Optimal Regulation of Virtual Power Plants

Emiliano Dall'Anese, *Member, IEEE*, Swaroop Guggilam, *Student Member, IEEE*, Andrea Simonetto, *Member, IEEE*, Yu Christine Chen, *Member, IEEE*, and Sairaj V. Dhople, *Member, IEEE*

***Abstract*—This paper develops a real-time algorithmic framework for aggregations of distributed energy resources (DERs) in distribution networks to provide regulation services in response to transmission-level requests. Leveraging online primal-dual-type methods for time-varying optimization problems and suitable linearizations of the nonlinear AC power-flow equations, we believe this work establishes a system-theoretic foundation to realize the vision of distribution-level *virtual power plants*. The optimization framework controls the output powers of dispatchable DERs such that, in aggregate, they respond to automatic generation control and/or regulation-services commands. This is achieved while concurrently regulating voltages within the feeder and maximizing customers' and utility's performance objectives. Convergence and tracking capabilities are analytically established under suitable modeling assumptions. Simulations are provided to validate the proposed approach.**

Distribution systems, virtual power plants, real-time optimization, optimization with feedback.

I. Introduction

Traditional approaches for regulating frequency and maintaining reliable operation of transmission systems leverage primary frequency response, automatic generation control (AGC), and regulation services provided by large-scale synchronous generators. In the future, dispatchable distributed energy resources (DERs) are envisioned to supplement generation-side capabilities, by providing additional flexibility in regulating frequency and maintaining reliable system operation [1], [2]. Towards realizing this vision, we develop an algorithmic framework for DER aggregations in distribution feeders to emulate a *virtual power plant* that effectively provides regulation services to the bulk system while guaranteeing power quality in the distribution network.

The main idea and technical approach are outlined with respect to the illustrative system in Fig. 1. The objective is to develop a real-time optimization architecture for DERs, so that the active power at the *feeder head*, P_0, is adjusted in real time to track setpoint $P_{0,\text{set}}$ (while we focus on active power, the framework can be extended to consider reactive-power setpoints too). For example, $P_{0,\text{set}}$ can be an AGC signal (scaled by a given feeder participation factor), a ramping signal, or a 5-minute dispatch commanded by the transmission system operator (e.g., flexible ramping products) [2]. The real-time algorithm is designed to track the setpoint $P_{0,\text{set}}$ at the feeder head, while concurrently: i) maximizing customers' and utility's performance objectives, and ii) ensuring that operational limits are enforced throughout the feeder. The algorithm is developed with the aid of online primal-dual-gradient methods applied to double-smoothed Lagrangian functions [3], [4] while relying on suitable linear approximations of the AC power-flow equations to bypass their nonlinearity. As shown in Fig. 1, the resultant operational strategy involves collecting measurements of pertinent voltages and powers in real time. These measurements and the setpoint $P_{0,\text{set}}$ are then utilized to dispatch individual DERs. Convergence and tracking capabilities of the proposed algorithms are analytically established.

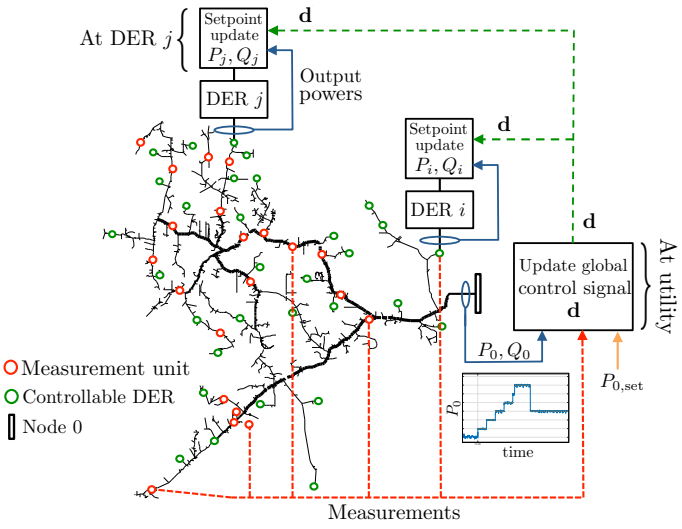

Fig. 1. Proposed architecture: output powers of the DERs (P_i's, Q_i's) are regulated in real time so that the active power at the *feeder head*, P_0 follows the setpoint reference $P_{0,\text{set}}$ while maximizing customers' and utility's performance objectives and ensuring that operational limits are enforced.

Prior works in this context have considered controlling aggregations of DERs such as thermostatically controlled loads [5]–[7] and other deferrable loads (e.g., pool pumps) [8] to track given power setpoints at the substation. A control framework for scheduling and provisioning of frequency reserves by aggregations of commercial buildings is proposed in [9]. Strategies to manage fleets of electric vehicles to provide services to the main grid are investigated in [10]. A framework for modeling of flexible loads as virtual batteries

Submitted: December 7, 2016. Revised: April 22, 2017. Re-revised: June 26, 2017. Accepted: August 15, 2017.

E. Dall'Anese is with the National Renewable Energy Laboratory, Golden, CO, USA. S. Guggilam and S. V. Dhople are with the University of Minnesota, Minneapolis, MN, USA. A Simonetto is with IBM Research Ireland, Dublin, Ireland. Y. C. Chen is with The University of British Columbia, Vancouver, Canada.

This work was supported by the U.S. Department of Energy under Contract No. DE-AC36-08GO28308 with the National Renewable Energy Laboratory. The efforts of E. Dall'Anese, S. Guggilam, and S. V. Dhople were supported by the Advanced Research Projects Agency-Energy (ARPA-E) under the Network Optimized Distributed Energy Systems (NODES) program.

is proposed in [11]. However, the optimization and control strategies proposed in [5]–[10] are *network-agnostic*, in the sense that they assume that all DERs are connected to one electrical node and power flows in the network are ignored. Further, analytical tracking results in time-varying operational conditions are not available in the majority of those works. While this simplifies the design of optimization and control strategies, such strategies do not allow for the regulation of voltage levels throughout the feeder. On the other hand, the proposed methodology enables tracking of power setpoints at the substation while enforcing voltage limits. An economic dispatch model is considered in [12], where distributed controllers are designed to meet a certain load profile over a finite time horizon while minimizing an aggregate cost; stability of the distributed algorithm is analyzed in a continuous-time setting (and, hence, the effects of communication delays and discrete-time operations are not included) and for time-invariant conditions.

A virtual-power-plant profit-maximization problem over a 24-hour period is considered in [13], based on an economic-dispatch model, and a stochastic adaptive optimization approach for virtual power plants participating in the day-ahead and the real-time energy markets is proposed in [14]. However, [13], [14] require solving optimization problems to convergence (i.e., they provide an offline solution method), and are not applicable to real-time control of DERs. A centralized real-time controller for distribution systems is proposed in [15], [16], based on projected-gradient methods; it is shown that the closed-loop system converges on average to a predetermined objective. Compared to [15], [16], our proposed framework affords a distributed implementation, it is based on an online primal-dual method (instead of barrier-type functions, which may lead to prologued constraint violations), and it is shown to be stable under dynamic operational conditions. The online algorithm in [17] does not utilize measurements of voltage magnitudes and powers at the substation; therefore, it may nor enforce voltage regulation and it may fail in tracking setpoints at the substation. Furthermore, the algorithm in [17] considers a diminishing stepsize rule, which is not suitable for the dynamic setting considered in the present paper.

A similar controller design strategy is proposed in [18]; however, the algorithm in [18] does not utilize measurements of the output powers of the DERs. In contrast, the proposed method is based on measurements of the DERs' output powers to promote adaptability and cope with slow-responding DERs. Compared to [18], the present paper provides suitable convergence claims for the considered tracking problem, it provides a more general approximate model for voltage magnitudes, and it provides an approximate model for the power flows at the substation. Finally, a real-time distributed algorithm for the optimal power flow problem is proposed in [19]; however, the distributed algorithm relies on nested loops (thus imposing stringent communication constraints) and it is applicable only to radial (balanced) distribution systems. Relative to [19], the proposed method requires a much simpler communication strategy, and it is applicable to generic systems with mesh or radial topologies.

The remainder of this manuscript is organized as follows. Section II outlines preliminaries and system model. The real-time algorithm is described in Section III, numerical results are presented in Section IV, and concluding remarks are in Section V. Additional modeling details, an extension to multiphase systems, and proofs are provided in the Appendix.

II. Preliminaries

A. Distribution-network Model

Consider a distribution feeder comprising $N+1$ nodes collected in the set $\mathcal{N} \cup \{0\}$, $\mathcal{N} := \{1, \ldots, N\}$, and lines represented by the set of edges $\mathcal{E} := \{(m,n)\} \subset (\mathcal{N} \cup \{0\}) \times (\mathcal{N} \cup \{0\})$.[1] Let $V_n \in \mathbb{C}$ and $I_n \in \mathbb{C}$ denote the phasors for the voltage and the current injected into node n, respectively, and define the N-dimensional complex vectors $\mathbf{v} := [V_1, \ldots, V_N]^{\mathsf{T}} \in \mathbb{C}^N$ and $\mathbf{i} := [I_1, \ldots, I_N]^{\mathsf{T}} \in \mathbb{C}^N$. Node 0 denotes the secondary of the distribution transformer. Using Ohm's and Kirchhoff's circuit laws, the following relationship can be established:

$$\begin{bmatrix} I_0 \\ \mathbf{i} \end{bmatrix} = \begin{bmatrix} y_{00} & \overline{\mathbf{y}}^{\mathsf{T}} \\ \overline{\mathbf{y}} & \mathbf{Y} \end{bmatrix} \begin{bmatrix} V_0 \\ \mathbf{v} \end{bmatrix}, \tag{1}$$

where $\mathbf{Y} \in \mathbb{C}^{N \times N}$, $\overline{\mathbf{y}} \in \mathbb{C}^{N \times 1}$, and $y_{00} \in \mathbb{C}$ are formed based on the system topology and the π-equivalent circuit of the distribution lines (see, e.g., [20, Chapter 6] for additional details on distribution line modeling). Finally, V_0 denotes the voltage at the secondary of the transformer/substation. A constant-power load model is utilized, and $P_{\ell,n}$ and $Q_{\ell,n}$ denote the real and reactive demands at node $n \in \mathcal{N}$ [20]. The active and reactive powers flowing into the feeder at the substation are denoted as P_0 and Q_0.

Let $\mathcal{G} \subseteq \mathcal{N}$ be a set of nodes where DERs are located, and denote by P_i and Q_i the real and reactive powers injected by the DER located at node $i \in \mathcal{G}$. We denote as $\mathcal{Y}_i \subset \mathbb{R}^2$ the set of possible setpoints (P_i, Q_i) for DER i; the set $\mathcal{Y}_i \subset \mathbb{R}^2$ captures hardware and operational constraints of the DER i, and it is assumed to be convex and compact. Some examples are provided next.

Photovoltaic (PV) systems: Let P_i^{av} denote the available real power from a PV system and let S_i be the rated apparent capacity. Then, the set $\mathcal{Y}_i$ is given by:

$$\mathcal{Y}_i = \left\{(P_i, Q_i) \colon 0 \leq P_i \leq P_i^{av}, P_i^2 + Q_i^2 \leq S_i^2\right\}.$$

The set $\mathcal{Y}_i$ is time varying since P_i^{av} depends on underlying irradiance conditions (it can be obtained via, for example, forecasting algorithms). The set $\mathcal{Y}_i$ can also be modified to account for power factor constraints.

[1] Upper-case (lower-case) boldface letters will be used for matrices (column vectors); $(\cdot)^{\mathsf{T}}$ for transposition; $(\cdot)^*$ complex-conjugate; and, $(\cdot)^{\mathsf{H}}$ complex-conjugate transposition. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex number, respectively, and $\mathrm{j} := \sqrt{-1}$. $|\cdot|$ denotes the absolute value of a number or the cardinality of a set. For a given $N \times 1$ vector $\mathbf{x} \in \mathbb{R}^N$, $\|\mathbf{x}\|_2 := \sqrt{\mathbf{x}^{\mathsf{H}}\mathbf{x}}$; $\mathrm{diag}(\mathbf{x})$ returns a $N \times N$ matrix with the elements of $\mathbf{x}$ in its diagonal. Further, $\mathrm{proj}_{\mathcal{Y}}\{\mathbf{x}\}$ denotes the projection of $\mathbf{x}$ onto the convex set $\mathcal{Y}$. Given a matrix $\mathbf{X} \in \mathbb{R}^{N \times M}$, $x_{m,n}$ denotes its (m,n)-th entry, $\mathbf{X}_{n,\cdot}$ denotes the nth row, and $\|\mathbf{X}\|_2$ denotes the ℓ_2-induced matrix norm. $\nabla_{\mathbf{x}} f(\mathbf{x})$ returns the gradient vector of $f(\mathbf{x})$ with respect to $\mathbf{x} \in \mathbb{R}^N$. Finally, $\mathbf{1}_N$ denotes the $N \times 1$ vector with all ones, and $\mathbf{0}_N$ denotes the $N \times 1$ vector with all zeros.

Energy storage systems: The set $\mathcal{Y}_i$ for an energy storage system is given by:

$$\mathcal{Y}_i = \left\{(P_i, Q_i)\colon P_i^{min} \le P_i \le P_i^{max}, P_i^2 + Q_i^2 \le S_i^2\right\},$$

for given limits P_i^{min}, P_i^{max}. These limits are updated during the operation of the battery based on the state of charge. For example, of the battery is fully charged, then $P_i^{min} < 0$ and $P_i^{max} = 0$.

Variable frequency drives: For DERs such as water pumps and supply fans of commercial HVAC systems, the set $\mathcal{Y}_i$ can be described as:

$$\mathcal{Y}_i = \left\{(P_i, Q_i)\colon P_i^{min} \le P_i \le P_i^{max}, Q_i = 0\right\},$$

for given limits P_i^{min}, P_i^{max}. These limits can be fixed or updated by local controllers on a regular basis [8].

Fuel cells: The set $\mathcal{Y}_i$ is given by:

$$\mathcal{Y}_i = \left\{(P_i, Q_i)\colon 0 \le P_i \le P_i^{max}, P_i^2 + Q_i^2 \le S_i^2\right\}$$

for given constant parameters P_i^{max} and S_i.

The operating region of small-scale diesel generators can be modeled using box constraints. For DERs with discrete levels of output powers (e.g., electric vehicle chargers with discrete charging levels), $\mathcal{Y}_i$ represents the convex envelope of the possible operating points; see e.g., [16].

B. Approximate models for the Power-flow Equations

Let $\mathbf{s}_{\text{inj}} := [S_1, \dots, S_N] \in \mathbb{C}^N$ collect the net powers injected at nodes $\mathcal{N}$, where $S_i = P_i - P_{\ell,i} + \mathrm{j}(Q_i - Q_{\ell,i})$ for $i \in \mathcal{G}$, and $S_i = -P_{\ell,i} - \mathrm{j}Q_{\ell,i}$ for $i \in \mathcal{N}\backslash\mathcal{G}$. Then, using (1), the complex-power injections can be compactly written as

$$\mathbf{s}_{\text{inj}} = \operatorname{diag}(\mathbf{v})\,\mathbf{i}^* = \operatorname{diag}(\mathbf{v})\,(\mathbf{Y}^*\mathbf{v}^* + \overline{\mathbf{y}}^* V_0^*)\,. \tag{2}$$

Assume that node 1 is connected to the substation via a distribution line, and recall that the power entering the feeder is given by $S_0 = V_0 I_{01}^*$, where I_{01} is the current flowing on the distribution line $(0,1)$. Particularly, with $y_{01} \in \mathbb{C}$ denoting the admittance of line $(0,1)$, $y_{00}^{\text{sh}} \in \mathbb{C}$ any passive shunt elements connected to node 0, and y_{01}^{sh} the shunt component of the line $(0,1)$, I_{01} is given by $I_{01} = y_{01}(V_0 - V_1) + y_0 V_0$, with $y_0 = g_0 + \mathrm{j}b_0 := y_{00}^{\text{sh}} + y_{01}^{\text{sh}}$. Thus, S_0 can be rewritten as

$$S_0 = |V_0|^2(y_{01}^* + y_0^*) - V_0(y_{01}^* V_1^*)\,. \tag{3}$$

Unfortunately, the nontrivial nonlinearities in (2) and (3) hinder the possibility of seeking analytical closed-form solutions to $\mathbf{v}$, P_0, and Q_0 (as a function of the network topology and composition, power injections, and voltage V_0).

To facilitate the design and analysis of computationally-tractable controllers that afford a real-time implementation, the proposed approach will therefore leverage pertinent linearization approaches for (2)–(3). Particularly, we will utilize the following linear approximation of the nodal-voltage magnitudes $|\mathbf{v}|$ (where the absolute value is taken entry-wise) and P_0, Q_0 as a function of the real and reactive power injections:

$$|\mathbf{v}| \approx \mathbf{A}\mathbf{p}_{\text{inj}} + \mathbf{B}\mathbf{q}_{\text{inj}} + \mathbf{c}\,, \tag{4}$$

$$\begin{bmatrix} P_0 \\ Q_0 \end{bmatrix} \approx \mathbf{M}\mathbf{p}_{\text{inj}} + \mathbf{N}\mathbf{q}_{\text{inj}} + \mathbf{o}\,, \tag{5}$$

where $\mathbf{p}_{\text{inj}} := \Re\{\mathbf{s}_{\text{inj}}\}$ and $\mathbf{q}_{\text{inj}} := \Im\{\mathbf{s}_{\text{inj}}\}$. The model parameters $\mathbf{A} \in \mathbb{R}^{N\times N}$, $\mathbf{B} \in \mathbb{R}^{N\times N}$, $\mathbf{M} \in \mathbb{R}^{2\times N}$, $\mathbf{N} \in \mathbb{R}^{2\times N}$, $\mathbf{c} \in \mathbb{R}^N$, and $\mathbf{o} \in \mathbb{R}^2$ can be obtained using suitable linearization methods for the AC power-flow equations. For example, one can leverage the approximation method proposed in [21], the method based on a first-order linear manifold approximant described in [22], [23], or the so-called "LinDistFlow" approximation [24]. In the Appendix, we will provide an extension of the approach of [21] to derive a more general approximate model for voltage magnitudes and powers at the substation. When the network model is not known, regression methods can be utilized based on real-time measurements of $\mathbf{v}$, P_0, Q_0, and $\mathbf{s}_{\text{inj}}$ (see, e.g., the recursive least-squares method in [25]).

The approximate model (4)–(5) facilitates the design of computationally-affordable algorithms. However, power setpoints obtain from (4)–(5) may cause violations of electrical limits. Section III will then show how to leverage appropriate measurements to enforce electrical limits while enabling effective tracking of setpoints at the substation.

III. Feeder as a Virtual Power Plant

A. Problem Formulation

Control actions are performed in a discrete-time fashion at time instants $\{t_k = k\tau\}_{k\in\mathbb{N}}$, where $\tau > 0$ is the time required to compute one closed-loop iteration of the control strategy illustrated in Fig. 1. The value of τ is limited by communication delays involved in collecting measurements of voltages and powers at the substation and when broadcasting the signal $\mathbf{d}$ to the individual DERs. Typically, τ can be on the order of subseconds to seconds [16], [18]. We start by formalizing a *time-varying* optimization problem to model operational objectives and constraints at each time instant t_k. To this end, define the following quantities related to voltage magnitudes and power at the substation (derived from (4) and (5)):

$$\underline{g}_n^{t_k}(\mathbf{p},\mathbf{q}) := V^{\min} - \overline{c}_n^{t_k} - \sum_{i\in\mathcal{G}}(a_{n,i}^{t_k}P_i + b_{n,i}^{t_k}Q_i), \tag{6a}$$

$$\overline{g}_n^{t_k}(\mathbf{p},\mathbf{q}) := \sum_{i\in\mathcal{G}}(a_{n,i}^{t_k}P_i + b_{n,i}^{t_k}Q_i) + \overline{c}_n^{t_k} - V^{\max}, \tag{6b}$$

$$P_0^{t_k}(\mathbf{p},\mathbf{q}) := \sum_{i\in\mathcal{G}}(m_{1,i}^{t_k}P_i + n_{1,i}^{t_k}Q_i) + \overline{o}_1^{t_k}, \tag{6c}$$

where it follows from (4) and (5) that:

$$\overline{c}_n^{t_k} := c_n^{t_k} - \sum_{i\in\mathcal{N}}(a_{n,i}^{t_k}P_{\ell,i}^{t_k} + b_{n,i}^{t_k}Q_{\ell,i}^{t_k}), \tag{7a}$$

$$\overline{o}_1^{t_k} := o_1^{t_k} - \sum_{i\in\mathcal{N}}(m_{1,i}^{t_k}P_{\ell,i}^{t_k} + n_{1,i}^{t_k}Q_{\ell,i}^{t_k})\,. \tag{7b}$$

Let $P_{0,\text{set}}^{t_k}$ be the setpoint specified for the active power at the substation at time t_k (with a positive sign representing power flowing into the feeder) and assume that the active power at the feeder head must track $P_{0,\text{set}}^{t_k}$ within a given tracking error $E^{t_k} > 0$; that is,

$$h^{t_k}|P_0^{t_k}(\mathbf{p},\mathbf{q}) - P_{0,\text{set}}^{t_k}| \le E^{t_k}, \tag{8}$$

where $h^{t_k}=1$ if the feeder is requested to follow the setpoint $P^{t_k}_{0,\text{set}}$ and $h^{t_k}=0$ otherwise. With these preliminaries in place, consider the following optimization problem:

$$(\text{P1}^{t_k}) \quad \min_{\mathbf{p},\mathbf{q}} \ \sum_{i\in\mathcal{G}} f_i^{t_k}(P_i,Q_i) \tag{9a}$$

$$\text{subject to } P_i,Q_i\in\mathcal{Y}_i^{t_k}, \qquad \forall\, i\in\mathcal{G} \tag{9b}$$

$$h^{t_k}(P_0^{t_k}(\mathbf{p},\mathbf{q})-P^{t_k}_{0,\text{set}})\le E^{t_k}, \tag{9c}$$

$$-h^{t_k}(P_0^{t_k}(\mathbf{p},\mathbf{q})-P^{t_k}_{0,\text{set}})\le E^{t_k}, \tag{9d}$$

$$\underline{g}_n^{t_k}(\mathbf{p},\mathbf{q})\le 0, \qquad \forall n\in\mathcal{M} \tag{9e}$$

$$\overline{g}_n^{t_k}(\mathbf{p},\mathbf{q})\le 0, \qquad \forall n\in\mathcal{M} \tag{9f}$$

where constraints (9e)–(9f) enforce voltage regulation (see (6a)–(6b)) at the M nodes in the set $\mathcal{M}\subset\mathcal{N}$ where measurements of the voltage magnitudes can be obtained. Functions $f_i^{t_k}(\cdot)$ can capture a variety of operational objectives for both the DERs and the utility/aggregator.

Problem (9) defines a sequence of time-varying optimal DER output-power setpoints $\{\mathbf{p}^{\text{opt},t_k},\mathbf{q}^{\text{opt},t_k}\}_{k\in\mathbb{N}}$. However, solving problem (9) at each time t_k in a *batch* fashion to compute the optimal DER output-powers is not feasible in practice. In fact, i) due to underlying computational and communication limits, it is possible that problem (9) may not be solved within τ seconds (especially if distributed solution strategies are utilized); and, ii) since problem (9) relies on approximate linear models, its solution may lead to voltage violations and poor tracking performance.

In the next section, we will develop a computationally affordable online algorithm that continuously pursues the optimal solution trajectory of (9). With the aid of appropriate measurements, the algorithm enables effective tracking of setpoints at the substation while enforcing voltage regulation.

B. Proposed Algorithm

To begin the algorithm design, we first establish a few pertinent technical assumptions regarding (9):

Assumption 1. Function $f_i^{t_k}(P_i,Q_i)$ is convex and continuously differentiable for each $i\in\mathcal{G}$ and for each t_k. Define further the gradient map $\mathbf{f}^{t_k}(\mathbf{p},\mathbf{q}):=[\nabla^{\mathsf{T}}_{[P_1,Q_1]}f_1^{t_k}(P_1,Q_1),\ldots,\nabla^{\mathsf{T}}_{[P_{N_\mathcal{G}},Q_{N_\mathcal{G}}]}f_{N_\mathcal{G}}^{t_k}(P_{N_\mathcal{G}},Q_{N_\mathcal{G}})]^{\mathsf{T}}$. Then, $\mathbf{f}^{t_k}:\mathbb{R}^{2N_\mathcal{G}}\to\mathbb{R}^{2N_\mathcal{G}}$ is Lipschitz continuous with constant L over $\mathcal{Y}^{t_k}:=\mathcal{Y}_1^{t_k}\times\ldots\times\mathcal{Y}_{N_\mathcal{G}}^{t_k}$ for all t_k. □

Assumption 2. For all $t_k\ge 0$, there exist $\{P_i,Q_i\in\mathcal{Y}_i^{t_k}\}_{i\in\mathcal{G}}$ such that constraints (9c)–(9f) can be satisfied. □

Since (9c)–(9f) are linear in $\mathbf{p},\mathbf{q}$, *Assumption 2* implies that Slater's condition holds. If equality constraints are included in (P1^{t_k}), then *Assumption 2* must be properly modified to presuppose the existence of a strictly feasible solution.

The proposed algorithm leverages primal-dual-gradient methods applied to regularized Lagrangian functions [3], [4], [18]. Let $\boldsymbol{\gamma}^{t_k}:=[\gamma_1^{t_k},\ldots,\gamma_M^{t_k}]^{\mathsf{T}}$ and $\boldsymbol{\mu}^{t_k}:=[\mu_1^{t_k},\ldots,\mu_M^{t_k}]^{\mathsf{T}}$ collect the dual variables associated with (9e) and (9f), respectively; similarly, let λ^{t_k} and ζ^{t_k} be the Lagrange multipliers associated with the constraints (9c)–(9d). With $\mathbf{d}:=$

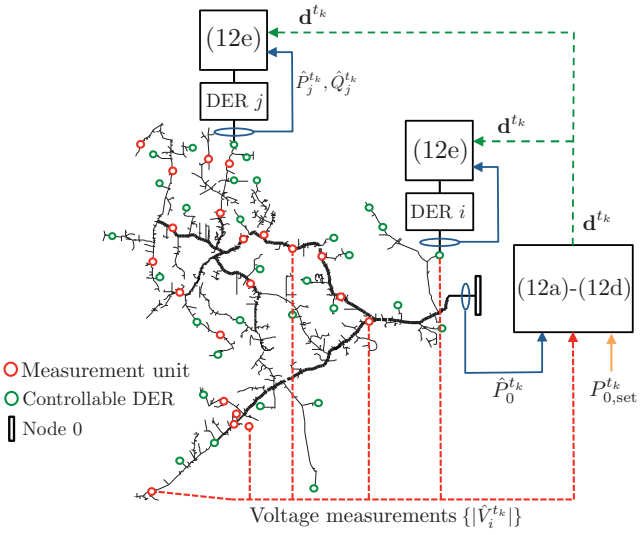

Fig. 2. Operating principles of the control framework (12) when implemented in a distributed fashion. Every τ seconds, measurements of voltage magnitudes and powers at the feeder head are acquired, setpoints for the feeder head are received (if any), and power commands of the DERs are updated via (12).

$\{\boldsymbol{\gamma},\boldsymbol{\mu},\lambda,\zeta\}$, consider the following regularized Lagrangian function associated with (P1^{t_k}):

$$\begin{aligned}\mathcal{L}^{t_k}(\mathbf{p},\mathbf{q},\mathbf{d}):=&\sum_{i\in\mathcal{G}} f_i^{t_k}(P_i,Q_i)\\&+\sum_{n\in\mathcal{M}}\Big(\gamma_n\underline{g}_n^{t_k}(\mathbf{p},\mathbf{q})+\mu_n\overline{g}_n^{t_k}(\mathbf{p},\mathbf{q})\Big)\\&+\lambda\left[h^{t_k}\left(P_0^{t_k}(\mathbf{p},\mathbf{q})-P^{t_k}_{0,\text{set}}\right)-E^{t_k}\right]\\&+\zeta\left[h^{t_k}\left(P^{t_k}_{0,\text{set}}-P_0^{t_k}(\mathbf{p},\mathbf{q})\right)-E^{t_k}\right]\\&+\frac{\nu}{2}\sum_{i\in\mathcal{G}}(P_i^2+Q_i^2)-\frac{\epsilon}{2}\|\mathbf{d}\|_2^2,\end{aligned} \tag{10}$$

where the constants $\nu>0$ and $\epsilon>0$ appearing in the regularization terms are design parameters. Based on (10), consider the following time-varying saddle-point problem:

$$\max_{\mathbf{d}\in\mathbb{R}_+^{2M+2}}\ \min_{\mathbf{p},\mathbf{q}\in\mathcal{Y}^{t_k}} \quad \mathcal{L}^{t_k}(\mathbf{p},\mathbf{q},\mathbf{d}) \tag{11}$$

and denote by $\mathbf{p}^{*,t_k},\mathbf{q}^{*,t_k},\mathbf{d}^{*,t_k}$ the *unique* primal-dual optimizers of (10) at time t_k. In lieu of (9), the proposed algorithm will track solutions of the time-varying saddle-point problem (11). In general, the solutions of (9) and (11) are expected to be different whenever $\nu\neq 0$ and/or $\epsilon\neq 0$. However, the advantages of (11) include desirable convergence properties that improve the ability to track $P^{t_k}_{0,\text{set}}$, and the possibility of acknowledging cost functions $\{f_i^{t_k}\}_{i\in\mathcal{G}}$ that are not strongly convex [3], [4]. The discrepancy between $\mathbf{p}^{\text{opt},t_k},\mathbf{q}^{\text{opt},t_k},\mathbf{d}^{\text{opt},t_k}$ and $\mathbf{p}^{*,t_k},\mathbf{q}^{*,t_k},\mathbf{d}^{*,t_k}$ can be bounded as in [3, Lemma 3.2] (particularly, the gap between the two solutions is proportional to $\sqrt{\epsilon}$).

Integral to the algorithm is the collection of the following *measurements* at each time t_k:
(m1) $\hat{V}_n^{t_k}$: measurement of the voltage $V_n^{t_k}$ at each $n\in\mathcal{M}$;
(m2) $\hat{P}_0^{t_k}$: measurement of the active power at the feeder head;
(m3) $\hat{P}_i^{t_k},\hat{Q}_i^{t_k}$: measurement of the active and reactive output powers of DER $i\in\mathcal{G}$.

Based on measurements (m1)–(m3), the sequential execution of the following steps defines the proposed algorithm:

Real-time Virtual-power-plant Regulation

[S1a] Collect voltage-magnitude measurements $\{|\hat{V}_n^{t_k}|\}_{n\in\mathcal{M}}$.
[S1b] Collect measurement of $\hat{P}_0^{t_k}$.
[S2a] For all $n \in \mathcal{M}$, update $\gamma_n^{t_{k+1}}$ and $\mu_n^{t_{k+1}}$ as follows:

$$\gamma_n^{t_{k+1}} = \text{proj}_{\mathbb{R}_+}\left\{\gamma_n^{t_k} + \alpha\left(V^{\min} - |\hat{V}_n^{t_k}| - \epsilon\gamma_n^{t_k}\right)\right\} \quad (12a)$$

$$\mu_n^{t_{k+1}} = \text{proj}_{\mathbb{R}_+}\left\{\mu_n^{t_k} + \alpha\left(|\hat{V}_n^{t_k}| - V^{\max} - \epsilon\mu_n^{t_k}\right)\right\} \quad (12b)$$

[S2b] For the feeder head, if $h^{t_k} = 1$ update dual variables as follows:

$$\lambda^{t_{k+1}} = \text{proj}_{\mathbb{R}_+}\left\{\lambda^{t_k} + \alpha(\hat{P}_0^{t_k} - P_{0,\text{set}}^{t_k} - E^{t_k} - \epsilon\lambda^{t_k})\right\} \quad (12c)$$

$$\zeta^{t_{k+1}} = \text{proj}_{\mathbb{R}_+}\left\{\zeta^{t_k} + \alpha(P_{0,\text{set}}^{t_k} - \hat{P}_0^{t_k} - E^{t_k} - \epsilon\zeta^{t_k})\right\} \quad (12d)$$

[S3a] Measure output powers $\hat{P}_i^{t_k}, \hat{Q}_i^{t_k}$ at each DER $i \in \mathcal{G}$.
[S3b] Update power setpoints at each DER $i \in \mathcal{G}$ as:

$$\begin{bmatrix} P_i^{t_{k+1}} \\ Q_i^{t_{k+1}} \end{bmatrix} = \text{proj}_{\mathcal{Y}_i^{t_k}}\left\{\begin{bmatrix} P_i^{t_k} \\ Q_i^{t_k} \end{bmatrix} - \alpha\nabla_{[P_i,Q_i]}\mathcal{L}^{t_k}(\mathbf{p},\mathbf{q},\mathbf{d})|_{\hat{P}_i^{t_k},\hat{Q}_i^{t_k},\mathbf{d}^{t_{k+1}}}\right\}, \quad (12e)$$

[S3c] Dispatch setpoints to each DER i, and return to **[S1a]**.

Steps **[S1]**–**[S3]** are performed at each time t_k. The stepsize α is a design parameters chosen as explained in Section III-C. The steps in (12) can be implemented in one of two ways:

- *Centralized implementation*: Steps **[S1]**–**[S3]** are implemented centrally at the utility/aggregator. The utility/aggregator collects measurements of voltages, DER output powers, and $P_0^{t_k}$, executes steps (12), and relays the power setpoints $P_i^{t_k}, Q_i^{t_k}$ to each DER $i \in \mathcal{G}$.
- *Distributed implementation*: As illustrated in Fig. 2, steps **[S1]**–**[S2]** are performed at the utility/aggregator, while step **[S3]** is implemented locally at individual DERs. The utility/aggregator collects measurements of voltages and $P_0^{t_k}$, computes $\mathbf{d}^{t_{k+1}}$, and subsequently broadcasts $\mathbf{d}^{t_{k+1}}$. Each DER updates $P_i^{t_k}, Q_i^{t_k}$ based on $\mathbf{d}^{t_{k+1}}$ and the (local) measurements $\hat{P}_i^{t_k}, \hat{Q}_i^{t_k}$.

The setpoints for the power at the feeder head $P_{0,\text{set}}^{t_k}$ and the target accuracy E^{t_k} are specified (and continuously updated) by the Independent System Operator (ISO) based on transmission-system operating requirements and ISO-utility market agreements.

The ability of the updates to track the optimizers $\mathbf{z}^{*,t_k} := \{\mathbf{p}^{*,t_k}, \mathbf{q}^{*,t_k}, \mathbf{d}^{*,t_k}\}$ of (11) will be analytically established and numerically verified next.

C. Convergence of Algorithm

In this section, convergence of the updates in (12) is established. Begin by noticing that there exists a constant G_g such that $\|\nabla_{[\mathbf{p},\mathbf{q}]}\underline{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q})\|_2 \le G_g$ and $\|\nabla_{[\mathbf{p},\mathbf{q}]}\overline{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q})\|_2 \le G_g$ for all $\mathbf{p},\mathbf{q} \in \mathcal{Y}^{t_k}$ and for all t_k, where $\underline{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q}) \in \mathbb{R}^M$ and $\overline{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q}) \in \mathbb{R}^M$ are vectors stacking all the functions $\underline{g}_n^{t_k}(\mathbf{p},\mathbf{q}), n \in \mathcal{M}$, and $\overline{g}_n^{t_k}(\mathbf{p},\mathbf{q}), n \in \mathcal{M}$, respectively. Also, define the scalar G_0 such that $\|\nabla_{[\mathbf{p},\mathbf{q}]}P_0(\mathbf{p},\mathbf{q})\| \le G_0$ for all $\mathbf{p},\mathbf{q} \in \mathcal{Y}^{t_k}$ and t_k. Then, define the constant $G := \max\{G_g, G_0\}$. For example, since functions $\underline{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q})$ and $\overline{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q})$ are linear [c.f. (4)], a possible bound G_g could be $G_g = \|[\mathbf{A}\ \mathbf{B}]\|_2$. Similarly, $G_0 = \|[\mathbf{M}_{1,\cdot}^\mathsf{T}\ \mathbf{N}_{1,\cdot}^\mathsf{T}]^\mathsf{T}\|_2$ and, thus, $G = \max\{\|[\mathbf{A}\ \mathbf{B}]\|_2, \|[\mathbf{M}_{1,\cdot}^\mathsf{T}\ \mathbf{N}_{1,\cdot}^\mathsf{T}]^\mathsf{T}\|_2\}$ (we recall that $\mathbf{M}_{1,\cdot}$ and $\mathbf{N}_{1,\cdot}$ denote the row 1 of matrices $\mathbf{M}$ and $\mathbf{N}$, respectively). This constant will be used next in the convergence result for the proposed algorithm.

With respect to the measurements (m1)–(m3), the following are assumed:

Assumption 3. There exists a scalar $0 \le e_p < +\infty$ such that:

$$\left\|\begin{bmatrix} \mathbf{p}^{t_k} \\ \mathbf{q}^{t_k} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{p}}^{t_k} \\ \hat{\mathbf{q}}^{t_k} \end{bmatrix}\right\|_2 \le e_p \quad (13)$$

for all t_k, $k \in \mathbb{N}$. □

Assumption 4. There exist constants $0 \le e_v < +\infty$ and $0 \le e_0 < +\infty$ such that

$$\|(\mathbf{A}^{t_k}\mathbf{p}_{\text{inj}}^{t_k} + \mathbf{B}^{t_k}\mathbf{q}_{\text{inj}}^{t_k} + \mathbf{a}^{t_k}) - |\hat{\mathbf{v}}^{t_k}|\|_2 \le e_v \quad (14)$$

$$|(\mathbf{M}_{1,\cdot}\mathbf{p}_{\text{inj}}^{t_k} + \mathbf{N}_{1,\cdot}\mathbf{q}_{\text{inj}}^{t_k} + o_1^{t_k}) - \hat{P}_0^{t_k}| \le e_0 \quad (15)$$

for all t_k, $k \in \mathbb{N}$. □

Assumption 3 provides a bound for the discrepancy between the commanded setpoint $(P_i^{t_k}, Q_i^{t_k})$ and the actual output powers of each DER i [16], [17], [26]. Particularly, the output powers may not coincide with the commanded setpoints $P_i^{t_k}, Q_i^{t_k}$ because: i) measurements may be prone to error, ii) the settling time of the DER output-powers may be larger than τ [17], or iii) $\mathcal{Y}_i^{t_k}$ may represent only an estimate of the actual operating region [16]. *Assumption 4* accounts for measurement and linearization errors. Steps in (12) represent a modified online primal-dual-gradient method applied to the saddle-point problem (11) where actual measurements from the distribution system replace the mathematical models for voltages and powers in (9c)–(9f). When using measurements (m1)–(m3) in (12), the optimal and dual updates involve inexact gradient steps [27]. Based on this observation, the error in the setpoint computation is bounded in the following lemma.

Lemma 1: Let $\{\mathbf{p}_{\text{ex}}^{t_{k+1}}, \mathbf{q}_{\text{ex}}^{t_{k+1}}\}$ be the (exact) primal iterates given by replacing $\{\hat{\mathbf{p}}^{t_k}, \hat{\mathbf{q}}^{t_k}\}$ with $\{\mathbf{p}^{t_k}, \mathbf{q}^{t_k}\}$ in (12e). Suppose that *Assumptions 1* and *3* hold. Then, whenever $\{\hat{\mathbf{p}}^{t_k}, \hat{\mathbf{q}}^{t_k}\} \ne \{\mathbf{p}^{t_k}, \mathbf{q}^{t_k}\}$, the error in the gradient in (12e) is such that:

$$\left\|\begin{bmatrix} \mathbf{p}^{t_k} \\ \mathbf{q}^{t_k} \end{bmatrix} - \begin{bmatrix} \mathbf{p}_{\text{ex}}^{t_k} \\ \mathbf{q}_{\text{ex}}^{t_k} \end{bmatrix}\right\|_2 \le \alpha(L+\nu)e_p \quad (16)$$

where L is the Lipschitz constant of the gradient map $\mathbf{f}^{t_k}(\mathbf{p},\mathbf{q})$. □

The main convergence result is established next.

Theorem 1: Consider the sequence $\{\mathbf{z}^{t_k}\} := \{\mathbf{p}^{t_k}, \mathbf{q}^{t_k}, \mathbf{d}^{t_k}\}$ generated by (12). Let *Assumptions 1–4* hold and suppose that, for fixed positive scalars $\epsilon, \nu > 0$, the stepsize $\alpha > 0$ is chosen such that

$$\rho(\alpha) := \sqrt{1 - 2\alpha\min\{\nu,\epsilon\} + \alpha^2 B} < 1 \quad (17)$$

where $B = (L+\nu+4G)^2 + 4(G+\epsilon)^2$. Then the sequence $\{\mathbf{z}^{t_k}\}$ converges Q-linearly to $\mathbf{z}^{*,t_k} = \{\mathbf{p}^{*,t_k}, \mathbf{q}^{*,t_k}, \mathbf{d}^{*,t_k}\}$ up to the asymptotic error bound given by:

$$\limsup_{t_k\to\infty} \|\mathbf{z}^{t_k} - \mathbf{z}^{*,t_k}\|_2 = \frac{1}{1-\rho(\alpha)}\Big(\alpha e + \sigma\Big) \qquad (18)$$

where $e := \sqrt{(L+\nu)^2 e_p^2 + 2e_v^2 + 2e_0^2}$ and $\sigma \geq 0$ a given constant such that $\|\mathbf{z}^{*,t_{k+1}} - \mathbf{z}^{*,t_k}\| \leq \sigma$ for all $t_k \geq 0$. □

The result (18) bounds the maximum discrepancy between the DER setpoints generated by the algorithm (12) and the time-varying optimizer of (11) at any time t_k. The bound (18) is a function of the parameters α, ϵ, and ν as well as the measurement and linearization errors. The parameters α, ϵ, and ν can be chosen to satisfy condition

$$0 < \alpha < 2\,\frac{\min\{\epsilon,\nu\}}{B} = \frac{\min\{\epsilon,\nu\}}{(L+\nu+4G)^2 + 4(G+\epsilon)^2}$$

to achieve Q-linear convergence.

The bound (18) depends on the underlying dynamics of the distribution system through σ. Particularly, σ captures the maximum difference between the solutions of (11) at two consecutive time instants t_k and t_{k+1} [4]; in the current problem formulation, variations in $\mathbf{z}^{*,t_k}$ are due to changes in the powers injected/consumed by non-controllable assets [cf. (7)], time-varying setpoints $P_{0,\text{set}}^{t_k}$, possibly time-varying cost functions $\{f_i^{t_k}(P_i, Q_i)\}$, and variations in the voltage limits When $\mathbf{z}^{*,t_k}$ varies slowly in time, bound (18) becomes tighter. Expression (18) provides an asymptotic bound for the tracking error; in the Appendix, we discuss how to obtain an upper bound on the tracking error at each iteration. The result (18) can also be interpreted as input-to-state stability, where the optimal trajectory $\{\mathbf{z}^{*,t_k}\}$ of the time-varying problem (9) is taken as a reference. Finally, notice that when $e = 0$ and $\sigma = 0$, the algorithm converges to the solution of the static optimization problem (11).

The proof of Theorem 1 is provided in the Appendix. Although (18) is related to the time-varying solution of the linearized problem (9), ongoing efforts are looking at establishing similar bounds against the time-varying solution of the nonconvex nonlinear counterpart to (9).

Remark 1 (local DER controller) The algorithm (12) produces setpoints $(P_i^{t_k}, Q_i^{t_k}) \in \mathcal{Y}_i^{t_k}$ for the output powers of the DERs. It is assumed that the DERs are endowed with controllers that are designed so that, upon receiving the setpoint, the output powers are driven to the commanded setpoints. Relevant dynamical models for the output powers of inverters operating in a grid-connected mode are discussed in e.g., [28], [29] and can be found in datasheets of commercially available DERs. Assumption 3 accounts for measurement errors and bounds the discrepancy between the commanded setpoint and the actual output power when updates of the setpoints may be performed faster than the DERs' settling times; Assumption 3 is valid, for example, when the DER's response to a step-change in the setpoint follows a first-order model [28], [29].

Remark 2 (feasibility of powers at the substation) Assumption 2 implies that the setpoints for the active power at the substation are feasible. The set of feasible setpoints for the active and reactive power at the substation can be assessed (and optimized) by solving suitable optimization problems at a slower time scale. See, for example, the multi-period optimization problem proposed in [30]. The feasible setpoints $P_{0,\text{set}}^{t_k}$ can be computed based on the operating regions for DERs $\mathcal{Y}_i^{t_k}$ as well as given operational limits. An alternative approach to assess the flexibility of aggregations of DERs is presented in [31].

IV. Numerical Experiments

A. Test Case 1

Consider a modified version of the IEEE 37-node test feeder shown in Fig. 3. The modified network is obtained by considering a single-phase equivalent, and by replacing the loads on phase "c" specified in the original dataset with real load data measured from feeders in a neighborhood called Anatolia in California during a week in August 2012 [32]. Time-series data for the non-controllable loads have a granularity of 1 second and are plotted in Fig. 4. Line impedances and shunt admittances are adopted from the original dataset. With reference to Fig. 3, it is assumed that eighteen PV systems are located at nodes 4, 7, 10, 13, 17, 20, 22, 23, 26, 28, 29, 30, 31, 32, 33, 34, 35, and 36. The rating of these inverters are 300 kVA for $i = 3$, 350 kVA for $i = 15, 16$, and 200 kVA for the remaining ones. The generation profiles are simulated based on the real solar irradiance data available in [32] and have a granularity of 1 second. As an instance, the power available from a PV system with capacity 50 kW is reported in Fig. 4. The dynamics of the output powers of the inverters are modeled using a first-order system; different values for the time constant of the first-order system will be considered throughout this section. Energy storage systems are placed at nodes 3 and 25, their maximum state of charge is 200 kWh, capacity is 50 kVA, and charging and discharging efficiencies are set to 90%. With these simulation settings, overvoltage conditions would occur if the PV inverters operate according to business-as-usual practices (that is, unity power factor and maximum power point). This give us the opportunity to corroborate the ability of the proposed algorithm to track setpoints for the active power at the substation, while concurrently enforcing voltage regulation. The voltage limits $V_{\max}$ and $V_{\min}$ are set to 1.05 pu and 0.95 pu, respectively.

For the controllers illustrated in Fig. 2, the parameters are set as $\nu = 10^{-3}$, $\epsilon = 10^{-4}$, and $\alpha = 0.1$. The target optimization objective (9a) is set to $f_n^{t_k}(P_n, Q_n) = c_p(P_{\text{av},n}^{t_k} - P_n^{t_k})^2 + c_q(Q_n^{t_k})^2$ for PV systems (with $P_{\text{av},n}^{t_k}$ denoting the maximum real power available at PV system n) and $f_n^{t_k}(P_n, Q_n) = c_p(P_n^{t_k})^2 + c_q(Q_n^{t_k})^2$ for the batteries. The coefficients are set to $c_p = 3, c_q = 1$ for the PV systems and $c_p = 1, c_q = 1$ for the batteries. With these functions, the PV systems minimize the power curtailment while the batteries minimize the deviation from a prescribed charging/discharging profile. Functions $\{f_n^{t_k}\}$, however, could accommodate a variety of alternative performance objectives, including the rewards from ancillary service provisioning [33], [34] (to be maximized). The only requirement for each func-

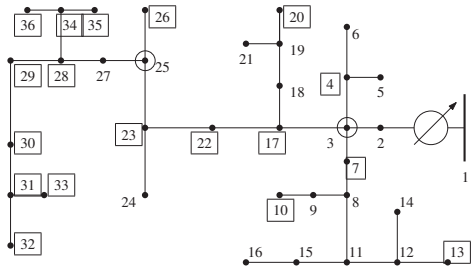

Fig. 3. IEEE 37-node feeder. The boxed nodes represent the location of PV systems and circles represent the location of battery systems for Test Case 1.

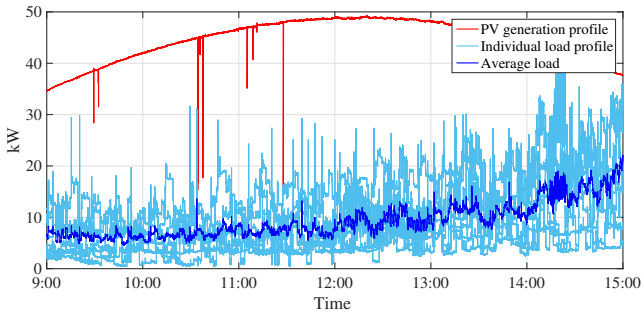

Fig. 4. Profile of loads and power available from the PV systems.

tion $f_n^{t_k}$ is to be convex and with a Lipschitz gradient in its domain.

Throughout this section, the proposed controllers are compared against a strategy where the active output powers of DERs are modified to track setpoints $P_{0,\text{set}}^{t_k}$ via the following rule (performed every τ seconds):

$$P_n^{t_{k+1}} = \text{proj}_{\mathcal{Y}_n^{t_k}} \left\{ P_n^{t_k} - \gamma_n (P_{0,\text{set}}^{t_k} - \hat{P}_0^{t_k}) \right\} \tag{19}$$

where $\gamma_n \geq 0$ is a pre-determined DER participation factor; for subsequent discussions, define $\gamma := \sum_{n \in \mathcal{G}} \gamma_n$. The control rule (19) is network agnostic and may lead to violations of voltage limits. Thus, (19) is complemented with a Volt/VAr heuristic with slope 1 and no deadband.

Assume that setpoints $P_{0,\text{set}}^{t_k}$ are received from 12:00 to 14:00 and they are depicted in red in Fig. 5. The trajectory $\{P_{0,\text{set}}^{t_k}\}$ includes a mix of 5-minute setpoints, 1-minute setpoints, ramp signals, and a command to keep $P_0(t)$ fixed for 1 hour and 5 minutes. Before 12:00 and after 14:00, $h^{t_k} = 0$ and the objective is to provide voltage regulation. Figure 5 illustrates the active power $P_0(t)$ at the feeder head achieved with the proposed controllers when $\tau = 330$ ms. Negative values for $P_0(t)$ indicate reverse power flows. It can be seen that the controllers (12) are such that $P_0(t)$ closely tracks $P_{0,\text{set}}^{t_k}$. A magnified version is provided in Fig. 6 to better illustrate the tracking accuracy. The trajectory for $P_{0,\text{set}}^{t_k}$ illustrated in Fig. 5 might not be realistic; however, such hypothetical trajectory is utilized in this test case in order to comprehensively test the tracking performance of the proposed framework for a variety of dispatch commands.

The tracking performance of the controllers (12) is compared with the strategy (19) in Fig. 7. The coefficients γ_n are set as $\gamma_n = \gamma/|\mathcal{G}|$ for a prescribed γ. When $\gamma \leq 1$, (19)

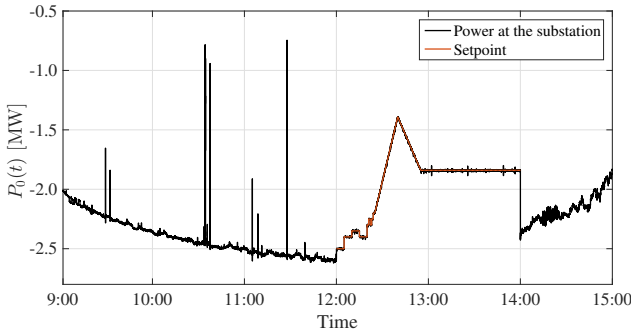

Fig. 5. Power at the substation $P_0(t)$ achieved with the proposed algorithm.

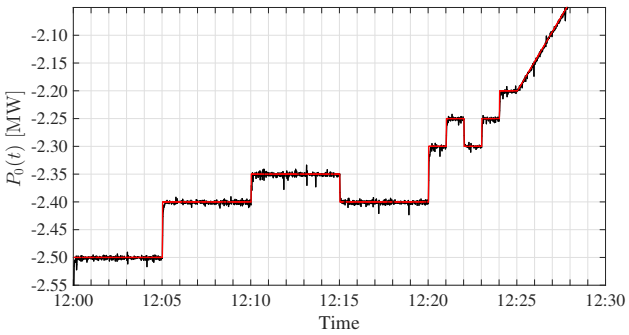

Fig. 6. Trajectory of $P_0(t)$ from 12:00 to 12:30.

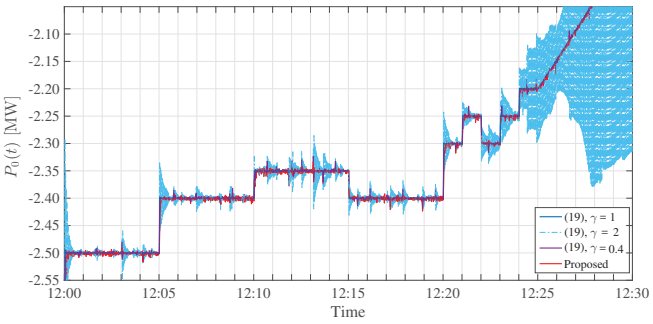

Fig. 7. Comparison with the controller (19) for various participation factors.

successfully tracks the setpoints $P_{0,\text{set}}^{t_k}$ and may converge to the new setpoint faster than the proposed controller when $\gamma = 1$. On the other hand, when $\gamma > 1$, (19) exhibits overshoot behavior that prevents the feeder from following ramping commands. This implies that to set γ properly, one requires knowledge of the number of participating DERs; on the other hand, the stepsize α in (12) depends only on given (and fixed) problem parameters. It is also worth reiterating that the strategy (19) does not address the voltage regulation problem and does not minimize functions $\{f_n^{t_k}(P_n, Q_n)\}$ as demonstrated next.

Figure 8 illustrates the voltage profiles obtained with the proposed controllers. The profiles are compared with the case where the controller (19) is complemented with a local Volt/VAr rule and with the case where the proposed algorithm is inplemented in a network-agnostic (NA) fashion; that is, voltage constraints are discarded, $\mathbf{M} = -\mathbf{I}$, $\mathbf{N} = \mathbf{0}$, and

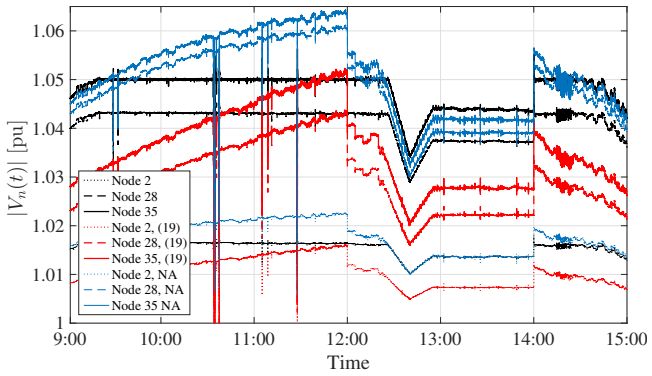

Fig. 8. Voltage profile $|V_n^t|$ in pu achieved by the proposed controllers.

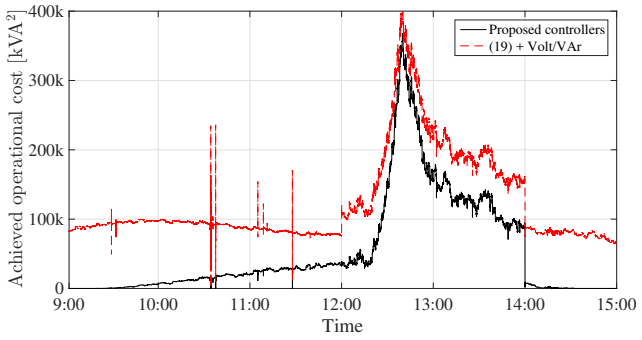

Fig. 9. Cost $\sum_n f_n^{t_k}(P_n, Q_n)$ achieved by (12) and (19).

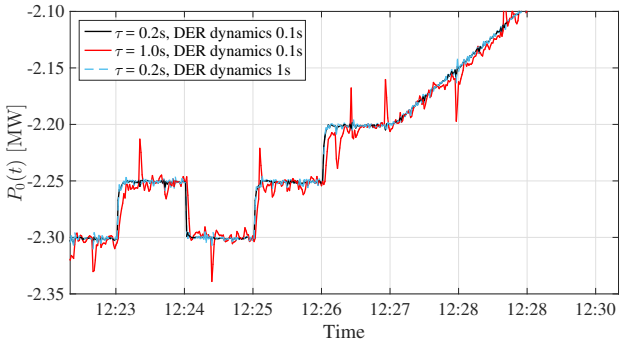

Fig. 10. Performance of the proposed control scheme for different values of τ and time constants of the inverters.

$\mathbf{o} = \mathbf{0}$. It can be seen that the proposed controllers enforce voltage regulation, and a flat voltage profile is obtained; on the other hand, the solution (19)+Volt/VAr may not confine voltages within bounds. The NA implementation exhibits similar tracking performances but leads to voltage violations. Figure 9 illustrates the cost $\sum_n f_n^{t_k}(P_n, Q_n)$ achieved by the proposed solution. It can be seen that, by encapsulating optimization objectives, the proposed method systematically achieves lower operational costs compared to the heuristic (19)+Volt/VAr.

Finally, Fig. 10 demonstrates the tracking capabilities of (12) for different values of τ and inverter time constants. We see that the proposed method is resilient to slow-responding DERs, and ensures tracking accuracy even when the dynamics of the inverters are on the order of 1 s. On the other hand, the tracking performance deteriorates when the time τ required to perform one closed-loop iteration in Fig. 2 increases.

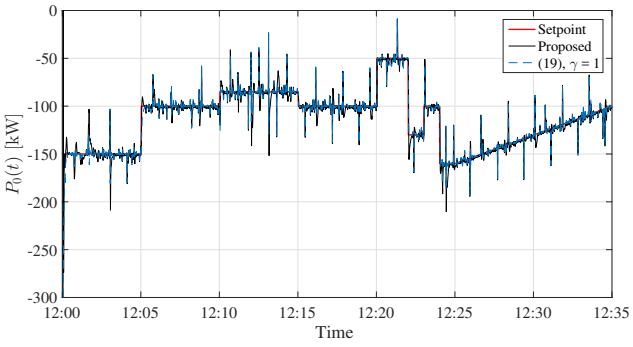

Fig. 11. Performance of the proposed algorithm in Test Case 2. Solution (19) is not stable for $\gamma = 2$ and it is not plotted.

In this test case, the average computational time for step (12e) was 0.17 ms on a MacBook with a 3.1 GHz Intel Core i7 and 16 GB 1867 MHz DDR3. Similar computational times were obtained for the update of the dual variables.

B. Test Case 2

We now consider a test case with a mix of residential-scale PV inverters with capacities of 3kVA and 5kVA, and commercial-scale PV installations with capacities of 10kVA, 50kVA and 100kVA. The number of PV inverters per node along with their aggregate capacity is summarized in Table I. We also consider three utility-scale energy storage systems, as summarized in Table I.

TABLE I
PV AND STORAGE SYSTEMS FOR TEST CASE II.

PV system		
Node	Units	Total Capacity [kVA]
4	5	23
7	7	25
17	4	20
20	4	20
22	4	20
23	4	20
26	4	120
28	2	20
29	2	20
30	2	20
33	2	20
34	2	20
35	3	20
36	3	300
Energy storage systems		
Node	Units	Total Capacity [kWh]
3	1	200
25	1	50
36	3	200

Fig. 11 shows the tracking performance of proposed algorithm (12). The performance is compared with the strategy (19) when the coefficients γ_n are set as $\gamma_n = \gamma/|\mathcal{G}|$, with $\gamma = 1$. Both algorithms exhibit good tracking capabilities, with a relative error of 2.25% and 2.1%. However, (19) leads to instability when $\gamma = 2$ (and it is not plotted for ease of readability).

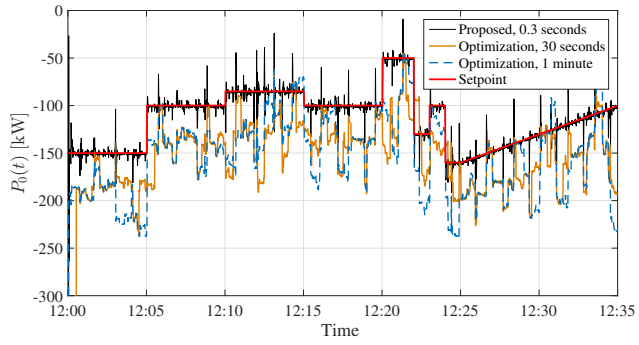

Fig. 12. Comparison with an offline optimization solution where power setpoints are commanded to the DERs only upon convergence of the algorithm; different convergence times are simulated.

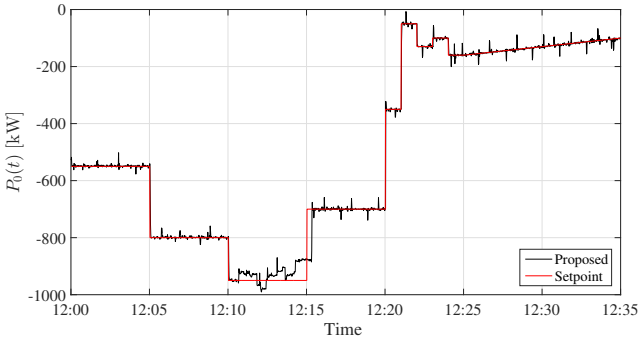

Fig. 13. Performance of the proposed algorithm when a setpoint for the active power at the substation is not feasible. The setpoint $P_{0,\text{set}}^{t_k} = -950$ kW given between 12:10 and 12:15 is not feasible.

Fig. 12 compares the tracking capabilities of the proposed real-time algorithm with an offline optimization solution where power setpoints are commanded to the DERs only upon convergence of the algorithm; see e.g., [9], [13], [14], [21]. It is assumed that the offline optimization method can converge to a solution in 30 seconds and in 1 minute respectively. The linearized optimal power flow method proposed in e.g. [21] is utilized. It can be seen that the optimization-based approach is not able to track the specified setpoints for the power at the substation, and the relative error is of 37% when a 30-second update is considered; this is because this approach cannot cope with variations in the non-controllable loads within the update interval. Further, it can be seen that approximation errors of the linearized model (5) introduce a bias in the tracking error.

Fig. 13 illustrates the case where a setpoint for the active power at the substation is not feasible. Particularly, the setpoint $P_{0,\text{set}}^{t_k} = -950$ kW given between 12:10 and 12:15 is not feasible, in the sense that the powers generated by the PV systems and the powers injected into the feeder by the batteries in aggregate are not sufficient for the power at the substation to reach $P_{0,\text{set}}^{t_k}$. As expected, it can be seen that the algorithm is not able to track the setpoint -950 kW at the substation; however, once a new feasible setpoint is given, the algorithm drives the active power at the substation at the desired value. It is worth pointing out that the solution (19) would not be able to track the setpoint -950 kW, and offline optimization solvers [9], [13], [14], [21] would not return a solution for the DERs' output powers.

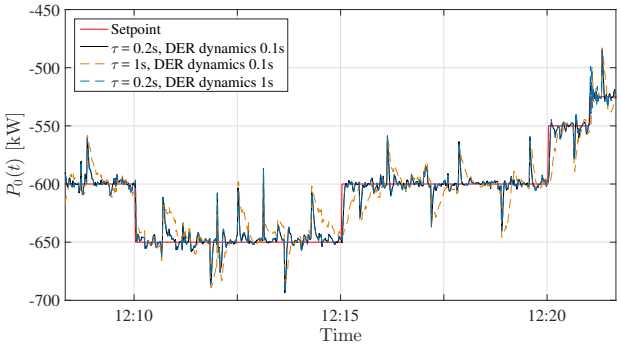

Fig. 14. Performance of the proposed control scheme for different values of τ and time constants of the inverters in Test Case II.

Finally, Fig. 14 demonstrates the tracking capabilities of (12) for different values of τ and inverter time constants. Similar to the results obtained in Test Case I, the proposed method is resilient to slow-responding DERs. On the other hand, for higher values of τ, the controllers respond in a slower manner to variations in the non-controllable loads and irradiance.

To further validate the performance gains with respect to an NA implementation, we consider an additional test case where distribution lines feature shunts elements; in the IEEE 37-node test feeder, this introduces a more pronounced diversification of the coefficients in the matrix $\mathbf{M}$ in (5). The tracking error of the proposed method turned out to be 1.8%, whereas the NA implementation exhibited a tracking error of 6.9%. The tracking error was computed by taking the time average of $|P_0^{t_k} - P_{0,\text{set}}^{t_k}|/P_{0,\text{set}}^{t_k}$. Overall, the proposed network-cognizant implementation allows one to tightly control voltages within limits while achieving higher tracking performance.

V. Concluding Remarks

This paper developed an algorithmic framework to enable distribution networks to emulate a virtual power plants that respond to regulation requests received from the transmission system. The controllers adjust the output powers of individual DERs in response to setpoints for the power at the feeder head, while concurrently regulating voltages within the feeder and maximizing customers' and utility's performance objectives. With respect to stability and tracking capabilities, analytical results were presented. Numerical experiments corroborated the analytical findings and assessed the tracking performance for different speeds of updates for the DER's commanded powers. It is shown that the proposed method is resilient to slow-responding DERs, and ensures tracking accuracy even when the dynamics of the inverters are on the order of seconds. It is also shown that the proposed approach outperforms traditional offline optimization approaches.

Future research endeavors will broaden the applicability of the proposed algorithm to account for DERs with discrete power commands (including on/off decisions), and will look at the development of real-time algorithm for nonconvex problem formulations. Finally, notice that the proposed real-time algorithm is in fact a myopic control strategy. We will

then pursue the development of online algorithms for time-varying multi-period optimization problems.

APPENDIX

A. Linear approximation

The parameters of the AC power-flow approximations (4)–(5) can be computed (and periodically updated) in multiple ways. Suitable linearization methods for the AC power-flow equations, such as the methods outlined in [21], [23] and the the so-called LinDistFlow approximation [24] can be utilized. Alternatively, the model parameters can be estimated via regression-based methods such as the online RLS algorithm, based on real-time measurements of voltages and powers flows. In this section, an example for a linear approximation method of the AC power-flow equations is outlined; particularly, we broaden the approach of [21] to provide a more general approximation of voltage magnitudes and derive an approximate relationship between powers at the substation and net injected powers throughout the feeder. For ease of exposition, the temporal index t_k is dropped.

Central to the linearization approach is to express the voltages $\mathbf{v}$ as $\mathbf{v} = \mathbf{v}_{\rm nom} + \mathbf{v}_{\rm e}$, where $\mathbf{v}_{\rm nom}$ is some nominal-voltage vector (i.e., the linearization point) determined *a priori*, and entries of $\mathbf{v}_{\rm e}$ capture perturbations around $\mathbf{v}_{\rm nom}$. With $\mathbf{v}_{\rm nom}$ appropriately determined, we need to solve for $\mathbf{v}_e$ that satisfies the following equation:

$$\mathbf{s} = \operatorname{diag}\left(\mathbf{v}_{\rm nom} + \mathbf{v}_e\right)\left(\mathbf{Y}^*(\mathbf{v}_{\rm nom} + \mathbf{v}_e)^* + \overline{\mathbf{y}}^* V_0^*\right). \quad (20)$$

Expanding (20), one gets

$$\begin{aligned}\mathbf{s} = {} & \operatorname{diag}\left(\mathbf{v}_{\rm nom}\right)\mathbf{Y}^*\mathbf{v}_{\rm nom}^* + \operatorname{diag}\left(\mathbf{v}_{\rm nom}\right)\mathbf{Y}^*\mathbf{v}_e^* \\ & + \operatorname{diag}\left(\mathbf{v}_e\right)\mathbf{Y}^*\mathbf{v}_{\rm nom}^* + \operatorname{diag}\left(\mathbf{v}_e\right)\mathbf{Y}^*\mathbf{v}_e^* \\ & + \operatorname{diag}\left(\mathbf{v}_{\rm nom}\right)\overline{\mathbf{y}}^* V_0^* + \operatorname{diag}\left(\mathbf{v}_e\right)\overline{\mathbf{y}}^* V_0^*. \end{aligned} \quad (21)$$

Neglecting the second-order term $\operatorname{diag}\left(\mathbf{v}_e\right)\mathbf{Y}^*\mathbf{v}_e^*$, and recognizing that

$$\begin{aligned}\operatorname{diag}\left(\mathbf{v}_e\right)\mathbf{Y}^*\mathbf{v}_{\rm nom}^* &= \operatorname{diag}\left(\mathbf{Y}^*\mathbf{v}_{\rm nom}^*\right)\mathbf{v}_e, \\ \operatorname{diag}\left(\mathbf{v}_e\right)\overline{\mathbf{y}}^* V_0^* &= V_0^* \operatorname{diag}\left(\overline{\mathbf{y}}^*\right)\mathbf{v}_e, \end{aligned} \quad (22)$$

and reorganizing terms, (21) can be compactly rewritten as

$$\mathbf{\Gamma}\mathbf{v}_e + \mathbf{\Xi}\mathbf{v}_e^* = \mathbf{s} - \mathbf{s}_{\rm nom}, \quad (23)$$

where $\mathbf{\Gamma} \in \mathbb{C}^{N\times N}$, $\mathbf{\Xi} \in \mathbb{C}^{N\times N}$, and $\mathbf{s}_{\rm nom} \in \mathbb{C}^N$ are given by

$$\mathbf{\Gamma} := \operatorname{diag}\left(\mathbf{Y}^*\mathbf{v}_{\rm nom}^* + \overline{\mathbf{y}}^* V_0^*\right), \quad (24)$$

$$\mathbf{\Xi} := \operatorname{diag}\left(\mathbf{v}_{\rm nom}\right)\mathbf{Y}^*, \quad (25)$$

$$\mathbf{s}_{\rm nom} := \operatorname{diag}\left(\mathbf{v}_{\rm nom}\right)\left(\mathbf{Y}^*\mathbf{v}_{\rm nom}^* + \overline{\mathbf{y}}^* V_0^*\right). \quad (26)$$

The next step consists solving for the voltage perturbation vector $\mathbf{v}_e$, using which one can recover an approximation to the actual solution $\mathbf{v}$. Thus, decomposing all quantities in (23) into their real and imaginary parts, one can solve for $\Re\{\mathbf{v}_{\rm e}\}$ and $\Im\{\mathbf{v}_{\rm e}\}$ (and hence, for $\mathbf{v}_e$) from

$$\begin{bmatrix} \Re\{\mathbf{v}_{\rm e}\} \\ \Im\{\mathbf{v}_{\rm e}\} \end{bmatrix} = \mathbf{H}\begin{bmatrix} \mathbf{p}_{\rm inj} \\ \mathbf{q}_{\rm inj} \end{bmatrix} - \mathbf{H}\begin{bmatrix} \mathbf{p}_{\rm nom} \\ \mathbf{q}_{\rm nom} \end{bmatrix}, \quad (27)$$

where $\mathbf{p}_{\rm nom} := \Re\{\mathbf{s}_{\rm nom}\}, \mathbf{q}_{\rm nom} := \Im\{\mathbf{s}_{\rm nom}\} \in \mathbb{R}^N$ denote the active- and reactive-power injected into the network *at* the nominal voltage, $\mathbf{v}_{\rm nom}$, and $\mathbf{H} \in \mathbb{R}^{2N\times 2N}$ is defined as follows:

$$\mathbf{H} := \begin{bmatrix} \Re\{\mathbf{\Gamma}\} + \Re\{\mathbf{\Xi}\} & -\Im\{\mathbf{\Gamma}\} + \Im\{\mathbf{\Xi}\} \\ \Im\{\mathbf{\Gamma}\} + \Im\{\mathbf{\Xi}\} & \Re\{\mathbf{\Gamma}\} - \Re\{\mathbf{\Xi}\} \end{bmatrix}^{-1}. \quad (28)$$

To aid subsequent discussions, we will find it useful to denote the $N \times N$ blocks that $\mathbf{H}$ is composed of by $\mathbf{H}^{(11)}$, $\mathbf{H}^{(12)}$, $\mathbf{H}^{(21)}$, and $\mathbf{H}^{(22)}$. In particular, this allows one to express

$$\Re\{\mathbf{v}_{\rm e}\} = \mathbf{H}^{(11)}\mathbf{p}_{\rm inj} + \mathbf{H}^{(12)}\mathbf{q}_{\rm inj} + \mathbf{h}_r, \quad (29a)$$

$$\Im\{\mathbf{v}_{\rm e}\} = \mathbf{H}^{(21)}\mathbf{p}_{\rm inj} + \mathbf{H}^{(22)}\mathbf{q}_{\rm inj} + \mathbf{h}_i, \quad (29b)$$

where $\mathbf{h}_r := -\mathbf{H}^{(11)}\mathbf{p}_{\rm nom} - \mathbf{H}^{(12)}\mathbf{q}_{\rm nom}$ and $\mathbf{h}_i := -\mathbf{H}^{(21)}\mathbf{p}_{\rm nom} - \mathbf{H}^{(22)}\mathbf{q}_{\rm nom}$.

Next, we leverage (27) to obtain (4). Begin by expressing:

$$\mathbf{v} = \mathbf{v}_{\rm nom} + \mathbf{v}_e = \operatorname{diag}\left(\mathrm{e}^{\mathrm{j}\boldsymbol{\theta}_{\rm nom}}\right)|\mathbf{v}_{\rm nom}| + \mathbf{v}_e. \quad (30)$$

Then, left multiplying (30) by $\operatorname{diag}(\mathrm{e}^{-\mathrm{j}\boldsymbol{\theta}_{\rm nom}})$ we obtain:

$$\begin{aligned}&\operatorname{diag}(\mathrm{e}^{-\mathrm{j}\boldsymbol{\theta}_{\rm nom}})\mathbf{v} = |\mathbf{v}_{\rm nom}| + \operatorname{diag}(\mathrm{e}^{-\mathrm{j}\boldsymbol{\theta}_{\rm nom}})\mathbf{v}_e \quad (31)\\ &= \operatorname{diag}(|\mathbf{v}_{\rm nom}|)\left(\mathbf{1}_N + \operatorname{diag}\left(\mathrm{e}^{-\mathrm{j}\boldsymbol{\theta}_{\rm nom}}\right)\operatorname{diag}\left(|\mathbf{v}_{\rm nom}|\right)^{-1}\mathbf{v}_e\right).\end{aligned}$$

Considering the (element-wise) magnitude on both sides above, it follows that:

$$|\mathbf{v}| = \operatorname{diag}(|\mathbf{v}_{\rm nom}|)\left|\mathbf{1}_N + \operatorname{diag}\left(\mathrm{e}^{-\mathrm{j}\boldsymbol{\theta}_{\rm nom}}\right)\operatorname{diag}\left(|\mathbf{v}_{\rm nom}|\right)^{-1}\mathbf{v}_e\right|. \quad (32)$$

Consider the approximation $|\mathbf{1}_N + \boldsymbol{\nu}| \approx \mathbf{1}_N + \Re(\boldsymbol{\nu})$ for $|\boldsymbol{\nu}| \prec\prec \mathbf{1}_N$ (element-wise) where $\boldsymbol{\nu} \in \mathbb{C}^N$. Since $\mathbf{v}_e$ represents a small perturbation around $\mathbf{v}_{\rm nom}$, it is reasonable to assume that

$$\left|\operatorname{diag}\left(\mathrm{e}^{-\mathrm{j}\boldsymbol{\theta}_{\rm nom}}\right)\operatorname{diag}\left(|\mathbf{v}_{\rm nom}|\right)^{-1}\mathbf{v}_e\right| \prec\prec \mathbf{1}_N. \quad (33)$$

Therefore, from (32), it follows that $|\mathbf{v}| \approx |\mathbf{v}_{\rm nom}| + \Re\left\{\operatorname{diag}(\mathrm{e}^{-\mathrm{j}\boldsymbol{\theta}_{\rm nom}})\mathbf{v}_e\right\}$ and, finally, from (27) one has that

$$|\mathbf{v}| = |\mathbf{v}_{\rm nom}| - \mathbf{\Theta}_{\rm nom}\mathbf{H}\begin{bmatrix} \mathbf{p}_{\rm nom} \\ \mathbf{q}_{\rm nom} \end{bmatrix} + \mathbf{\Theta}_{\rm nom}\mathbf{H}\begin{bmatrix} \mathbf{p}_{\rm inj} \\ \mathbf{q}_{\rm inj} \end{bmatrix}, \quad (34)$$

where $\mathbf{H}$ is defined in (28), and we define $\mathbf{\Theta}_{\rm nom} \in \mathbb{R}^{N\times 2N}$ as $\mathbf{\Theta}_{\rm nom} := \left[\operatorname{diag}(\cos(\boldsymbol{\theta}_{\rm nom}))\ \operatorname{diag}(\sin(\boldsymbol{\theta}_{\rm nom}))\right]$. Notice that (34) expresses the vector of node-voltage magnitudes as a linear function of the active- and reactive-power injections in the network. Equation (4) can be obtained upon setting $\mathbf{c} = |\mathbf{v}_{\rm nom}| - \mathbf{\Theta}_{\rm nom}\mathbf{H}[\mathbf{p}_{\rm nom}^{\mathsf{T}}, \mathbf{q}_{\rm nom}^{\mathsf{T}}]^{\mathsf{T}}$ and appropriately including the entries of matrix $\mathbf{\Theta}_{\rm nom}\mathbf{H}$ into $\mathbf{A}$ and $\mathbf{B}$.

Approximation (5) is derived next. With $\{\mathbf{e}_i \in \mathbb{R}^N\}_{i=1}^N$ denoting the vector basis for $\mathbb{R}^N$, it follows that V_1 can be rewritten as $V_1 = e_1^{\mathsf{T}}(\mathbf{v}_{\rm nom} + \Re\{\mathbf{v}_{\rm e}\} + \mathrm{j}\Im\{\mathbf{v}_{\rm e}\})$ and, thus:

$$\begin{aligned}S_0 = {} & |V_0|^2(y_{01}^* + y_0^*) \\ & - V_0[y_{01}^*(e_1^{\mathsf{T}}(\mathbf{v}_{\rm nom} + \Re\{\mathbf{v}_{\rm e}\} + \mathrm{j}\Im\{\mathbf{v}_{\rm e}\}))^*]. \end{aligned} \quad (35)$$

Substituting (29) in (35) and rearranging terms, the approximate linear relationship (5) between the power at the feeder

head $S_0 = P_0 + \mathrm{j}Q_0$ and the net power injections $\mathbf{p}_{\rm inj}, \mathbf{q}_{\rm inj}$ can be derived by setting $\mathbf{M}$, $\mathbf{N}$, and $\mathbf{o}$ as:

$$\begin{bmatrix} \mathbf{M} \\ \mathbf{N} \end{bmatrix} = \begin{bmatrix} -\psi_1 & 0 & \psi_2 & 0 \\ \psi_2 & 0 & \psi_1 & 0 \\ 0 & -\psi_1 & 0 & \psi_2 \\ 0 & \psi_2 & 0 & \psi_1 \end{bmatrix} \begin{bmatrix} \mathbf{H}_{1,\cdot}^{(11)} \\ \mathbf{H}_{1,\cdot}^{(12)} \\ \mathbf{H}_{1,\cdot}^{(21)} \\ \mathbf{H}_{1,\cdot}^{(22)} \end{bmatrix}, \quad (36)$$

$$\mathbf{o} = |V_0|^2 \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} g_{01}, g_0, b_{01}, b_0 \end{bmatrix}^\mathsf{T} + \begin{bmatrix} -\psi_1 & \psi_2 & -\psi_1 & \psi_2 \\ \psi_2 & \psi_1 & \psi_2 & \psi_1 \end{bmatrix} \begin{bmatrix} \Re\{V_{\rm nom,1}\} \\ \Im\{V_{\rm nom,1}\} \\ -\mathbf{H}_{1,\cdot}^{(11)}\mathbf{p}_{\rm nom} - \mathbf{H}_{1,\cdot}^{(12)}\mathbf{q}_{\rm nom} \\ -\mathbf{H}_{1,\cdot}^{(21)}\mathbf{p}_{\rm nom} - \mathbf{H}_{1,\cdot}^{(22)}\mathbf{q}_{\rm nom} \end{bmatrix}$$

where the following scalars have been defined for brevity:

$$\psi_1 = |V_0|(\cos(\theta_0)g_{01} + \sin(\theta_0)b_{01}) \quad (37)$$

$$\psi_2 = |V_0|(\cos(\theta_0)b_{01} - \sin(\theta_0)g_{01})\,. \quad (38)$$

It is worth pointing out that the linear model presented in this subsection can be extended to the multiphase unbalanced case. For example, the fixed-point methodologies presented in [35] can be utilized to derive approximate linear relationships of voltage magnitudes and powers at the substation in multiphase settings. Alternatively, the first-order Taylor method proposed in [22] can be utilized.

B. Proof of Lemma 1

Recall that the *exact update* of the primal variables can be obtained by replacing the power measurements $\{\hat{\mathbf{p}}^{t_k}, \hat{\mathbf{q}}^{t_k}\}$ with the iterates $\{\mathbf{p}^{t_k}, \mathbf{q}^{t_k}\}$ in (12e); that is,

$$\begin{bmatrix} \mathbf{p}_{\rm ex}^{t_{k+1}} \\ \mathbf{q}_{\rm ex}^{t_{k+1}} \end{bmatrix} = {\rm proj}_{\mathcal{Y}^{t_k}} \left\{ \begin{bmatrix} \mathbf{p}^{t_k} \\ \mathbf{q}^{t_k} \end{bmatrix} - \alpha \nabla_{[\mathbf{p},\mathbf{q}]} \mathcal{L}^{t_k}(\mathbf{p},\mathbf{q},\mathbf{d})|_{\mathbf{p}^{t_k},\mathbf{q}^{t_k},\mathbf{d}^{t_k}} \right\} \quad (39)$$

where $\mathcal{Y}^{t_k} = \mathcal{Y}_1^{t_k} \times \ldots \times \mathcal{Y}_N^{t_k}$ is the Cartesian product of the operating regions of the DERs. For brevity, let $\mathbf{u}^{t_k} := [(\mathbf{p}^{t_k})^\mathsf{T}, (\mathbf{q}^{t_k})^\mathsf{T}]^\mathsf{T}$ and $\mathbf{u}_{\rm ex}^{t_k} := [(\mathbf{p}_{\rm ex}^{t_k})^\mathsf{T}, (\mathbf{q}_{\rm ex}^{t_k})^\mathsf{T}]^\mathsf{T}$; further, collect the measured output powers in the vector $\hat{\mathbf{u}}^{t_k} := [(\hat{\mathbf{p}}^{t_k})^\mathsf{T}, (\hat{\mathbf{q}}^{t_k})^\mathsf{T}]^\mathsf{T}$ and recall that $\mathbf{d}^{t_k} = [(\boldsymbol{\gamma}^{t_k})^\mathsf{T}, (\boldsymbol{\mu}^{t_k})^\mathsf{T}, \lambda^{t_k}, \zeta^{t_k}]^\mathsf{T}$. Leveraging the non-expansive property of the projection operator, and using the bounds in *Assumption 1* and *3*, it follows that

$$\begin{aligned}
&\left\|\mathbf{u}^{t_k} - \mathbf{u}_{\rm ex}^{t_k}\right\|_2 \\
&\le \left\|\alpha\left(\mathcal{L}^{t_k}(\mathbf{u},\mathbf{d})|_{\mathbf{u}^{t_k},\mathbf{d}^{t_k}} - \mathcal{L}^{t_k}(\mathbf{u},\mathbf{d})|_{\hat{\mathbf{u}}^{t_k},\mathbf{d}^{t_k}}\right)\right\|_2 && (40a)\\
&= \left\|\alpha\left(\mathbf{f}^{t_k}(\mathbf{u})|_{\mathbf{u}^{t_k}} - \mathbf{f}^{t_k}(\mathbf{u})|_{\hat{\mathbf{u}}^{t_k}} + \nu(\mathbf{u}^{t_k} - \hat{\mathbf{u}}^{t_k})\right)\right\|_2 && (40b)\\
&\le \alpha\left\|\mathbf{f}^{t_k}(\mathbf{u})|_{\mathbf{u}^{t_k}} - \mathbf{f}^{t_k}(\mathbf{u})|_{\hat{\mathbf{u}4}^{t_k}}\right\|_2 + \alpha\nu\left\|\mathbf{u}^{t_k} - \hat{\mathbf{u}}^{t_k}\right\|_2 && (40c)\\
&\le \alpha L\left\|\mathbf{u}^{t_k} - \hat{\mathbf{u}}^{t_k}\right\|_2 + \alpha\nu\left\|\mathbf{u}^{t_k} - \hat{\mathbf{u}}^{t_k}\right\|_2 && (40d)\\
&\le \alpha L e_p + \alpha\nu e_p && (40e)
\end{aligned}$$

where the first term on the right hand side of (40d) follows from the Lipschitz continuity of the gradient map $\mathbf{f}^{t_k}(\mathbf{p},\mathbf{q})$. Step (40e) then follows from (13).

C. Proof of Theorem 1

Begin by defining the following time-varying operator $\boldsymbol{\Phi}^{t_k}$:

$$\boldsymbol{\Phi}^{t_k} : \{\mathbf{z}^{t_k}\} \mapsto \begin{bmatrix} \nabla_{[\mathbf{p},\mathbf{q}]}\mathcal{L}^{t_k}(\mathbf{p},\mathbf{q},\mathbf{d})|_{\mathbf{p}^{t_k},\mathbf{q}^{t_k},\mathbf{d}^{t_k}} \\ -(\underline{\mathbf{g}}^{t_k}(\mathbf{p}^{t_k},\mathbf{q}^{t_k}) - \epsilon\boldsymbol{\gamma}^{t_k}) \\ -(\bar{\mathbf{g}}^{t_k}(\mathbf{p}^{t_k},\mathbf{q}^{t_k}) - \epsilon\boldsymbol{\mu}^{t_k}) \\ -(P_0(\mathbf{p}^{t_k},\mathbf{q}^{t_k}) - P_{0,\rm set}^{t_k} - E - \epsilon\lambda^{t_k}) \\ -(P_{0,\rm set}^{t_k} - P_0(\mathbf{p}^{t_k},\mathbf{q}^{t_k}) - E - \epsilon\zeta^{t_k}) \end{bmatrix},$$

where $\boldsymbol{\gamma}^{t_k}$ and $\boldsymbol{\mu}^{t_k}$ are vectors collecting the dual variables $\{\gamma_n^{t_k}\}$ and $\{\mu_n^{t_k}\}$, respectively, and $\mathbf{z}^{t_k} := [(\mathbf{u}^{t_k})^\mathsf{T}, (\mathbf{d}^{t_k})^\mathsf{T}]^\mathsf{T}$, with $\mathbf{u}^{t_k} = [(\mathbf{p}^{t_k})^\mathsf{T}, (\mathbf{q}^{t_k})^\mathsf{T}]^\mathsf{T}$. The map $\boldsymbol{\Phi}^{t_k}$ represents the counterpart of (12) when the voltage and power measurements are replaced by functions $\underline{\mathbf{g}}^{t_k}(\cdot)$, $\bar{\mathbf{g}}^{t_k}(\cdot)$, and $P_0(\cdot)$ evaluated at the most-up-to-date iterates $\mathbf{p}^{t_k}$ and $\mathbf{q}^{t_k}$. Recall that, since $\mathcal{Y}^{t^{t_k}}$ is compact and $\mathbf{g}^{t_k}(\mathbf{p},\mathbf{q}) \in \mathbb{R}^M$ and $\bar{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q}) \in \mathbb{R}^M$ are linear in $\mathbf{p},\mathbf{q}$, it follows that there exists a constant G_g such that $\|\nabla_{[\mathbf{p},\mathbf{q}]}\underline{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q})\|_2 \le G_g$ and $\|\nabla_{[\mathbf{p},\mathbf{q}]}\bar{\mathbf{g}}^{t_k}(\mathbf{p},\mathbf{q})\|_2 \le G_g$ for all $\mathbf{p},\mathbf{q} \in \mathcal{Y}^{t_k}$ and for all t_k. Further, there exist a scalar G_0 such that $\|\nabla_{[\mathbf{p},\mathbf{q}]}P_0(\mathbf{p},\mathbf{q})\|_2 \le G_0$ for all $\mathbf{p},\mathbf{q} \in \mathcal{Y}^{t_k}$ and t_k. Let $G = \max\{G_g, G_0\}$. Then, the following holds.

Lemma 2: The map $\boldsymbol{\Phi}^{t_k}$ is strongly monotone with constant $\eta = \min\{\nu,\epsilon\}$, and Lipschitz over $\mathcal{Y}^{t_k} \times \mathbb{R}^{2M+2}$ with constant $\bar{B} = [(L+\nu+4G)^2 + 4(G+\epsilon)^2]^{\frac{1}{2}}$. □

Strong monotonicity can be shown by noticing that:

$$\begin{aligned}
&(\boldsymbol{\Phi}^{t_k}(\mathbf{z}_1) - \boldsymbol{\Phi}^{t_k}(\mathbf{z}_2))^\mathsf{T}(\mathbf{z}_1 - \mathbf{z}_2) \\
&\quad \ge \nu\|\mathbf{u}_1 - \mathbf{u}_2\|_2^2 + \epsilon(\|\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2\|_2^2 + \|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|_2^2 \\
&\quad + |\lambda_1 - \lambda_2|_2^2 + |\zeta_1 - \zeta_2|^2) && (41a)\\
&\quad \ge \min\{\nu,\epsilon\}\|\mathbf{z}_1 - \mathbf{z}_2\|_2 && (41b)
\end{aligned}$$

where (41a) was derived by using the facts that: i) the regularized Lagrangian is strongly convex with constant ν for all t_k, and ii) the term

$$\begin{bmatrix} -\underline{\mathbf{g}}^{t_k}(\mathbf{u}_1) + \underline{\mathbf{g}}^{t_k}(\mathbf{u}_2) \\ -\bar{\mathbf{g}}^{t_k}(\mathbf{u}_1) + \bar{\mathbf{g}}^{t_k}(\mathbf{u}_2) \\ -P_0(\mathbf{u}_1) + P_0(\mathbf{u}_2) \\ P_0(\mathbf{u}_1) - P_0(\mathbf{u}_2) \end{bmatrix}^\mathsf{T} (\mathbf{u}_1 - \mathbf{u}_2)$$

leads to a quadratic form that is not strongly convex. Lipschitz continuity of the map in its domain can be shown by noticing that:

$$\begin{aligned}
&\|\boldsymbol{\Phi}^{t_k}(\mathbf{z}_1) - \boldsymbol{\Phi}^{t_k}(\mathbf{z}_2)\|_2 \\
&\le (L+\nu)\|\mathbf{u}_1 - \mathbf{u}_1\|_2 + 2(G_g + G_0)\|\mathbf{u}_1 - \mathbf{u}_1\|_2 \\
&\quad + \epsilon\|\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2\|_2 + \epsilon\|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|_2 + \epsilon|\lambda_1 - \lambda_2| + \epsilon|\zeta_1 - \zeta_2| \\
&\quad + \sum_{i=1}^{M}(|\gamma_{1,i} - \gamma_{2,i}|\|\nabla_{\mathbf{u}} g_i^{t_k}\|_2 + |\mu_{1,i} - \mu_{2,i}|\|\nabla_{\mathbf{u}}\bar{g}_i^{t_k}\|_2) \\
&\quad + |\lambda_1 - \lambda_2|\|\nabla_{\mathbf{u}} P_0\|_2 + |\zeta_1 - \zeta_2|\|\nabla_{\mathbf{u}} P_0\|_2 && (42a)\\
&\le (L + \nu + 2G_g + 2G_0)\|\mathbf{u}_1 - \mathbf{u}_1\|_2 \\
&\quad + \epsilon\|\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2\|_2 + \epsilon\|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|_2 + \epsilon|\lambda_1 - \lambda_2| + \epsilon|\zeta_1 - \zeta_2| \\
&\quad + G_0(\|\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2\|_2 + \|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|_2) \\
&\quad + G_0(|\lambda_1 - \lambda_2| + |\zeta_1 - \zeta_2|) && (42b)
\end{aligned}$$

where we used the fact that $\|\nabla_{\mathbf{x}} h(\mathbf{x}_1) - \nabla_{\mathbf{x}} h(\mathbf{x}_2)\|_2 = 0$ when the function $h(\mathbf{x})$ is linear in $\mathbf{x}$ to derive (42a), and

we used Hölder inequality to obtain (42b) from (42a). With $G = \max\{G_g, G_0\}$ and using Hölder inequality, it follows that

$$\|\mathbf{\Phi}^{t_k}(\mathbf{z}_1) - \mathbf{\Phi}^{t_k}(\mathbf{z}_2)\|_2 \leq \bar{B}\|\mathbf{z}_1 - \mathbf{z}_1\|_2\,. \tag{43}$$

The results of Lemma 2 are utilized next to prove the bound in (18). To this end, define the time-varying operator $\mathbf{\Phi}_e^{t_k}$ as:

$$\mathbf{\Phi}_e^{t_k} : \{\mathbf{z}^{t_k}\} \mapsto \begin{bmatrix} \nabla_{[\mathbf{p},\mathbf{q}]}\mathcal{L}^{t_k}(\mathbf{p},\mathbf{q},\mathbf{d})|_{\hat{\mathbf{p}}^{t_k},\hat{\mathbf{q}}^{t_k},\mathbf{d}^{t_k}} \\ -(V^{\min}\mathbf{1} - |\hat{\mathbf{v}}^{t_k}| - \epsilon\boldsymbol{\gamma}^{t_k}) \\ -(|\hat{\mathbf{v}}^{t_k}| - V^{\max}\mathbf{1} - \epsilon\boldsymbol{\mu}^{t_k}) \\ -(\hat{P}_0^{t_k} - P_{0,\text{set}}^{t_k} - E - \epsilon\lambda^{t_k}) \\ -(P_{0,\text{set}}^{t_k} - \hat{P}_0^{t_k} - E - \epsilon\zeta^{t_k}) \end{bmatrix},$$

where $\hat{\mathbf{v}}^{t_k}$ is a vector collecting measurements of the voltage collected at time t_k. Using $\mathbf{\Phi}_e^{t_k}$, the steps of the algorithm can be compactly rewritten as

$$\mathbf{z}^{t_{k+1}} = \text{proj}_{\mathcal{Y}^{t_k}\times\mathbb{R}_+^M\times\mathbb{R}_+^M\times\mathbb{R}_+\times\mathbb{R}_+}\left\{\mathbf{z}^{t_k} - \alpha\mathbf{\Phi}_e^{t_k}(\mathbf{z}^{t_k})\right\}. \tag{44}$$

By standard optimality conditions, the optimizer is a fixed point of the iterations (44), i.e., $\mathbf{z}^{*,t_{k-1}} = \text{proj}_{\mathcal{Y}^{t_{k-1}}\times\mathbb{R}_+^M\times\mathbb{R}_+^M\times\mathbb{R}_+\times\mathbb{R}_+}\left\{\mathbf{z}^{*,t_{k-1}} - \alpha\mathbf{\Phi}^{t_k}(\mathbf{z}^{*,t_{k-1}})\right\}$. consider then writing:

$$\begin{aligned}\|\mathbf{z}^{t_k} - \mathbf{z}^{*,t_{k-1}}\|_2 = \\ \Big\|\text{proj}_{\mathcal{Y}^{t_{k-1}}\times\mathbb{R}_+^M\times\mathbb{R}_+^M\times\mathbb{R}_+\times\mathbb{R}_+}\left\{\mathbf{z}^{t_{k-1}} - \alpha\mathbf{\Phi}_e^{t_{k-1}}(\mathbf{z}^{t_{k-1}})\right\} \\ - \text{proj}_{\mathcal{Y}^{t_{k-1}}\times\mathbb{R}_+^M\times\mathbb{R}_+^M\times\mathbb{R}_+\times\mathbb{R}_+}\left\{\mathbf{z}^{*,t_{k-1}} - \alpha\mathbf{\Phi}^{t_{k-1}}(\mathbf{z}^{*,t_{k-1}})\right\}\Big\|_2\end{aligned} \tag{45}$$

and utilize the non-expansivity property of the projection operator to obtain

$$\begin{aligned}\|\mathbf{z}^{t_k} - \mathbf{z}^{*,t_{k-1}}\|_2 \leq \|\mathbf{z}^{t_{k-1}} - \alpha\mathbf{\Phi}_e^{t_{k-1}}(\mathbf{z}^{t_{k-1}}) \\ - \mathbf{z}^{*,t_{k-1}} + \alpha\mathbf{\Phi}^{t_{k-1}}(\mathbf{z}^{*,t_k})\|_2\,.\end{aligned} \tag{46}$$

By utilizing real measurements in the map $\mathbf{\Phi}_e^{t_{k-1}}(\mathbf{z}^{t_{k-1}})$, it follows that:

$$\mathbf{\Phi}_e^{t_{k-1}}(\mathbf{z}^{t_{k-1}}) - \mathbf{\Phi}^{t_{k-1}}(\mathbf{z}^{t_{k-1}}) = \mathbf{e}^{t_{k-1}} \tag{47}$$

where the vector $\mathbf{e}^{t_k}$ captures measurement errors as well as model mismatches, and it is defined as:

$$\mathbf{e}^{t_k} := \begin{bmatrix} \nabla_{[\mathbf{p},\mathbf{q}]}\mathcal{L}^{t_k}(\mathbf{p}^{t_k},\mathbf{q}^{t_k},\mathbf{d}^{t_k}) - \nabla_{[\mathbf{p},\mathbf{q}]}\mathcal{L}^{t_k}(\hat{\mathbf{p}}^{t_k},\hat{\mathbf{q}}^{t_k},\mathbf{d}^{t_k}) \\ \underline{\mathbf{g}}^{t_k}(\mathbf{p}^{t_k},\mathbf{q}^{t_k}) - (V^{\min}\mathbf{1} - |\hat{\mathbf{v}}^{t_k}|) \\ \bar{\mathbf{g}}^{t_k}(\mathbf{p}^{t_k},\mathbf{q}^{t_k}) - (|\hat{\mathbf{v}}^{t_k}| - V^{\max}\mathbf{1}) \\ P_0(\mathbf{p}^{t_k},\mathbf{q}^{t_k}) - \hat{P}_0^{t_k} \\ -P_0(\mathbf{p}^{t_k},\mathbf{q}^{t_k}) + \hat{P}_0^{t_k} \end{bmatrix}$$

From an optimization standpoint, the vector $\mathbf{e}^{t_k}$ models errors in the computation of the gradients that are due to measurements of voltages and powers at the substation. From Assumption 3 and Assumption 4, and using the result of Lemma 1, the norm of $\mathbf{e}^{t_k}$ can be bounded as:

$$\|\mathbf{e}^{t_k}\|_2^2 \leq (L+\nu)^2 e_p^2 + 2e_v^2 + 2e_0^2\,. \tag{48}$$

The proof now follows steps similar to [18]. Particularly, expand the right-hand side of (46) as

$$\begin{aligned}\|\mathbf{z}^{t_{k-1}} - \alpha\mathbf{\Phi}^{t_{k-1}}(\mathbf{z}^{t_{k-1}}) - \mathbf{z}^{*,t_{k-1}} + \alpha\mathbf{\Phi}^{t_{k-1}}(\mathbf{z}^{*,t_{k-1}}) - \alpha\mathbf{e}^{t_{k-1}}\|_2 \leq \\ \|\mathbf{z}^{t_{k-1}} - \alpha\mathbf{\Phi}^{t_{k-1}}(\mathbf{z}^{t_{k-1}}) - \mathbf{z}^{*,t_{k-1}} + \alpha\mathbf{\Phi}^{t_{k-1}}(\mathbf{z}^{*,t_{k-1}})\|_2 + \\ \|\alpha\mathbf{e}^{t_{k-1}}\|_2\,.\end{aligned} \tag{49}$$

Using the results of Lemma 2, we can write

$$\begin{aligned}\|\mathbf{z}^{t_{k-1}} - \alpha\Phi^{t_{k-1}}(\mathbf{z}^{t_{k-1}}) - \mathbf{z}^{*,t_{k-1}} + \alpha\Phi^{t_{k-1}}(\mathbf{z}^{*,t_{k-1}})\|_2^2 \leq \\ (1 - 2\alpha\eta + \alpha^2 L_{\nu,\epsilon}^2)\|\mathbf{z}^{t_{k-1}} - \mathbf{z}^{*,t_{k-1}}\|_2^2.\end{aligned} \tag{50}$$

and, by putting together the results in (46), (50), and (48), we have that

$$\begin{aligned}\|\mathbf{z}^{t_k} - \mathbf{z}^{*,t_{k-1}}\|_2 \leq \alpha\sqrt{(L+\nu)^2 e_p^2 + 2e_v^2 + 2e_0^2} \\ + \sqrt{1 - 2\alpha\eta + \alpha^2\bar{B}^2}\|\mathbf{z}^{t_{k-1}} - \mathbf{z}^{*,t_{k-1}}\|_2.\end{aligned} \tag{51}$$

Let $\rho(\alpha) := \sqrt{1 - 2\alpha\eta + \alpha^2\bar{B}^2}$ and notice that B in (18) is given by $B = \bar{B}^2$. Given a constant $\sigma \geq 0$ such that $\|\mathbf{z}^{*,t_{k+1}} - \mathbf{z}^{*,t_k}\| \leq \sigma$ for all $t_k \geq 0$, and by using the triangle inequality, it follows that

$$\begin{aligned}\|\mathbf{z}^{t_k} - \mathbf{z}^{*,t_k}\|_2 &= \|\mathbf{z}^{t_k} - \mathbf{z}^{*,t_k} - \mathbf{z}^{*,t_{k-1}} + \mathbf{z}^{*,t_{k-1}}\|_2 \\ &\leq \|\mathbf{z}^{t_k} - \mathbf{z}^{*,t_{k-1}}\|_2 + \sigma \\ &\leq \rho(\alpha)\|\mathbf{z}^{t_{k-1}} - \mathbf{z}^{*,t_{k-1}}\|_2 + \alpha e + \sigma.\end{aligned} \tag{52}$$

If $\rho(\alpha) < 1$, then (52) is a contraction and (18) readily follows.

Notice that an upper bound on the tracking error at each iteration can be obtained from (52).

D. Application to multiphase systems

Consider an approximate linear model

$$|\mathbf{v}| \approx \mathbf{A}\mathbf{p}_{\text{inj}} + \mathbf{B}\mathbf{q}_{\text{inj}} + \mathbf{c}\,, \tag{53}$$

$$\left[\ P_0^a, Q_0^a, P_0^b, Q_0^b, P_0^c, Q_0^c\ \right]^{\mathsf{T}} \approx \mathbf{M}\mathbf{p}_{\text{inj}} + \mathbf{N}\mathbf{q}_{\text{inj}} + \mathbf{o}\,, \tag{54}$$

where $\mathbf{v}$ collects the voltages *per phase* and per node, $\mathbf{p}_{\text{inj}}$ and $\mathbf{q}_{\text{inj}}$ are vectors collecting the net injected active and reactive powers per phase and per node (with $\mathbf{A}$, $\mathbf{B}$, $\mathbf{M}$, $\mathbf{N}$, $\mathbf{a}$, $\mathbf{o}$ of appropriate dimensions), and where (P_0^ϕ, Q_0^ϕ) denote the active and reactive powers flowing into the feeder on phase ϕ. Model (53)–(54) can be obtained as shown in [22], [35] ór by following steps similar to Appendix A. Suppose that a setpoint for for active power at the substation on *each phase* is given at time t_k, and denote the 3×1 vector collecting the setpoints as $\mathbf{p}_{0,\text{set}}^{t_k} := [P_{0,\text{set}}^{a,t_k}, P_{0,\text{set}}^{b,t_k}, P_{0,\text{set}}^{c,t_k}]$.

With $(P_i^{\phi,t_k}, Q_i^{\phi,t_k})$ denoting the setpoint for a DER at phase $\phi \subseteq \{a,b,c\}$ of node i, consider the implementation of algorithm (12) per phase and node shown next:

[S1a] Collect voltage-magnitude measurements $\{|\hat{V}_n^{\phi,t_k}|\}_{\phi\subseteq\{a,b,c\},n\in\mathcal{M}}$.

[S1b] Collect measurement of $\hat{P}_0^{\phi,t_k}$, $\phi = \{a,b,c\}$.

[S2a] For every phase ϕ of node $n \in \mathcal{M}$, update $\gamma_n^{\phi,t_{k+1}}$ and $\mu_n^{\phi,t_{k+1}}$ as follows:

$$\gamma_n^{\phi,t_{k+1}} = \text{proj}_{\mathbb{R}_+}\left\{\gamma_n^{\phi,t_k} + \alpha\left(V^{\min} - |\hat{V}_n^{\phi,t_k}| - \epsilon\gamma_n^{\phi,t_k}\right)\right\} \tag{55a}$$

$$\mu_n^{\phi,t_{k+1}} = \text{proj}_{\mathbb{R}_+}\left\{\mu_n^{\phi,t_k} + \alpha\left(|\hat{V}_n^{\phi,t_k}| - V^{\max} - \epsilon\mu_n^{\phi,t_k}\right)\right\} \tag{55b}$$

[S2b] For the feeder head, if $h^{t_k} = 1$ update dual variables associated with each phase as:

$$\lambda^{\phi,t_{k+1}} = \text{proj}_{\mathbb{R}_+}\left\{\lambda^{\phi,t_k} + \alpha(\hat{P}_0^{\phi,t_k} - P_{0,\text{set}}^{\phi,t_k} - E^{t_k} - \epsilon\lambda^{\phi,t_k})\right\} \tag{55c}$$

$$\zeta^{\phi,t_{k+1}} = \text{proj}_{\mathbb{R}_+}\left\{\zeta^{\phi,t_k} + \alpha(P_{0,\text{set}}^{\phi,t_k} - \hat{P}_0^{\phi,t_k} - E^{t_k} - \epsilon\zeta^{\phi,t_k})\right\} \tag{55d}$$

[S3a] Measure output powers $\hat{P}_i^{\phi,t_k}, \hat{Q}_i^{\phi,t_k}$ at DER at phase ϕ of $i \in \mathcal{G}$.

[S3b] Update power setpoints for each DER as:

$$\begin{bmatrix} P_i^{\phi,t_{k+1}} \\ Q_i^{\phi,t_{k+1}} \end{bmatrix} = \text{proj}_{\mathcal{Y}_i^{\phi,t_k}}\left\{\begin{bmatrix} P_i^{\phi,t_k} \\ Q_i^{\phi,t_k} \end{bmatrix} - \alpha\nabla_{[P_i^\phi, Q_i^\phi]}\mathcal{L}^{t_k}(\mathbf{p},\mathbf{q},\mathbf{d})|_{\hat{P}_i^{\phi,t_k}, \hat{Q}_i^{\phi,t_k}, \mathbf{d}^{t_{k+1}}}\right\}, \tag{55e}$$

[S3c] Command setpoints to each DER and return to **[S1a]**.

The convergence result (18) of Theorem 1 applies to the algorithm (55) upon redefining the error bound e_0 in *Assumption 4* as:

$$\|\check{\mathbf{M}}\mathbf{p}_{\text{inj}}^{t_k} + \check{\mathbf{N}}^p\mathbf{q}_{\text{inj}}^{t_k} + \check{\mathbf{o}}^{t_k} - \hat{\mathbf{p}}_0^{t_k}\|_2 \leq e_0 \tag{56}$$

where $\hat{\mathbf{p}}_0^{t_k} := [\hat{P}_0^{a,t_k}, \hat{P}_0^{b,t_k}, \hat{P}_0^{c,t_k}]^\intercal$, $\check{\mathbf{M}}$ and $\check{\mathbf{N}}$ are matrices collecting the rows 1, 3, and 5 of matrices $\mathbf{M}$ and $\mathbf{N}$, respectively, and $\check{\mathbf{o}}^{t_k} := [o_1^{t_k}, o_3^{t_k}, o_5^{t_k}]^\intercal$.

REFERENCES

[1] V. Donde, M. A. Pai, and I. A. Hiskens, "Simulation and optimization in an AGC system after deregulation," *IEEE Trans. on Power Systems*, vol. 16, no. 3, pp. 481–489, Aug 2001.

[2] K. H. Abdul-Rahman, H. Alarian, M. Rothleder, P. Ristanovic, B. Vesovic, and B. Lu, "Enhanced system reliability using flexible ramp constraint in CAISO market," in *IEEE Power and Energy Society General Meeting*, July 2012.

[3] J. Koshal, A. Nedić, and U. Y. Shanbhag, "Multiuser optimization: Distributed algorithms and error analysis," *SIAM J. on Optimization*, vol. 21, no. 3, pp. 1046–1081, 2011.

[4] A. Simonetto and G. Leus, "Double smoothing for time-varying distributed multiuser optimization," in *IEEE Global Conf. on Signal and Information Processing*, Dec. 2014.

[5] E. Vrettos, J. L. Mathieu, and G. Andersson, "Control of thermostatic loads using moving horizon estimation of individual load states," in *Power Systems Computation Conf.,*, Aug 2014.

[6] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent, "Aggregate flexibility of thermostatically controlled loads," *IEEE Trans. on Power Systems*, vol. 30, no. 1, pp. 189–198, Jan 2015.

[7] S. J. Crocker and J. L. Mathieu, "Adaptive state estimation and control of thermostatic loads for real-time energy balancing," in *2016 American Control Conf.*, July 2016, pp. 3557–3563.

[8] S. P. Meyn, P. Barooah, A. Bušić, Y. Chen, and J. Ehren, "Ancillary service to the grid using intelligent deferrable loads," *IEEE Trans. on Automatic Control*, vol. 60, no. 11, pp. 2847–2862, Nov. 2015.

[9] E. Vrettos, F. Oldewurtel, and G. Andersson, "Robust energy-constrained frequency reserves from aggregations of commercial buildings," *IEEE Trans. on Power Systems*, vol. 31, no. 6, pp. 4272–4285, Nov 2016.

[10] M. A. Ortega-Vazquez, F. Bouffard, and V. Silva, "Electric vehicle Aggregator/System operator coordination for charging scheduling and services procurement," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1806–1815, May 2013.

[11] J. T. Hughes, A. D. Domínguez-García, and K. Poolla, "Identification of virtual battery models for flexible loads," *IEEE Trans. on Power Systems*, vol. 31, no. 6, pp. 4660–4669, Nov 2016.

[12] A. Cherukuri and J. Cortés, "Distributed coordination of DERs with storage," 2016, [Online] Available at: http://arxiv.org/pdf/1605.00721.

[13] D. Jigoria-Oprea, G. Vuc, and M. Litcanu, "Optimal management of a virtual power plant," in *IEEE International Conference on Environment and Electrical Engineering*, June 2016.

[14] A. Baringo and L. Baringo, "A stochastic adaptive robust optimization approach for the offering strategy of a virtual power plant," *IEEE Trans. on Power Systems*, 2016.

[15] A. Bernstein, L. Reyes-Chamorro, J.-Y. 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.

[16] 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," [Online] Available at: https://arxiv.org/abs/1511.08628.

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

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

[19] Y. Liu, J. Hours, G. Stathopoulos, and C. Jones, "Real-time distributed algorithms for nonconvex optimal power flow," in *American Control Conf.*, May 2017, [Online] Available at: https://infoscience.epfl.ch/record/225467.

[20] W. H. Kersting, *Distribution System Modeling and Analysis*. 2nd ed., Boca Raton, FL: CRC Press, 2007.

[21] S. S. Guggilam, E. Dall'Anese, Y. C. Chen, S. V. Dhople, and G. B. Giannakis, "Scalable optimization methods for distribution networks with high PV integration," *IEEE Trans. on Smart Grid*, vol. 7, no. 4, pp. 2061–2070, July 2016.

[22] 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.

[23] S. Bolognani and F. Dörfler, "Fast power system analysis via implicit linearization of the power flow manifold," in *2015 53rd Annual Allerton Conf. on Communication, Control, and Computing*, 2015, pp. 402–409.

[24] 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.

[25] D. Angelosante, J. A. Bazerque, and G. B. Giannakis, "Online adaptive estimation of sparse signals: Where RLS meets the ℓ_1-norm," *IEEE Trans. on Signal Processing*, vol. 58, no. 7, pp. 3436–3447, July 2010.

[26] A. Jokić, M. Lazar, and 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.

[27] D. P. Bertsekas and J. N. Tsitsiklis, "Gradient convergence in gradient methods with errors," *SIAM J. on Optimization*, vol. 10, no. 3, pp. 627–642, Jul. 1999.

[28] A. Yazdani and R. Iravani, *Voltage-Sourced Converters in Power Systems: Modeling, Control, and Applications*. John Wiley & Sons, 2010.

[29] H. Li, X. Yan, S. Adhikari, D. T. Rizy, F. Li, and P. Irminger, "Real and reactive power control of a three-phase single-stage PV system and PV voltage stability," in *PES General Meeting*, San Diego, CA, 2012.

[30] E. Polymeneas and S. Meliopoulos, "Aggregate modeling of distribution systems for multi-period OPF," in *Power Systems Computation Conference*, June 2016, pp. 1–8.

[31] S. Barot and J. A. Taylor, "A concise, approximate representation of a collection of loads described by polytopes," *International Journal of Electrical Power & Energy Systems*, vol. 84, pp. 55–63, 2017.

[32] J. Bank and J. Hambrick, "Development of a high resolution, real time, distribution-level metering system and associated visualization modeling, and data analysis functions," National Renewable Energy Laboratory, Tech. Rep. NREL/TP-5500-56610, May 2013.

[33] G. Heffner, C. Goldman, B. Kirby, and M. Kintner-Meyer, "Loads providing ancillary services: Review of international experience," in *Technical Report LBNL–62701* , May 2007.

[34] R. Hledik and J. Lazar, "Distribution system pricing with distributed energy resources," in *Technical Report LBNL–1005180* , May 2016.

[35] A. Bernstein, C. Wang, E. Dall'Anese, J.-Y. Le Boudec, and C. Zhao, "Load-flow in multiphase distribution networks: Existence, uniqueness, and linear models," 2017, [Online] Available at: http://arxiv.org/abs/1702.03310.

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 Theory and Signal Processing. Current applications pertain to distributed optimization and control of large-scale complex energy systems and power-grid analytics.

Swaroop S. Guggilam (S'15) received the Bachelor's degree in electrical engineering from the Veermata Jijabai Technological Institute, Mumbai, Maharashtra, India, in 2013 and M.S. degree in electrical engineering from the University of Minnesota, Minneapolis, MN, USA, in 2015. He is currently working toward the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, MN, USA.

His research interests include distribution networks, optimization, power system modeling, analysis, and control for increasing renewable integration.

Andrea Simonetto (M'12) received the Ph.D. degree in systems and control from Delft University of Technology, Delft, The Netherlands, in 2012. He is currently a research staff member in the optimization and control group at IBM Research Ireland, Dublin, Ireland.

His research interests include large-scale centralized and distributed optimization with applications in smart energy and intelligent transportation systems.

Yu Christine Chen (S'10–M'15) received the B.A.Sc. degree in engineering science from the University of Toronto, Toronto, ON, Canada, in 2009, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 2011 and 2014, respectively.

She is currently an Assistant Professor with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada, where she is affiliated with the Electric Power and Energy Systems Group. Her research interests include power system analysis, monitoring, and control.

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.