Feedback-Based Projected-Gradient Method for Real-Time Optimization of Aggregations of Energy Resources

Preprint

Emiliano Dall'Anese and Andrey Bernstein
National Renewable Energy Laboratory

Andrea Simonetto
IBM Research Ireland

Presented at the 5th IEEE Global Conference on Signal and Information Processing (GlobalSIP)
Montreal, Quebec, Canada
November 14–16, 2017

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Conference Paper
NREL/CP-5D00-68500
November 2017

Contract No. DE-AC36-08GO28308

NOTICE

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at SciTech Connect http:/www.osti.gov/scitech

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
OSTI http://www.osti.gov
Phone: 865.576.8401
Fax: 865.576.5728
Email: reports@osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312
NTIS http://www.ntis.gov
Phone: 800.553.6847 or 703.605.6000
Fax: 703.605.6900
Email: orders@ntis.gov

Cover Photos by Dennis Schroeder: (left to right) NREL 26173, NREL 18302, NREL 19758, NREL 29642, NREL 19795.

NREL prints on paper that contains recycled content.

FEEDBACK-BASED PROJECTED-GRADIENT METHOD FOR REAL-TIME OPTIMIZATION OF AGGREGATIONS OF ENERGY RESOURCES

Emiliano Dall'Anese[1]*, Andrey Bernstein*[1]*, and Andrea Simonetto*[2]

[1]National Renewable Energy Laboratory, Golden, CO, USA

[2]IBM Research Ireland, Dublin, Ireland

ABSTRACT

This paper develops an online optimization method to maximize operational objectives of distribution-level distributed energy resources (DERs), while adjusting the aggregate power generated (or consumed) in response to services requested by grid operators. The design of the online algorithm is based on a projected-gradient method, suitably modified to accommodate appropriate measurements from the distribution network and the DERs. By virtue of this approach, the resultant algorithm can cope with inaccuracies in the representation of the AC power flows, it avoids pervasive metering to gather the state of noncontrollable resources, and it naturally lends itself to a distributed implementation. Optimality claims are established in terms of tracking of the solution of a well-posed time-varying convex optimization problem.

1. INTRODUCTION

We address the problem of optimizing the operation of aggregations of heterogeneous energy resources connected to (a portion of) a distribution system — a research task that has gained significant interest from academic and industrial sectors because of the increased deployment of distributed energy resources (DERs) along with the shaping of (albeit futuristic) distribution-level ancillary-service markets. We focus on *real-time* optimization methods, wherein the term "real time" refers to an operational setting in which the power setpoints of the DERs are updated on a second or subsecond timescale to maximize the operational objectives while coping with the variability of ambient conditions and noncontrollable energy assets.

A main challenges in this context is related to the computational complexity, which may render infeasible the solution of optimization problems on a second or subsecond timescale to compute the optimal power setpoints to the DERs [1–3]. Further, when the optimization problem is solved in a distributed fashion, multiple communication rounds are necessary in order to converge to the optimal setpoints for the DERs. Another challenge is related to the need for pervasive metering to collect measurements of the state of non-controllable assets, which serve as inputs to the optimization task.

To address these challenges, we start from the formulation of a time-varying convex optimization problem that takes into account the operational objectives of DERs as well as target setpoints for the aggregate power provided by the DERs. We then design an online algorithm based on a projected-gradient method by suitably modifying the gradient method to accommodate appropriate measurements from the distribution network and the DERs. The resultant algorithm can cope with inaccuracies in the representation of the AC power flows, avoids pervasive metering to gather the state of non-controllable resources, and affords a distributed implementation.

The idea of accommodating measurements into primal-dual-type methods goes back to [4–6] and it was applied to real-time optimization of power systems in [7], wherein a centralized controller was developed based on projected-gradient methods. Online algorithms were developed in [8], [9], [10], and [11] to find solutions of AC optimal power flow (OPF) problems, with [10] establishing results in terms of tracking of solutions of a time-varying linearized AC OPF and [11] in terms of tracking of solutions of a time-varying relaxed AC OPF. Recently, a projected-gradient method on the (static) power flow manifold was proposed in [12, 13].

The method outlined in this paper builds on [7, 10], but: i) it is implemented for multiphase systems; ii) it accounts for DERs with nonconvex operational sets; and, iii) it involves a simplified implementation that can be used even in the case when only active powers are controlled using the proposed algorithm but reactive powers follow existing Volt/VAr rules [14]. Related to ii), the operational sets of DERs are convexified for the purpose of setpoint computation, whereas implementable setpoints are computed by utilizing a variant of the error-diffusion algorithm [15–17].

2. PROBLEM FORMULATION

Consider a multiphase distribution network consisting of one slack bus and N PQ buses. Leveraging the model proposed in [18], each multiphase node $j \in \mathcal{N} := \{1, \ldots, N\}$ can feature wye-connected or delta-connected (aggregations of) DERs. For ease of exposition, suppose that all the nodes are three-phase, and let $\mathbf{s}_j^Y := (s_j^a, s_j^b, s_j^c)^\mathsf{T}$ denote the vector net injected powers at phases $\phi \in \{a, b, c\}$ at node j from grounded-wye-connected DERs. Similarly, let $\mathbf{s}_j^\Delta := (s_j^{ab}, s_j^{bc}, s_j^{ca})^\mathsf{T}$ denote the power injections of the delta-connected sources, where $s_j^\phi \in \mathbb{C}$, $\phi \in \{ab, bc, ca\}$, denotes the net complex power injected from delta-connected DERs. When devices are modeled at the secondary side of the distribution transformer, $\mathbf{s}_j^Y$ and $\mathbf{s}_j^\Delta$ denote the line-to-line and line-to-ground connections, respectively [18]. Finally, we let $\mathbf{p}_0 := (p_0^a, p_0^b, p_0^c)^\mathsf{T} \in \mathbb{R}^3$ denote the active power flow at the point of connection with the rest of the grid.

For future developments, let $\mathbf{x}_{j,\phi} := (\Re\{s_j^\phi\}, \Im\{s_j^\phi\})^\mathsf{T}$ collect the active and reactive setpoints of the DER at node j, where $\phi \in \{a, b, c\}$ for wye connections and $\phi \in \{ab, bc, ca\}$ for delta connections. To further simplify the notation, consider stacking in the vector $\mathbf{x} := (\{\mathbf{x}_{j,\phi}\}, \phi \in \{a, b, c\} \cup \{ab, bc, ca\}, j = 1, \ldots, N)^\mathsf{T} \in \mathbb{R}^{12N}$ the active and reactive power setpoint of DERs at all phases and nodes. If no controllable DERs are present at a given location, the corresponding vector $\mathbf{x}_{j,\phi}$ is trivially set to $\mathbf{0}$.

To facilitate the development of computationally-affordable algorithms, we postulate the following approximate linear relation-

The work of E. Dall'Anese and A. Bernstein was supported by the Advanced Research Projects Agency-Energy (ARPA-E) under the Network Optimized Distributed Energy Systems (NODES) program.

ship:

$$\widetilde{\mathbf{p}}_0(\mathbf{x}) = \mathbf{M}(\mathbf{x}-\mathbf{x}_\ell)+\mathbf{m} := \sum_{j=1}^{N}\sum_{\phi\in\mathcal{P}} \mathbf{M}_{j,\phi}(\mathbf{x}_{j,\phi}-\mathbf{x}_{j,\phi,\ell})+\mathbf{m}, \quad (1)$$

where $\mathcal{P} := \{a, b, c\} \cup \{ab, bc, ca\}$ for brevity, $\mathbf{x}_{j,\phi,\ell}$ represents the noncontrollable loads, and the model parameters $\{\mathbf{M}_{j,\phi}\} \in \mathbb{R}^{3\times 3}$ and $\mathbf{m} \in \mathbb{R}^3$ can be obtained as shown in, e.g., [19, 20].

Consider a discrete-time operational setting wherein the setpoints of the DERs are updated at time instants $t_k = \delta k$, with $k \in \mathbb{N}$ and sampling period $\delta > 0$ based on the specific implementation requirements (e.g., second, subsecond, or a few seconds) [7, 8, 10]. At each time t_k, the performance objectives of a DER at phase $\phi \in \mathcal{P}$ and node j are modeled by means of a time-varying convex and differentiable function $C_{j,\phi}^{(k)} : \mathbb{R}^2 \to \mathbb{R}$, which is to be minimized at each time t_k. The set of feasible power setpoints of the same DER is denoted as $\mathcal{Y}_{j,\phi}^{(k)} \subseteq \mathbb{R}^2$. We assume that the set $\mathcal{Y}_{j,\phi}^{(k)}$ is convex and compact; we will address nonconvex sets (and, in particular, discrete sets) in Section 3.2. When a group of DERs are located at a node, $s_j^\phi \in \mathbb{C}$ represents the net power generated, and $\mathcal{Y}_{j,\phi}^{(k)}$ is given by the Minkowski sum of the operating regions of the individual DERs. Finally, consider a setpoint $\mathbf{p}_0^{\text{set},(k)}$ for the aggregate power at the point of connection $\mathbf{p}_0(\mathbf{x}^{(k)})$, which can be given by a utility company or an aggregator incentivizing customers to provide services.

With these definitions in place, consider the following optimization problem to compute the DERs' setpoints at time t_k:

$$\text{(P1)}^{(k)} \quad \min_{\mathbf{x}} \quad \sum_{j=1}^{N}\sum_{\phi\in\mathcal{P}} C_{j,\phi}^{(k)}(\mathbf{x}_{j,\phi}) + \frac{\gamma}{2}\left\|\mathbf{p}_0^{\text{set},(k)} - \widetilde{\mathbf{p}}_0(\mathbf{x})\right\|_2^2 \quad (2a)$$

$$\text{s. to :} \quad \mathbf{x}_{j,\phi} \in \mathcal{Y}_{j,\phi}^{(k)}, \ \forall\ \phi \in \mathcal{P},\ j \in \mathcal{N} \quad (2b)$$

where $\widetilde{\mathbf{p}}_0(\mathbf{x})$ is given by (1), and $\gamma > 0$ is a design parameter that influences the ability to track the reference signal $\{\mathbf{p}_0^{\text{set},(k)}, k \in \mathbb{N}\}$.

Problem $(\text{P1})^{(k)}$ is a *time-varying* convex optimization problem; however, solving $(\text{P1})^{(k)}$ in batch fashion at each time t_k might be impractical because of the following three main challenges:

- **c1: Complexity.** For real-time implementations (e.g., when δ is on the order of a second or subsecond), it might be unfeasible to solve $(\text{P1})^{(k)}$ to convergence, especially in a distributed setting.
- **c2. Model inaccuracy.** The linear model (1) provides only an approximate relationship between power injections and $\mathbf{p}_0(\mathbf{x}^{(k)})$. As a result, the optimal solution of $(\text{P1})^{(k)}$ might not necessarily track $\mathbf{p}_0^{\text{set},(k)}$ accurately enough.
- **c3. Pervasive metering.** Solving $(\text{P1})^{(k)}$ requires collecting measurements of the (aggregate) noncontrollable loads $\mathbf{x}_{j,\phi,\ell}$ at all locations in real time.

We will design a *feedback-based online algorithm* that tracks the optimal solution of $(\text{P1})^{(k)}$ over time, while coping with model inaccuracies and avoiding ubiquitous monitoring.

3. FEEDBACK-BASED ONLINE ALGORITHM

3.1. Design of the Online Algorithm

We begin with a modeling assumption pertaining to the cost functions functions $\{C_{j,\phi}^{(k)}(\mathbf{x}_{j,\phi})\}$.

Assumption 1. The functions $C_{j,\phi}^{(k)}$ are convex and continuously differentiable for each ϕ, j and k. Let

$$\mathbf{f}^{(k)}(\mathbf{x}) := \left(\left\{\nabla C_{j,\phi}^{(k)}(\mathbf{x}_{j,\phi}), \phi \in \mathcal{P}, j = 1, \ldots, N\right\}\right)^{\mathsf{T}} \quad (3)$$

denote the column-vector gradient map. We assume that:

A1.i. There exists $\eta > 0$ such that $(\mathbf{f}^{(k)}(\mathbf{x}) - \mathbf{f}^{(k)}(\mathbf{x}'))^{\mathsf{T}}(\mathbf{x}-\mathbf{x}') \geq \eta\|\mathbf{x}-\mathbf{x}'\|_2^2$, for all $\mathbf{x}, \mathbf{x}' \in \mathcal{Y}^{(k)}$, where $\mathcal{Y}^{(k)}$ is given by the Cartesian product of the sets $\{\mathcal{Y}_{j,\phi}^{(k)}, \phi \in \mathcal{P}, j = 1, \ldots N\}$; and,

A1.ii. The gradient map $\mathbf{f}^{(k)}$ is Lipschitz continuous with constant L over the compact set $\mathcal{Y}^{(k)}$ for all k; that is $\|\mathbf{f}^{(k)}(\mathbf{x}) - \mathbf{f}^{(k)}(\mathbf{x}')\|_2 \leq L\|\mathbf{x}-\mathbf{x}'\|_2$ for all $\mathbf{x}, \mathbf{x}' \in \mathcal{Y}^{(k)}$. □

Assumption **A1.i** states that the cost $\sum_{j=1}^{N}\sum_{\phi\in\mathcal{P}} C_{j,\phi}^{(k)}(\mathbf{x}_{j,\phi})$ is strongly convex in $\mathbf{x}$. Notice that if the cost function specified for the DERs is not strongly convex, a regularization term such as $(\eta/2)\|\mathbf{x} - \mathbf{x}_{\text{reg}}\|_2^2$, where $\mathbf{x}_{\text{reg}}$ is a given point and $\eta > 0$, can be added; see e.g., [10, 21].

Consider then the following projected gradient method to *track* the optimal solution of $(\text{P1})^{(k)}$ over time:

$$\mathbf{x}_{j,\phi}^{(k+1)} = \mathsf{Proj}_{\mathcal{Y}_{j,\phi}^{(k)}}\Big\{\mathbf{x}_{j,\phi}^{(k)} - \alpha\Big(\nabla C_{j,\phi}^{(k)}(\mathbf{x}_{j,\phi}^{(k)}) + \gamma\mathbf{M}_{j,\phi}^{\mathsf{T}}\left[\widetilde{\mathbf{p}}_0(\mathbf{x}^{(k)}) - \mathbf{p}_0^{\text{set},k}\right]\Big)\Big\} \quad \forall\ \phi \in \mathcal{P}, j \in \mathcal{N} \quad (4)$$

where $\alpha > 0$ is the stepsize and $\mathsf{Proj}_{\mathcal{Y}}\{\mathbf{y}\} := \arg\min_{\mathbf{x}\in\mathcal{Y}} \|\mathbf{x}-\mathbf{y}\|_2$ is the projection of $\mathbf{y}$ onto the convex set $\mathcal{Y}$. By suitably choosing the stepsize, (4) can address the challenge c1 identified above by provably tracking the time-varying solution of $(\text{P1})^{(k)}$ [22]. However, (4) does not address c2-c3, as it relies on an approximate representation of the powers flows at the point of coupling and requires collecting measurements of $\mathbf{x}_{j,\phi,\ell}$ from all the nodes. Further, the projected gradient method (4) requires knowledge of the entire vector $\mathbf{x}^{(k)}$ at every DER j; thus, it does not afford a distributed implementation.

To address the challenges c2-c3 identified above and enable a distributed algorithmic implementation, the idea is to suitably modify (4) to accommodate *measurements* from the distribution system – i.e., the "feedback". To this end, let $\widehat{\mathbf{p}}_0^{(k)}$ and $\widehat{\mathbf{x}}_{j,\phi}^{(k)}$ denote, respectively, the measurement of the power flows at the point of connection and the output power of DER located at phase ϕ of node j at time step k. In the spirit of [7, 10], we then propose the following feedback-based online algorithm:

$$\mathbf{x}_{j,\phi}^{(k+1)} = \mathsf{Proj}_{\mathcal{Y}_{j,\phi}^{(k)}}\Big\{\widehat{\mathbf{x}}_{j,\phi}^{(k)} - \alpha\Big(\nabla C_{j,\phi}^{(k)}(\widehat{\mathbf{x}}_{j,\phi}^{(k)}) + \gamma\mathbf{M}_{j,\phi}^{\mathsf{T}}\left[\widehat{\mathbf{p}}_0^{(k)} - \mathbf{p}_0^{\text{set},k}\right]\Big)\Big\} \quad \forall\ \phi \in \mathcal{P}, j \in \mathcal{N}. \quad (5)$$

Note that the update (5) does not require information regarding the noncontrollable loads, as the effect of $\mathbf{x}_{j,\phi,\ell}$ on the power at the point of coupling is already captured in the measurement $\widehat{\mathbf{p}}_0^{(k)}$.

In lieu of the linearlized model $\widetilde{\mathbf{p}}_0(\mathbf{x})$, (5) utilized measurements $\widehat{\mathbf{p}}_0^{(k)}$ of the power at the point of coupling; the only information needed is the matrix $\mathbf{M}$ that models network-related sensitivities, and which can be updated at a slower timescale [7, 10]. Further, update (5) affords a distributed implementation provided that the measurement $\widehat{\mathbf{p}}_0^{(k)}$ is broadcasted to every DER. Convergence and tracking capabilities of (5) will be assessed in Section 4. But first, an extension to DERs with nonconvex operating regions is presented.

3.2. Nonconvex Resources

Suppose that $\mathcal{Y}_{j,\phi}^{(k)}$ is nonconvex for a given DER. This is the case, for example, for electric vehicles (EVs with discrete charging levels, energy storage systems with minimum power factor constraints, and electric water heaters,to mention a few. Consider then replacing $\mathcal{Y}_{j,\phi}^{(k)}$ in (5) with its *convex hull* $\mathsf{ch}\mathcal{Y}_{j,\phi}^{(k)}$. Once the potential setpoint

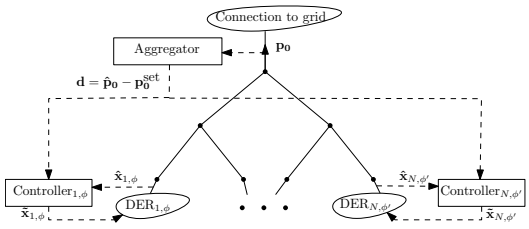

Fig. 1: Online algorithm: distributed implementation.

$\mathbf{x}_{j,\phi}^{(k+1)} \in \mathsf{ch}\mathcal{Y}_{j,\phi}^{(k)}$ is computed, an additional step is required to generate a feasible setpoint within the set $\mathcal{Y}^{(k)}$.

A naïve approach is to choose a *closest point* to $\mathbf{x}_{j,\phi}^{(k+1)}$ in $\mathcal{Y}_{j,\phi}^{(k)}$. However, as shown in, e.g., [23], this approach might lead to poor tracking performance. On the other hand, we propose selecting a feasible point $\widetilde{\mathbf{x}}_{j,\phi}^{(k+1)}$ that *on average* coincides with $\mathbf{x}_{j,\phi}^{(k+1)}$. This would lead a notion of optimality with respect to the optimal solution of the relaxed problem $(\text{P1})^{(k)}$ in terms of energy produced/consumed.

Once $\mathbf{x}_{j,\phi}^{(k+1)}$ is computed from (5) using $\mathsf{ch}\mathcal{Y}_{j,\phi}^{(k)}$, one of the following additional steps can be performed to obtain an implementable point $\widetilde{\mathbf{x}}_{j,\phi}^{(k+1)} \in \mathcal{Y}_{j,\phi}^{(k)}$:

(i) **Randomized rounding.** Choose $\widetilde{\mathbf{x}}_{j,\phi}^{(k+1)}$ randomly using a probability measure induced by the distance of $\mathbf{x}_{j,\phi}^{(k+1)}$ to points in $\mathcal{Y}_{j,\phi}^{(k)}$.

(ii) **Error diffusion.** Choose $\widetilde{\mathbf{x}}_{j,\phi}^{(k+1)}$ using the following rule:

$$\widetilde{\mathbf{x}}_{j,\phi}^{(k+1)} \in \mathsf{Proj}_{\mathcal{Y}_{j,\phi}^{(k)}}\{\mathbf{x}_{j,\phi}^{(k+1)} + \mathbf{e}_{j,\phi}^{(k)}\}, \tag{6}$$

where $\mathbf{e}_{j,\phi}^{(k)} := \sum_{\ell=1}^{k}(\mathbf{x}_{j,\phi}^{(\ell)} - \widetilde{\mathbf{x}}_{j,\phi}^{(\ell)})$ is the accumulated error up to time step k [23].

The first option is suitable for discrete sets $\mathcal{Y}_{j,\phi}^{(k)}$, whereas the second option is equally applicable to the cases of discrete and continuous sets. In this paper, we utilize the second option due to its simplicity and applicability to a large family of nonconvex sets.

4. ANALYSIS

Recall that $\mathcal{Y}^{(k)}$ is the Cartesian product of the sets $\{\mathcal{Y}_{j,\phi}^{(k)}, \phi \in \mathcal{P}, j = 1, \dots N\}$ and define the following mapping $\mathbf{\Phi}^{(k)} : \mathbb{R}^{12N} \to \mathbb{R}^{12N}$ as $\mathbf{\Phi}^{(k)}(\mathbf{x}) := \mathbf{f}^{(k)}(\mathbf{x}) + \gamma \mathbf{M}^{\mathsf{T}}\left[\widetilde{\mathbf{p}}_0(\mathbf{x}) - \mathbf{p}_0^{\text{set},k}\right]$. This way, the projected-gradient step (4) can be compactly rewritten as $\mathbf{x}^{(k+1)} = \mathsf{Proj}_{\mathcal{Y}^{(k)}}\{\mathbf{x}^{(k)} - \alpha \mathbf{\Phi}^{(k)}(\mathbf{x}^{(k)})\}$. Regarding the map $\mathbf{\Phi}^{(k)}(\mathbf{x})$, we have the following result.

Lemma 1. *Under Assumption 1, the mapping $\mathbf{\Phi}^{(k)}$ is strongly monotone with constant η, and is Lipschitz continuous over $\mathcal{Y}^{(k)}$ with constant $L_\Phi := L + \gamma\|\mathbf{M}^T\mathbf{M}\|_2$.*

Let $\mathbf{x}^{*,k}$ denote the unique optimizer of $(\text{P1})^{(k)}$ at time t_k. To characterize analytically the tracking properties of (5) in terms of distance from the optimal trajectory $\{\mathbf{x}^{*,k}, k \in \mathbb{N}\}$, define

$$\sigma^{(k)} := \|\mathbf{x}^{*,k+1} - \mathbf{x}^{*,k}\|_2 \tag{7}$$

and notice that the sequence $\{\sigma^{(k)}\}$ represents a measure of the time-variability of the optimization problem $(\text{P1})^{(k)}$. Further, since (5) relies on measurements, define the following quantities:

$$e_x^{(k)} := \|\mathbf{x}^{(k)} - \widehat{\mathbf{x}}^{(k)}\|_2, \quad e_0^{(k)} := \|\widetilde{\mathbf{p}}_0(\widehat{\mathbf{x}}^{(k)}) - \widehat{\mathbf{p}}_0^{(k)}\|_2\,. \tag{8}$$

Note that $e_x^{(k)}$ captures (i) measurements errors, (ii) discrepancy between setpoints and actual output powers of the DERs, and (iii) the error introduced by using the error-diffusion algorithm (6) in the case of nonconvex resources. Finally, $e_0^{(k)}$ captures measurement errors as well as approximation errors introduced by (1). We assume that the error sequences are uniformly bounded.

Assumption 2. There exist finite constants $\sigma^{\max}$, $e_x^{\max}$, and $e_0^{\max}$ such that, for all k, $\sigma^{(k)} \le \sigma^{\max}$, $e_x^{(k)} \le e_x^{\max}$, and $e_0^{(k)} \le e_0^{\max}$.

Tracking properties of (5) are characterized next.

Theorem 1. *Consider the sequence $\{\mathbf{x}^{(k)}\}$ generated by* (5)*. Let Assumption 1 hold. If the stepsize $\alpha > 0$ is chosen such that $\alpha < 2/L_\Phi$, then*

$$\|\mathbf{x}^{(k)} - \mathbf{x}^{*,k}\|_2 \le (\rho(\alpha))^k \|\mathbf{x}^{(0)} - \mathbf{x}^{*,0}\|_2 + \sum_{j=0}^{k-1}(\rho(\alpha))^j \Delta^{(k-j-1)} \tag{9}$$

where $\rho(\alpha) := \max\{|1 - \alpha\,\eta|, |1 - \alpha\,L_\Phi|\} < 1$, and $\Delta^{(k)} := \rho(\alpha)e_x^{(k)} + \alpha\gamma\|\mathbf{M}\|_2 e_0^{(k)} + \sigma^{(k)}$, and $e_x^{(k)}$, $e_0^{(k)}$, and $\sigma^{(k)}$ are given by (8) *and* (7)*, respectively.*

Proof. Define the following time-varying mapping

$$\mathbf{\Phi}_e^{(k)}(\mathbf{x}) := \mathbf{f}^{(k)}(\mathbf{x}) + \gamma \mathbf{M}^{\mathsf{T}}\left[\widehat{\mathbf{p}}_0^{(k)} - \mathbf{p}_0^{\text{set},k}\right] \tag{10}$$

which allows us to rewrite (5) as $\mathbf{x}^{(k+1)} = \mathsf{Proj}_{\mathcal{Y}^{(k)}}\{\widehat{\mathbf{x}}^{(k)} - \alpha \mathbf{\Phi}_e^k(\widehat{\mathbf{x}}^{(k)})\}$. Consider the norm $\|\mathbf{x}^{(k)} - \mathbf{x}^{*,k-1}\|_2$, which captures the distance between $\mathbf{x}^{(k)}$ and the optimizer $\mathbf{x}^{*,k-1}$ of $(\text{P1})^{(k-1)}$. By standard optimality conditions, we have $\mathbf{x}^{*,k-1} = \mathsf{Proj}_{\mathcal{Y}^{(k-1)}}\left\{\mathbf{x}^{*,k-1} - \alpha \mathbf{\Phi}^k(\mathbf{x}^{*,k-1})\right\}$. It follows then that

$$\begin{aligned}\|\mathbf{x}^{(k)} - \mathbf{x}^{*,k-1}\|_2 &= \Big\|\mathsf{Proj}_{\mathcal{Y}^{(k-1)}}\left\{\widehat{\mathbf{x}}^{(k-1)} - \alpha\mathbf{\Phi}_e^{(k-1)}(\widehat{\mathbf{x}}^{(k-1)})\right\} \\ &\quad - \mathsf{Proj}_{\mathcal{Y}^{(k-1)}}\left\{\mathbf{x}^{*,k-1} - \alpha\mathbf{\Phi}^{(k-1)}(\mathbf{x}^{*,k-1})\right\}\Big\|_2 \\ &\le \|\widehat{\mathbf{x}}^{(k-1)} - \alpha\mathbf{\Phi}_e^{(k-1)}(\widehat{\mathbf{x}}^{(k-1)}) - \mathbf{x}^{*,k-1} + \alpha\mathbf{\Phi}^{(k-1)}(\mathbf{x}^{*,k-1})\|_2,\end{aligned} \tag{11}$$

where the inequality follows by the nonexpansivity property of the projection operator. Observe that

$$\begin{aligned}&\left\|\mathbf{\Phi}_e^{(k-1)}(\widehat{\mathbf{x}}^{(k-1)}) - \mathbf{\Phi}^{(k-1)}(\widehat{\mathbf{x}}^{(k-1)})\right\|_2 \\ &\qquad \le \gamma\|\mathbf{M}^{\mathsf{T}}(\widehat{\mathbf{p}}_0^{(k-1)} - \widetilde{\mathbf{p}}_0(\widehat{\mathbf{x}}^{(k-1)}))\|_2 \le \gamma\|\mathbf{M}\|_2 e_0^{(k-1)}.\end{aligned} \tag{12}$$

We can now expand and bound the right-hand side of (11) as

$$\begin{aligned}\|\widehat{\mathbf{x}}^{(k-1)} - \alpha\mathbf{\Phi}_e^{(k-1)}(\widehat{\mathbf{x}}^{(k-1)}) - \mathbf{x}^{*,k-1} + \alpha\mathbf{\Phi}^{(k-1)}(\mathbf{x}^{*,k-1})\|_2 \le \\ \|\widehat{\mathbf{x}}^{(k-1)} - \alpha\mathbf{\Phi}^{(k-1)}(\widehat{\mathbf{x}}^{(k-1)}) - \mathbf{x}^{*,k-1} + \alpha\mathbf{\Phi}^{(k-1)}(\mathbf{x}^{*,k-1})\|_2 \\ + \alpha\gamma\|\mathbf{M}\|_2 e_0^{(k-1)},\end{aligned} \tag{13}$$

where we have used the triangle inequality. By using the results of Lemma 1 and the fact that $\mathbf{\Phi}$ is a gradient mapping, the first term in the right-hand side of (13) can be written as

$$\begin{aligned}\|\widehat{\mathbf{x}}^{(k-1)} - \alpha\mathbf{\Phi}^{(k-1)}(\widehat{\mathbf{x}}^{(k-1)}) - \mathbf{x}^{*,k-1} + \alpha\mathbf{\Phi}^{(k-1)}(\mathbf{x}^{*,k-1})\|_2 \le \\ \rho(\alpha)\|\widehat{\mathbf{x}}^{(k-1)} - \mathbf{x}^{*,k-1}\|_2.\end{aligned} \tag{14}$$

where we let $\rho(\alpha) := \max\{|1 - \alpha\,\eta|, |1 - \alpha\,L_\Phi|\}$ (see [24]).

By putting together (11), (13), and (14), we have that

$$\begin{aligned}\|\mathbf{x}^{(k)}-\mathbf{x}^{*,k-1}\|_2 &\le \rho(\alpha)\|\widehat{\mathbf{x}}^{(k-1)}-\mathbf{x}^{*,k-1}\|_2+\alpha\gamma\|\mathbf{M}\|_2 e_0^{(k-1)}\\ &\le \rho(\alpha)\|\mathbf{x}^{(k-1)}-\mathbf{x}^{*,k-1}\|_2+\rho(\alpha)\|\mathbf{x}^{(k-1)}-\widehat{\mathbf{x}}^{(k-1)}\|_2\\ &\quad+\alpha\gamma\|\mathbf{M}\|_2 e_0^{(k-1)}\\ &= \rho(\alpha)\|\mathbf{x}^{(k-1)}-\mathbf{x}^{*,k-1}\|_2+\rho(\alpha)e_x^{(k-1)}+\alpha\gamma\|\mathbf{M}\|_2 e_0^{(k-1)},\end{aligned} \tag{15}$$

where the equality follows from (8). Next, consider the distance between $\mathbf{x}^{(k)}$ and the optimizer of (P1)$^{(k)}$, i.e., $\|\mathbf{x}^{(k)}-\mathbf{x}^{*,k}\|_2$. By using the triangle inequality and (15), it follows that

$$\begin{aligned}\|\mathbf{x}^{(k)}-\mathbf{x}^{*,k}\|_2 &\le \|\mathbf{x}^{(k)}-\mathbf{x}^{*,k-1}\|_2+\sigma^{(k-1)}\\ &\le \rho(\alpha)\|\mathbf{x}^{(k-1)}-\mathbf{x}^{*,k-1}\|_2+\rho(\alpha)e_x^{(k-1)}\\ &\quad+\alpha\gamma\|\mathbf{M}\|_2 e_0^{(k-1)}+\sigma^{(k-1)}.\end{aligned} \tag{16}$$

If $\rho(\alpha)<1$, then (16) represents a contraction, and applying it recursively, we obtain (9). □

Theorem 1 has the following immediate corollary.

Corollary 1. *Under the conditions of Theorem 1, we have that*

$$\limsup_{k\to\infty}\|\mathbf{x}^{(k)}-\mathbf{x}^{*,k}\|_2 \le \frac{\Delta^{\max}}{1-\rho(\alpha)}, \tag{17}$$

where $\Delta^{\max}:=\rho(\alpha)e_x^{\max}+\alpha\gamma\|\mathbf{M}\|_2 e_0^{\max}+\sigma^{\max}$.

Theorem 1 provides an *a-posteriori bound* as it is formulated in terms of the actual realization of the errors at each iteration.

We conclude the section by stating a result from [23] establishing average tracking properties of the error-diffusion rule (6). To this end, we introduce some pertinent definitions and assumptions regarding buses j and phases ϕ at which error diffusion is performed. The *Voronoi cell* associated with a set $\mathcal{Y}\subseteq\mathbb{R}^2$ and a point $\mathbf{x}\in\mathcal{Y}$ is defined as $\mathcal{V}_{\mathcal{Y}}(\mathbf{x}):=\{\mathbf{y}\in\mathbb{R}^2:\ \|\mathbf{x}-\mathbf{y}\|\le\|\mathbf{x}'-\mathbf{y}\|,\forall\mathbf{x}'\in\mathcal{Y}\}$.

Assumption 3. Consider the collection of *bounded* Voronoi cells of $\mathcal{Y}_{j,\phi}^{(k)}$, $k=1,2,...$:

$$\left\{\mathcal{V}_{\mathcal{Y}_{j,\phi}^{(k)}}(\mathbf{x}):\ \mathbf{x}\in\mathcal{Y}_{j,\phi}^{(k)},\left|\mathcal{V}_{\mathcal{Y}_{j,\phi}^{(k)}}(\mathbf{x})\right|<\infty,k=1,2,\dots\right\}.$$

The sizes of these bounded Voronoi cells are *uniformly bounded*. □

Assumption 4. The collection $\{\mathsf{ch}\mathcal{Y}_{j,\phi}^{(k)},k=1,2,...\}$ is a collection of *polytopes* such that:
A4.i The sizes of the polytopes are *uniformly bounded*; and,
A4.ii The set of outgoing normals to the faces of the polytopes is *finite*. □

Theorem 2 (Theorem 2 in [23])**.** *Under Assumptions 3 and 4, for each resource located at (j,ϕ), there exists a finite constant $E_{j,\phi}$ such that $\|\mathbf{e}_{j,\phi}^{(k)}\|_2\le E_{j,\phi}$ for all k. Consequently,*

$$\left\|\frac{1}{k}\sum_{\ell=1}^{k}\mathbf{x}_{j,\phi}^{(\ell)}-\frac{1}{k}\sum_{\ell=1}^{k}\widetilde{\mathbf{x}}_{j,\phi}^{(\ell)}\right\|_2\le\frac{E_{j,\phi}}{k} \tag{18}$$

and $\|\mathbf{x}_{j,\phi}^{(k)}-\widetilde{\mathbf{x}}_{j,\phi}^{(k)}\|_2\le 2E_{j,\phi}$ for all k.

Notice that Theorem 2 can be used to upper bound the error $e_x^{(k)}$ defined in (8). In particular, assume that the measurement error is upper bounded as $\|\widetilde{\mathbf{x}}_{j,\phi}^{(k)}-\widehat{\mathbf{x}}_{j,\phi}^{(k)}\|_2\le e^{\text{meas}}$. Then, we have that $\|\mathbf{x}_{j,\phi}^{(k)}-\widehat{\mathbf{x}}_{j,\phi}^{(k)}\|_2\le\|\mathbf{x}_{j,\phi}^{(k)}-\widetilde{\mathbf{x}}_{j,\phi}^{(k)}\|_2+\|\widetilde{\mathbf{x}}_{j,\phi}^{(k)}-\widehat{\mathbf{x}}_{j,\phi}^{(k)}\|_2\le 2E_{j,\phi}+e^{\text{meas}}$.

Finally, notice that extending the framework to time-varying nonconvex costs and constraints is a subject of ongoing efforts.

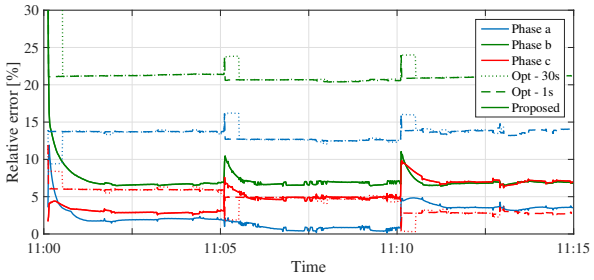

Fig. 2: Example of tracking error. Step changes in the setpoints $\mathbf{p}_0^{\text{set},k}$ occur at 11:05 and 11:10.

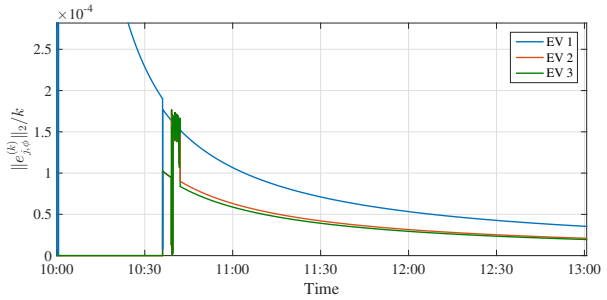

Fig. 3: Evolution of (18) for three representatives EVs.

5. ILLUSTRATIVE NUMERICAL RESULTS

Numerical tests are performed by using the IEEE 37-node test feeder [25]. Similar to e.g., [7, 8, 10, 12], we first convert all constant-current and constant-impedance loads in the IEEE data set into constant-power loads. We then create a power profile over the course of a day based on data provided by a utility in California. We populate the feeder with photovoltaic (PV) systems, energy storage systems, as well as EVs. Particularly, we consider 32 PV systems with a collective capacity of 960 kVA, 8 utility-scale batteries with an aggregate inverter nameplate capacity of 3600 kVA and state-of-charge of 4800 kWh, and 9 EVs with sizes 60, 80, and 130 kWh. Level-2 charging stations are presupposed, with discrete charging levels of 10, 20, 40, 60, 80, and 100% of the maximum charging capability of 7.2 kW.

For the algorithm illustrated in Fig. 1, the stepsize is set $\alpha=0.1$ and $\gamma=5$. The DER-related cost functions are set to $C_{j,\phi}^{(k)}=5(P_{\text{av},i,\phi}^{(k)}-P_{j,\phi}^{(k)})^2+(Q_{j,\phi}^{(k)})^2$ for PV systems (with $P_{\text{av},j,\phi}^{(k)}$ denoting the maximum real power available), $C_{j,\phi}^{(k)}=(P_{j,\phi}^{(k)})^2+(Q_{j,\phi}^{(k)})^2$, for the batteries, and $C_{j,\phi}^{(k)}=4(P_{j,\phi}^{(k)}-P_{\text{max},i,\phi})^2$ for the EVs, where $P_{\text{max},i,\phi}$ is the maximum charging rate. With this setting, the DER are incentivized to follow a given profile for the powers at the point of coupling, while minimizing the power curtailed from the PV systems and the deviation from a predetermined (dis)charging profilr for the batteries. A minimum charging rate is set for the EVs so that they can be fully charged at the time specified by the drivers.

FFig. 2 compares the tracking error when (P1)$^{(k)}$ is solved on an offline fashion (labeled as "Opt") every second and every 30 seconds, and the proposed method. For the same simulation setting, the proposed method achieves a better tracking accuracy. Fig. 3 corroborates the claims of Theorem 2 for three EVs (one EV arrives at the charging station at 10:00 and the other two arrive at 10:30).

6. REFERENCES

[1] J. A. Taylor, S. V. Dhople, and D. S. Callaway, "Power systems without fuel," *Renewable & Sustainable Energy Reviews*, vol. 57, pp. 1322–1336, May 2016.

[2] S. Paudyaly, C. A. Canizares, and K. Bhattacharya, "Three-phase distribution OPF in smart grids: Optimality versus computational burden," in *2nd IEEE PES Intl. Conf. and Exhibition on Innovative Smart Grid Technologies*, Manchester, UK, Dec. 2011.

[3] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter VAR control in distribution systems with high PV penetration," in *IEEE PES General Meeting*, San Diego, CA, Jul. 2012.

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

[5] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *Proc. of 50th IEEE Conf. on Decision and Control*, Orlando, FL, Dec. 2011.

[6] F. D. Brunner, H.-B Durr, and C. Ebenbauer, "Feedback design for multi-agent systems: A saddle point approach," in *Proc. of 51st IEEE Conf. on Decision and Control*, Maui, HI, Dec 2012, pp. 3783–3789.

[7] 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 set points. part I: Framework," *Electric Power Systems Research*, vol. 125, no. August, pp. 254–264, 2015.

[8] L. Gan and S. H. Low, "An online gradient algorithm for optimal power flow on radial networks," *IEEE J. on Selected Areas in Commun.*, vol. 34, no. 3, pp. 625–638, March 2016.

[9] E. Dall'Anese, S. V. Dhople, and G. B. Giannakis, "Photovoltaic inverter controller seeking ac optimal power flow solutions," *IEEE Trans. Power Syst.*, 2015, to appear. [Online] Available at: http://arxiv.org/abs/1501.00188.

[10] E. Dall'Anese and A. Simonetto, "Optimal power flow pursuit," *IEEE Trans. on Smart Grid*, May 2016.

[11] Y. Tang, K. Dvijotham, and S. Low, "Real-time optimal power flow," *IEEE Trans. on Smart Grid*, 2017.

[12] A. Hauswirth, S. Bolognani, G. Hug, and F. Dorfler, "Projected gradient descent on Riemannian manifolds with applications to online power system optimization," in *54th Annual Allerton Conference on Communication, Control, and Computing*, Sept 2016, pp. 225–232.

[13] A. Hauswirth, A. Zanardi, S. Bolognani, G. Hug, and F. Dorfler, "Online optimization in closed loop on the power flow manifold," in *12th IEEE PES PowerTech conference*, 2017.

[14] T. Basso, "IEEE 1547 and 2030 standards for distributed energy resources interconnection and interoperability with the electricity grid," National Renewable Energy Laboratory, 2014.

[15] H. Inose, Y. Yasuda, and J. Murakami, "A telemetering system by code modulation–Δ-Σ modulation," *IRE Transactions on Space Electronics and Telemetry*, , no. 3, pp. 204–209, 1962.

[16] R. W. Floyd and L. Steinberg, "An adaptive algorithm for spatial grey scale," in *Proc. Dig. SID International Symp.*, Los Angeles, California, 1975, pp. 36–37.

[17] D. Anastassiou, "Error diffusion coding for a/d conversion," *Circuits and Systems, IEEE Transactions on*, vol. 36, no. 9, pp. 1175–1186, Sep 1989.

[18] 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.

[19] A. Bernstein and E. Dall'Anese, "Linear power-flow models in multiphase distribution networks," in *The 7th IEEE Intl. Conf. on Innovative Smart Grid Technologies*, Sep. 2017.

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

[21] 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.

[22] A. Simonetto and G. Leus, "Distributed asynchronous time-varying constrained optimization," in *48th Asilomar Conference on Signals, Systems and Computers*, Nov 2014, pp. 2142–2146.

[23] A. Bernstein, N. Bouman, and J.-Y. Le Boudec, "Real-time minimization of average error in the presence of uncertainty and convexification of feasible sets," 2016, [Online] Available at: arXiv:1612.07287.

[24] E. K. Ryu and S. Boyd, "Primer on monotone operator methods," *Appl. Comput. Math.*, vol. 15, no. 1, pp. 3–43, Jan 2016.

[25] W. H. Kersting, "Radial distribution test feeders," in *IEEE Power Engineering Society Winter Meeting*, 2001, vol. 2, pp. 908–912.