople (S'09-M'13) received the B.S., M.S., and Ph.D. degrees in electrical engineering, in 2007, 2009, and 2012, respectively, from the University of Illinois, Urbana-Champaign. He is currently an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Minnesota (Minneapolis), where he is affiliated with the Power and Energy Systems research group. His research interests include modeling, analysis, and control of power electronics and power systems with a focus on renewable integration. Dr. Dhople received the National Science Foundation CAREER Award in 2015. He currently serves as an Associate Editor for the IEEE Transactions on Energy Conversion.



Zi Xu received the B.Sc. degree in applied mathematics from the Hunan Normal University, Hunan, China, in 2003. She then studied nonlinear optimization in the Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, and received the Ph.D. degree in optimization in 2008. After her graduation, she has been with the Department of Mathematics in Shanghai University, Shanghai, China, and became an Associate Professor in 2012. Her research interests include optimization algorithm, complexity analysis and various optimization applications.