Distributed Optimal Power Flow Algorithms over Time-Varying Communication Networks

Madi Zholbaryssov, Alejandro D. Domínguez-García, Senior Member, IEEE

Abstract—In this paper, we consider the problem of optimally coordinating the response of a group of distributed energy resources (DERs) in distribution systems by solving the so-called optimal power flow (OPF) problem. The OPF problem is concerned with determining an optimal operating point, at which some cost function, e.g., generation cost or power losses, is minimized, and operational constraints are satisfied. To solve the OPF problem, we propose distributed algorithms that are able to operate over time-varying communication networks and have geometric convergence rate. We solve the second-order cone program (SOCP) relaxation of the OPF problem for radial distribution systems, which is formulated using the so-called DistFlow model. Theoretical results are further supported by the numerical simulations.

I. INTRODUCTION

The growing integration of renewable generation inevitably raises concerns about the efficacy of the traditional centralized electricity paradigm in meeting global energy needs in a reliable and sustainable way. It is envisioned that present-day power grids, which are predominantly based on centralized energy sources, will eventually transition towards more decentralized architecture, where electric power is mainly produced from distributed energy resources (DERs). One of the obstacles in fulfilling this vision is the need for effective control strategies for coordinating a large population of DERs.

In order to adequately respond to rapid and large fluctuations in renewable generation, DERs will need to more frequently adjust their set-points, which will require real-time control systems to run and process data more often. To coordinate a large number of DERs, it will also be required to process a large volume of data in real-time. The traditional centralized approach operates by collecting the data in a central processing unit requiring a denser communication network with high-speed communication channels, which also needs to be secure to prevent cyber attackers from stealing sensitive private information. Hence, the centralized approach may not be feasible because of the resulting communication overhead. By contrast, a distributed approach processes the data locally, thereby dispensing with the need for moving the data to the central node. However, it is more difficult to solve the OPF problem in a distributed way, since communication latency and random data packet losses might prevent the distributed approach from converging to an optimal solution.

In this work, we consider the standard OPF problem for balanced distribution systems with high penetration of DERs, where each DER is operated within its capacity constraints and is endowed with a generation cost. The objective of the OPF problem is to determine an optimal operating point at which some cost function, e.g., total generation cost or power loss, is minimized, and operational constraints are satisfied. One of the main objectives of this work is to design distributed OPF solvers with a certain level of resiliency to communication latency and random data packet losses required to properly coordinate DERs over less reliable communication networks. Another important objective is to ensure that these algorithms have fast convergence rate in order to quickly update the set-points of DERs and provide a fast response to rapid and large transients in renewable generation.

A vast body of work has focused on solving the OPF problem for distribution systems. Earlier works (see, e.g., [1]–[3]) focused on dealing with the non-convexity of the OPF problem, and proposed semidefinite program (SDP) and SOCP relaxations, which were shown to be exact for radial networks under some conditions. A few works proposed distributed approaches for solving the OPF problem over time-invariant communication networks (see, e.g., [4]–[7]). In [7], the authors proposed a distributed algorithm for solving the SOCP relaxation of the OPF problem for balanced radial networks, which is based on the alternating direction method of multipliers (ADMM). The asynchronous ADMM was proposed in [8] to solve the OPF problem over time-varying communication networks. There exists another body of works (see, e.g., [9]–[19]) that focused on solving simplified OPF problems over time-varying communication networks, in which most operational constraints were neglected. One line of works (see, e.g., [9]–[16]) focused on the DER coordination problem with only total active power balance constraint and generation capacity constraints being considered. In addition to these constraints, another line of works (see, e.g., [17]–[19]) also considered line flow constraints.

Our starting point in the design of the algorithms is a primal-dual algorithm for solving the system of optimality conditions also known as the Lagrangian system. We then develop distributed versions of this primal-dual algorithm by having bus agents closely emulate the iterations of the primal-dual algorithm, where each agent maintains and updates only local variables. The resulting distributed primal-dual algorithms converge geometrically fast when operated over unreliable communication networks. Each algorithm can be viewed as a feedback interconnection of the primal-dual algorithm representing the nominal system and the disturbance generated due to the nature of the distributed implementation.
The key ingredient for establishing the convergence results is to show that both systems are finite-gain stable, which then allows us to use the small-gain theorem (see, e.g., [20]) and show the convergence of the feedback interconnected system. The small-gain-theorem-based analysis first appeared in [21] in the context of distributed algorithms for solving an unconstrained consensus optimization problem.

II. PRELIMINARIES

In this section, we formulate the OPF problem and outline the communication network models adopted in this work.

A. OPF Problem Formulation

We consider a balanced radial distribution system, the topology of which can be described by a directed graph $G_p = (V_p, E_p)$, where $V_p := \{1, 2, \ldots, n\}$ denotes the set of $n$ buses (nodes), and $E_p := \{e_1, e_2, \ldots, e_{n-1}\}$ denotes the set of electrical lines, with $e_k = (i, j) \in E_p$ if node $i$ is located upstream from node $j$, i.e., node $i$ is closer to the distribution substation, represented by node 1. Let $N_i^+ = \{j : (i, j) \in E_p\}$ and $N_i^- = \{l : (l, i) \in E_p\}$ denote the sets of (downstream) out-neighbors and (upstream) in-neighbors of node $i$, respectively. We define a node-to-edge incidence matrix, $M \in \mathbb{R}^{n \times |E_p|}$, with $M_{ik} = 1$ and $M_{jk} = -1$, if $e_k = (i, j) \in E_p$, and $M_{ik} = 0$ and $M_{jk} = 0$, otherwise. Also, let $M_0 \in \mathbb{R}^{n \times |E_p|}$ contain the entries of $M$, each corresponding to an upstream end, with $M_{0ik} = 1$ if $e_k = (i, j) \in E_p$ for some $j \in V_p$, and $M_{0ik} = 0$, otherwise. Similarly, $N_0 \in \mathbb{R}^{n \times |E_p|}$ contains the entries of $M$, each corresponding to a downstream end, with $N_{0ik} = -1$ if $e_k = (j, i) \in E_p$, for some $j \in V_p$, and $N_{0ik} = 0$, otherwise. [Note that $M = M_0 + N_0$.]

Let $r_{ij}$ and $x_{ij}$ denote the series resistance and reactance of line $(i, j)$, $R := \text{diag}(\{r_{ij}\}_{(i,j) \in E_p})$, and $X := \text{diag}(\{x_{ij}\}_{(i,j) \in E_p})$. Let $l_i^{(p)}$ and $g_i^{(q)}$ denote the active power demand and supply at node $i$, $p_{ij}$ and $q_{ij}$ denote the active and reactive power flow out of node $i$ through line $(i, j)$, i.e., $p_{ij} + q_{ij} \sqrt{-1}$ is the sending end. Let $V_i^e$ denote the voltage magnitude at node $i$, and $v_i := V_i^e$. Let $I_{ij}$ denote the current magnitude through line $(i, j) \in E_p$, and $I_{ij} := I_{ij}^2$. If $G_p$ is radial, the AC power flow equations can be exactly represented via the standard DistFlow model (see, e.g., [2], [22]):

\[
\begin{align*}
0 = g_i^{(p)} - l_i^{(p)} - M_p + N_0 R, & \quad \text{(1a)} \\
0 = g_i^{(q)} - l_i^{(q)} - M_q + N_0 X, & \quad \text{(1b)} \\
0 = M^T v - 2 R p + 2 X q + (R^2 + X^2) \ell, & \quad \text{(1c)} \\
0 = p \circ p + q \circ q - M_0^T v \circ \ell, & \quad \text{(1d)}
\end{align*}
\]

where $\circ$ denotes an element-wise multiplication, $g_i^{(p)} = [g_i^{(p)}, \ldots, g_i^{(p)}]^T$, $g_i^{(q)} = [g_i^{(q)}, \ldots, g_i^{(q)}]^T$, $v = [v_1, \ldots, v_n]^T$, $p = \{p_{ij}\}_{(i,j) \in E_p}$, $q = \{q_{ij}\}_{(i,j) \in E_p}$, and $\ell = \{\ell_{ij}\}_{(i,j) \in E_p}$.

Next, we specify the operational constraints to be satisfied in the problem formulation. We impose the following operating limits on the power outputs of the DERs:

\[
g_i^{(p)\min} \leq g_i^{(p)} \leq g_i^{(p)\max}, \quad (2a)
\]

For an inverter-interfaced DER at node $i$, the maximum amount of active and reactive power that can be produced is limited by the apparent power capability, $s_i^\max$, of an inverter:

\[
(g_i^{(p)})^2 + (g_i^{(q)})^2 \leq (s_i^\max)^2. \quad (3)
\]

To simplify the exposition, we will not consider (3) in the problem formulation; however, the proposed algorithms can handle the constraint (3).

Also, the voltage levels and line currents need to be within the following operating limits:

\[
v_{\min} \leq v \leq v_{\max}, \quad (4)
\]

\[
0 \leq \ell \leq \ell_{\max}. \quad (5)
\]

Then, the OPF problem can be formulated as follows:

\[
\text{OPF} : \min f(g_i^{(p)}) := \sum_{i=1}^n f_i(g_i^{(p)})
\]

\[
\text{over } g_i^{(p)}, g_i^{(q)}, p, q, v, \ell
\]

\[
\text{subject to } (1), (2), (4), (5)
\]

where $f_i(\cdot)$ denotes the cost function associated with the electric power generated by the DER at bus $i$. We make the following assumption regarding the cost function.

**Assumption 1.** Each cost function $f_i(\cdot)$ is twice differentiable and strongly convex with parameter $m > 0$, i.e., $f_i''(x) \geq m$, $\forall x \in [g_i^{(p)\min}, g_i^{(p)\max}]$, $\forall i \in V_p$.

B. SOCP Relaxation Of The OPF Problem

Because of the nonlinear equality constraint (1d), the OPF problem (6) is non-convex. For radial networks, it has been shown in [2], [3] that under certain assumptions when (1d) is relaxed to the second-order cone constraint, namely,

\[
p \circ p + q \circ q \leq M_0^T v \circ \ell, \quad (7)
\]

the OPF problem (6) admits an exact second-order cone program (SOCP) relaxation given below:

\[
\text{SOCP} : \text{minimize } f(g_i^{(p)})
\]

\[
\text{over } g_i^{(p)}, g_i^{(q)}, p, q, v, \ell
\]

\[
\text{subject to } (1a)–(1c), (2), (4), (5), (7)
\]

For our further analysis, we introduce additional variable $\epsilon = \{\epsilon_{ij}\}_{(i,j) \in E_p}$, and break the constraint (1c) into the following equivalent constraints:

\[
0 = \epsilon - 2 R p + 2 X q + (R^2 + X^2) \ell, \quad (9a)
\]

\[
0 = \epsilon - M^T v. \quad (9b)
\]

The proposed algorithms are designed to solve a regularized approximation of the OPF problem (8), where we add a regularization term to the objective function. The purpose of the problem regularization is to allow us to establish the convergence results. However, if the regularization term is small, there is practically no difference between the solutions.
of \((8)\) and its regularized approximation, which we provide below:

\[
\text{rSOCP : minimize } f(g^{(p)}) + \rho \| \|v\|_2^2 + \rho \|v - 1\|_2^2 \\
\text{over } g^{(p)}, g^{(p)}, p, q, v, \ell, \varepsilon
\]  

subject to (1a–1b, 2, 3, 4, 5, 7, 8).

where \(\| \cdot \|_2\) is the Euclidean norm, and \(\rho \| \|v\|_2^2 + \rho \|v - 1\|_2^2\) is the regularization term that also allows us to penalize the line currents and the deviation of the bus voltages from the nominal voltage, 1 pu. To this end, we develop a distributed algorithm that solves rSOCP for radial distribution systems.

C. Communication Network Models

Next, we introduce the model describing the communication network that enables the information exchange between nodes of the distribution system. We assume that the topology of the nominal communication network coincides with the topology of the power network. We consider (i) bidirectional and (ii) unidirectional communication models.

1) Bidirectional Communication Model: Let \(G_0 = (V, E, 0)\) denote an undirected graph, where \(E_0\) is the set of all available bidirectional communication links, with \((i, j) \in E_0\) if \((i, j)\) or \((j, i) \in E_p\). During time period \((t_k, t_{k+1})\), successful data transmissions among nodes can be captured by the undirected graph \(G^{(c)}: (V, E^{(c)})\), where \(E^{(c)} \subseteq E_0\) is the set of active communication links, with \((i, j) \in E^{(c)}\) if nodes \(i\) and \(j\) receive information from each other during time period \((t_k, t_{k+1})\). We make the following standard assumption regarding the connectivity of the network (see, e.g., [23]).

Assumption 2. \(\{i, j\} \in \bigcup_{k=0}^{B-1} E^{(c)}\), \(\forall (i, j) \in E_p\), for some positive integer \(B\).

Assumption 2 requires that a communication link \((i, j)\) is active at least once every \(B\) iterations. Note that communication graph \(G^{(c)}\) is not necessarily connected at any given time instant \(k\).

2) Unidirectional Communication Model: Let \(G_0 = (V, E, 0)\) denote a directed graph, where \(E_0\) is the set of all available unidirectional communication links, with \((i, j) \in E_0\) if \((i, j)\) or \((j, i) \in E_p\). During time period \((t_k, t_{k+1})\), successful data transmissions among nodes can be captured by the directed graph \(G^{(c)}: (V, E^{(c)})\), where \(E^{(c)} \subseteq E_0\) is the set of active communication links, with \((i, j) \in E^{(c)}\) if node \(j\) receives information from node \(i\) during time period \((t_k, t_{k+1})\). We make the following assumption regarding the connectivity of the network.

Assumption 3. \(\{i, j\} \in \bigcup_{k=0}^{B-1} E^{(c)}\) and \((j, i) \in \bigcup_{k=0}^{B-1} E^{(c)}\), \(\forall (i, j) \in E_p\), for some positive integer \(B\).

Unlike the bidirectional communication network, the unidirectional communication network does not necessarily allow nodes to exchange information simultaneously, within the same time period \((t_k, t_{k+1})\), if a communication link between them is active.

D. Simplified OPF

Many of the basic ideas behind the proposed distributed algorithms can be more effectively presented by considering a simplified OPF problem, referred to as the security-constrained economic dispatch (SCED) problem given below:

\[
\text{SCED : minimize } f(g^{(p)}) \\
\text{over } g^{(p)}, p
\]

subject to (11a–11b, 11c, 11d, 11e).

where \(p_{\text{max}}\) denotes the vector of the line capacities, and \(p_{\text{min}} = -p_{\text{max}}\). Then, the basic ideas can be directly carried over to design distributed algorithms that solve rSOCP. In the remainder, we first present the distributed primal-dual algorithms that solve SCED. Then, we extend these algorithms to solve rSOCP.

E. The Small-Gain Theorem

In the following, we give a brief overview of the main analysis tool used in later developments—the small-gain theorem (see, e.g., [20] Theorem 5.6) for discrete-time systems. For the forthcoming developments, we adopt the appropriate metric for measuring energy content of the signals of interest. For a given sequence of iterates, \(\{x[k]\}_\infty^{k=0}\), where \(x[k] \in \mathbb{R}^n\), consider the following norm (previously used in [21]):

\[
\|x\|^2_{\alpha, K} := \max_{0 \leq k \leq K} \alpha^{-k}\|x[k]\|_2^2
\]

for some \(\alpha \in (0, 1)\), where \(\| \cdot \|_2\) is the Euclidean norm. If \(\|x\|^2_{\alpha, K}\) is bounded for all \(K \geq 0\), then, \(\alpha^{-k}\|x[k]\|_2^2\) is bounded for all \(k \geq 0\), and, thus, it follows that \(x[k]\) converges to zero at a geometric rate \(O(\alpha^k)\).

Now, consider a feedback connection of two discrete-time systems \(H_1\) and \(H_2\) such that

\[
e_{2}^[k+1] = H_1(e_1[k]),
\]

\[
e_2[k+1] = H_2(e_2[k]).
\]

We assume that \(H_1\) and \(H_2\) are finite-gain stable in the sense of the norm \(\| \cdot \|^2_{\alpha, K}\), namely, the following relations hold:

\[
\|e_2\|^2_{\alpha, K} \leq \gamma_1 \|e_1\|^2_{\alpha, K} + \beta_1,
\]

\[
\|e_1\|^2_{\alpha, K} \leq \gamma_2 \|e_2\|^2_{\alpha, K} + \beta_2,
\]

for some nonnegative constants \(\beta_1, \beta_2, \gamma_1,\) and \(\gamma_2\). From (12), we have that

\[
\|e_2\|^2_{\alpha, K} \leq \gamma_1 \|e_1\|^2_{\alpha, K} + \beta_1
\]

\[
\leq \gamma_1 \gamma_2 \|e_2\|^2_{\alpha, K} + \gamma_1 \beta_2 + \beta_1,
\]

which by rearranging (13) yields

\[
\|e_2\|^2_{\alpha, K} \leq \frac{\gamma_1 \beta_2 + \beta_1}{1 - \gamma_1 \gamma_2}.
\]

Similarly,

\[
\|e_1\|^2_{\alpha, K} \leq \frac{\gamma_2 \beta_1 + \beta_2}{1 - \gamma_1 \gamma_2}.
\]
Then, if $\gamma_1 \gamma_2 < 1$, $\|e_1\|_2^{a,K}$ and $\|e_2\|_2^{a,K}$ are bounded, and $e_1[k]$ and $e_2[k]$ converge to zero at a geometric rate $O(a^k)$.

### III. Distributed SCED Over Time-Varying Communication Graphs

The purpose of this section is to present the key ideas behind the distributed primal-dual algorithm for solving rSOCP by considering a simpler problem, namely, SCED. We first consider the case of undirected communication graphs. Then, we tackle the general case of directed communication graphs.

#### A. Time-Varying Undirected Communication Graphs

Let $L(g^{(p)}, p, \lambda)$ denote the Lagrangian for SCED given by

$$L(g^{(p)}, p, \lambda) = f(g^{(p)}) + \lambda^T (g^{(p)} - l^{(p)} - Mp) + \frac{\rho}{2} \|g^{(p)} - l^{(p)} - Mp\|_2^2,$$

where $\rho > 0$ is a regularization parameter, $\lambda$ denotes the Lagrange multiplier associated with the power balance constraint \([11c]\). Our starting point for solving SCED is the following primal-dual algorithm \([24]\) Section 4.4) with additional projection:

\begin{align*}
g^{(p)}[k + 1] &= \left[g^{(p)}[k] - s \frac{\partial L[k]}{\partial g^{(p)}}\right]_{g^{(p)} \text{ max}}^{\frac{\partial L[k]}{\partial g^{(p)}}} \text{ min}, \tag{14a} \\
p[k + 1] &= \left[p[k] - s \frac{\partial L[k]}{\partial p}\right]_{p \text{ max}} \text{ min}, \tag{14b} \\
\lambda[k + 1] &= \lambda[k] + s \frac{\partial L[k]}{\partial \lambda}, \tag{14c}
\end{align*}

where $s > 0$ is a stepsize, and $[x^1, x^2]$ denotes the projection onto the box $[x_1, x_2]$, for $x_1, x_2 \in \mathbb{R}^n$. Let $\{g^{(p)}[i], p^*[i], \lambda^*[i]\}^T$ denote the equilibrium of \([14]\), where $g^{(p)}[i] := [g_1^{(p)}[i], \ldots, g_n^{(p)}[i]]^T$, $p^*[i] := [p_{i1}^*[i], \ldots, p_{ii}^*[i]]^T$, and $\lambda^*[i] := [\lambda_1^*[i], \ldots, \lambda_n^*[i]]^T$. In the distributed version of \([14]\), each node $i$ maintains the estimates of only local optimal quantities, namely, the local optimal power injection, $g^{(p)}[i]$, out-going flows, $p_{i1}^*[i]$ $(i, j) \in \mathcal{E}_p$, in-coming flows, $p_{i1}^*[i]$ $(l, i) \in \mathcal{E}_p$, and the Lagrange multiplier, $\lambda^*[i]$. Let $g^{(p)}[i]$, $p_{i1}^*[i]$, $\lambda^*[i]$ denote the estimates of $g^{(p)}[i]$, $p_{i1}^*[i]$, $\lambda^*[i]$ maintained at node $i$ at time instant $k$. Note that $p_{i1}^*[i]$, $(i, j) \in \mathcal{E}_p$, is estimated by nodes $i$ and $j$. Let

$$x^{(i)}[k] := [g^{(p)}[i], \{p_{i1}^*[i], (i,j) \in \mathcal{E}_p\}, \{p_{i1}^*[i], (i,l) \in \mathcal{E}_p\}, \lambda^*[i]]^T$$

denote the vector of the estimates of all local optimal primal and dual variables, denoted by

$$x^{(i)*} := [g^{(p)}[i], \{\hat{p}_{i1}^*[i], (i,j) \in \mathcal{E}_p\}, \{\hat{p}_{i1}^*[i], (i,l) \in \mathcal{E}_p\}, \lambda_i^*[i]]^T,$$

maintained at node $i$ at time instant $k$. We let node $i$ perform the updates based on its local Lagrangian given by

$$L^{(i)}(x^{(i)}) := f(\hat{b}_i^{(p)}) + \lambda_i^* \hat{b}_i^{(p)} + \frac{\rho}{2} \hat{b}_i^{(p), 2},$$

where

$$\hat{b}_i^{(p)} := g_i^{(p)} - l_i^{(p)} - \sum_{j \in \mathcal{N}_i^+} \hat{p}_{ij} + \sum_{l \in \mathcal{N}_i^-} \hat{p}_{il}.$$

Then, node $i$ updates $g^{(p)}[i]$ and $\lambda_i[k]$ as follows:

\begin{align*}
g^{(p)}[i][k + 1] &= \left[g^{(p)}[i][k] - s \frac{\partial L^{(i)}[k]}{\partial g^{(p)}}\right]_{g^{(p)} \text{ max}}^{\frac{\partial L^{(i)}[k]}{\partial g^{(p)}}} \text{ min}, \tag{15} \\
\lambda_i[k + 1] &= \lambda_i[k] + s \frac{\partial L^{(i)}[k]}{\partial \lambda_i}, \tag{16}
\end{align*}

where $L^{(i)}[k] := L^{(i)}(x^{(i)}[k])$. Next, we explain how node $i$ updates its local flow estimates, $\hat{p}_{ij}[i]$, $(i, j) \in \mathcal{E}_p$, and $\hat{p}_{ij}[i]$, $(i, l) \in \mathcal{E}_p$. Consider $(i, j) \in \mathcal{E}_p$, and note that $\hat{p}_{ij}[i]$ and $\hat{p}_{ij}[i]$ are the estimates of $p_{ij}^*[i]$ maintained by nodes $i$ and $j$, respectively. To ensure that the estimates $\hat{p}_{ij}[i]$ and $\hat{p}_{ij}[i]$ converge to the same value, nodes $i$ and $j$ need to exchange the estimates, compute the average, and use it in their updates as follows:

\begin{align*}
\hat{p}_{ij}[i][k + 1] &= \left(1 - a_{ij}[k]\right)\hat{p}_{ij}[i][k] + a_{ij}[k]\hat{p}_{ij}[j][k] \\
&- s\bar{y}_{ij}[k]\hat{p}_{ij}^{\text{ max}}, \tag{17} \\
\hat{p}_{ij}[j][k + 1] &= \left(1 - a_{ij}[k]\right)\hat{p}_{ij}[j][k] + a_{ij}[k]\hat{p}_{ij}[i][k] \\
&- s\bar{y}_{ij}[k]\hat{p}_{ij}^{\text{ max}}, \tag{18}
\end{align*}

where

$$a_{ij}[k] = \begin{cases} 0.5 & \text{if } (i, j) \in \mathcal{E}_c[k], \\ 0 & \text{otherwise}, \end{cases}$$

and $\bar{y}_{ij}[k]$ and $\bar{y}_{ij}[k]$ are the estimates of the gradient $\frac{\partial L}{\partial p_{ij}}$ maintained at nodes $i$ and $j$, respectively. One way to estimate the gradient can be purely based on the local Lagrangian (local information):

$$\bar{y}_{ij}[k] = \frac{\partial L^{(i)}[k]}{\partial p_{ij}}, \quad \bar{y}_{ij}[k] = \frac{\partial L^{(i)}[k]}{\partial p_{ij}}.$$

However, leveraging only local information, as in \([19]\), results in slow (asymptotic) convergence (to be demonstrated numerically in Section III-D). A better approach is to let each node track the gradient by also using the neighbor’s information:

\begin{align*}
\hat{y}_{ij}[k + 1] &= \left(1 - a_{ij}[k]\right)\hat{y}_{ij}[k] + a_{ij}[k]\bar{y}_{ij}[k] \\
&+ 2\left(\frac{\partial L^{(i)}[k + 1]}{\partial \hat{p}_{ij}} - \frac{\partial L^{(i)}[k]}{\partial \hat{p}_{ij}}\right), \tag{20a} \\
\bar{y}_{ij}[k + 1] &= \left(1 - a_{ij}[k]\right)\bar{y}_{ij}[k] + a_{ij}[k]\hat{y}_{ij}[k] \\
&+ 2\left(\frac{\partial L^{(i)}[k + 1]}{\partial \hat{p}_{ij}} - \frac{\partial L^{(i)}[k]}{\partial \hat{p}_{ij}}\right), \tag{20b}
\end{align*}

with

$$\hat{y}_{ij}[0] = 2\frac{\partial L^{(i)}[0]}{\partial \hat{p}_{ij}}, \quad \bar{y}_{ij}[0] = 2\frac{\partial L^{(i)}[0]}{\partial \hat{p}_{ij}},$$

where $\hat{y}_{ij}[k]$ and $\bar{y}_{ij}[k]$ are updated so that their average

$$\frac{1}{2} \left(\hat{y}_{ij}[k] + \bar{y}_{ij}[k]\right) = \frac{\partial L^{(i)}[k]}{\partial \hat{p}_{ij}} + \frac{\partial L^{(i)}[k]}{\partial \hat{p}_{ij}} = -\lambda_i[k] + \lambda_{ij}[k]$$
has exactly the same form as
\[ \frac{\partial L}{\partial p_{ij}} \bigg|_{g^{(p)}[k], p[k], \lambda[k]} = -\lambda_i[k] + \lambda_j[k], \]
used in algorithm (4). This idea of tracking the gradient, which appeared in [21] for solving an unconstrained multi-agent optimization problem, allows us to more closely emulate the updates in algorithm (14), and achieve faster (geometric) convergence rate.

Below, we provide the iterations run by node \(i\):
\begin{align}
g^{(p)}_i[k+1] &= g^{(p)}_i[k] - s \frac{\partial L^{(i)}[k]}{\partial p^{(i)}} \bigg|_{g^{(p)}_i[k]} \frac{\partial L^{(i)}[k+1]}{\partial p^{(i)}}, \quad (22a) \\
p^{(i)}_i[k+1] &= \left(1 - a_{ij}[k]\right)\hat{p}_{ij}[k] + a_{ij}[k]p^{(i)}_i[k] - s\hat{g}_{ij}[k] \frac{\partial g^{(i)}_{ij}[k]}{\partial p^{(i)}}, \quad (22b) \\
\hat{y}_{ij}[k+1] &= \left(1 - a_{ij}[k]\right)\hat{y}_{ij}[k] + a_{ij}[k]\hat{y}_{ij}[k] + 2 \left( \frac{\partial L^{(i)}[k+1]}{\partial \lambda_i[k]} - \frac{\partial L^{(i)}[k]}{\partial \lambda_i[k]} \right), \quad (22d) \\
\hat{\lambda}_i[k+1] &= \frac{\partial L^{(i)}[k+1]}{\partial \lambda_i[k]} \quad \text{(22f)}
\end{align}

\(B. \text{ Feedback Interconnection Representation of the Distributed Primal-Dual Algorithm}\)

In the following, we represent (22) as a feedback interconnection of a nominal system, denoted by \(H_1\), and a disturbance system, denoted by \(H_2\), which allows us to utilize the small-gain theorem for convergence analysis purposes. To this end, let
\[
g^{(p)}[k] := [g^{(p)}_1[k], \ldots, g^{(p)}_n[k]]^\top, \lambda[k] := [\lambda_1[k], \ldots, \lambda_n[k]]^\top, \\
p^{(i)}[k] := [\hat{p}_{ij}[k]]_{(i,j) \in E_p}, \hat{p}^{(i)}[k] := [\hat{p}_{ij}[k]]_{(i,j) \in E_p}, \\
y^{(i)}[k] := [\hat{y}_{ij}[k]]_{(i,j) \in E_p}, \hat{y}^{(i)}[k] := [\hat{y}_{ij}[k]]_{(i,j) \in E_p}, \\
p[k] := \frac{1}{2} (\hat{p}[k] + \bar{p}[k]), \bar{p}[k] := \frac{1}{2} (\hat{p}[k] + \bar{p}[k]).
\]

By using (22), we write the iterations for \(g^{(p)}[k], \bar{p}[k], \text{ and } \lambda[k]\), which constitute the nominal system, \(H_1\), given by:
\[
H_1 : g^{(p)}[k+1] = g^{(p)}[k] - s \frac{\partial L^{(i)}[k+1]}{\partial g^{(p)}}, \quad (23a) \\
\bar{p}[k+1] = \frac{1}{2} \left[ \bar{p}[k] - s \frac{\partial L^{(i)}[k]}{\partial p} + e_p[k] \right] \frac{\partial L^{(i)}[k]}{\partial p}, \quad (23b) \\
\lambda[k+1] = \lambda[k] + s \frac{\partial L^{(i)}[k]}{\partial \lambda} + e_\lambda[k], \quad (23c)
\]

where
\[
L^{(i)}[k] := L(g^{(p)}[k], \bar{p}[k], \lambda[k]), \\
e_p[k] := -s\hat{p}(M\bar{p}[k] - M_0\bar{p}[k] - N_0\hat{p}[k]), \\
e_\lambda[k] := s(M\bar{p}[k] - M_0\bar{p}[k] - N_0\hat{p}[k]),
\]
to obtain (23b), we used the fact that
\[
\frac{\partial L^{(i)}[k]}{\partial p} = \bar{p}[k].
\]

We note that \(e[k] := [e_p[k]^\top, e_\lambda[k]^\top, e_\lambda[k]^\top]^\top\) results from \((\hat{p}[k], \bar{p}[k])\) and \((\hat{y}[k], \bar{y}[k])\) deviating from their respective average, \(\bar{p}[k]\) and \(\bar{y}[k]\); without \(e[k]\), the nominal system \(H_1\) has exactly the same form as (14). Now, we define the disturbance system, \(H_2\), as follows:
\[
H_2 : \hat{y}_{ij}[k+1] = \left(1 - a_{ij}[k]\right)\hat{y}_{ij}[k] + a_{ij}[k]\hat{y}_{ij}[k] + 2 \left( \frac{\partial L^{(i)}[k+1]}{\partial p_{ij}}, \hat{y}_{ij}[k] + a_{ij}[k]\hat{y}_{ij}[k] + 2 \left( \frac{\partial L^{(i)}[k+1]}{\partial p_{ij}} \right), \quad (24c)
\]
\[
\hat{\omega}[k+1] = \frac{1}{2} (\hat{p}[k] + \bar{p}[k]), \bar{p}[k] := \frac{1}{2} (\hat{p}[k] + \bar{p}[k]).
\]

Then, as illustrated in Fig. 1, algorithm (22) can be viewed as a feedback interconnection of \(H_1\) and \(H_2\). In the following, we find the relationship between the loop gain of the feedback system and the stepsize \(s\) establishing that the loop gain can be reduced by decreasing \(s\). This enables us to apply the small-gain theorem and show that the feedback loop does not amplify the energy of the convergence error, but, on the contrary, the error eventually decays to zero, if the loop gain is sufficiently small.

\(C. \text{ Convergence Analysis}\)

We first establish that the systems \(H_1\) and \(H_2\) are finite-gain stable:
\[
R_1. \|z\|_2^\alpha \leq \alpha_1 \|e\|_2^\alpha + \beta_1 \text{ for some positive } \alpha_1 \text{ and } \beta_1, \\
R_2. \|e\|_2^\alpha \leq \alpha_2 \|z\|_2^\alpha + \beta_2 \text{ for some positive } \alpha_2 \text{ and } \beta_2,
\]
algorithm (22), we have that
\[ \|z\|_2^{a,K} \leq \beta, \] (28)
for some \( a \in (0, 1) \), and sufficiently small \( s > 0 \). In particular, \( x^{(i)}[k] \) converges to \( x^{(i)*} \), \( \forall i \), at a geometric rate \( O(a^k) \).

Finally, we show that \((g^{(p)*}, p^*)\) is the solution of SCED.

**Lemma 1.** Consider \((g^{(p)*}, p^*, \lambda^*)\), namely, the equilibrium of (14). Then, \((g^{(p)*}, p^*)\) is the solution of SCED.

**Proof.** At the equilibrium, we have that
\[
\begin{align*}
g^{(p)*} &= \left[g^{(p)*} - s\nabla f(g^{(p)*}) - s\lambda^* \right]g^{(p)\text{max}}, \\
p^* &= [p^* + sM^T \lambda^*]p^{\text{max}}, \\
\lambda^* &= \lambda^* + s(g^{(p)*} - l(p) - M p^*). 
\end{align*}
\]

Then, the following relations hold:
\[
\begin{align*}
0 &= \nabla f(g^{(p)*}) + \lambda^* + \mu^* - \nu^*, \quad \text{(29a)} \\
0 &= M^T \lambda^* - \alpha^* + \beta^*, \quad \text{(29b)} \\
0 &= g^{(p)*} - l(p) - M p^*, \quad \text{(29c)} \\
0 &= \mu^* (g^{(p)*} - g^i), \quad i = 1, \ldots, n, \quad \text{(29d)} \\
0 &= \alpha^*_i (p^i_{ij} - p^*_{ij}), \quad i = 1, \ldots, n, \quad \text{(29f)} \\
0 &= \beta^*_i (p^i_{ij} - p^*_{ij}), \quad (i, j) \in E_p, \quad \text{(29g)}
\end{align*}
\]

for some non-negative \( \mu^*, \nu^*, \alpha^*_i, \beta^*_i \), \( i = 1, \ldots, n, \alpha^*_i, \beta^*_i \), \( (i, j) \in E_p \), where \( \mu^* = [\mu^*_1, \ldots, \mu^*_n]^T \), \( \nu^* = [\nu^*_1, \ldots, \nu^*_n]^T \), \( \alpha^* = \{\alpha^*_i\}_{i,j} \in E_p \), and \( \beta^* = \{\beta^*_i\}_{i,j} \in E_p \). Noticing that (29) represents the Karush-Kuhn-Tucker (KKT) conditions for SCED, it follows from (24) Proposition 3.3.1 that \((g^{(p)*}, p^*)\) is the solution of SCED.

**D. Numerical Simulations**

Next, we present numerical results to illustrate the performance of the distributed primal-dual algorithm (22) over undirected graph \( G^{(i)}[k] \) using the IEEE 69-bus radial test system (28).

A subset of buses are designated to have a DER. For a DER at bus \( i \), we choose \( f_i(p_i) = a_i p_i^2 \), where \( a_i > 0 \) is randomly selected. It is assumed that each communication link becomes inactive with probability 0.4. The algorithm uses a constant stepsize \( s = 2 \times 10^{-2} \) and \( \rho = 2 \). For initialization, we use \( g^{(p)}[0] = l_p[0], \hat{p}_{ij}[0] = 0, \bar{p}_{ij}[0] = 0, (i, j) \in E_p \), and
\[
\hat{y}_{ij}[0] = 2 \frac{\partial L(i)}{\partial p_{ij}}, \quad \bar{y}_{ij}[0] = 2 \frac{\partial L(i)}{\partial \bar{p}_{ij}}.
\]

In the distributed implementation, communicating data takes much longer than one iteration executed by a computing device. Rather than the total number of iterations, the number of communication attempts can serve as a more appropriate performance metric to evaluate the practical usefulness of the algorithm. We believe that it is reasonable to assume that a computing device is able to perform a number of iterations
(less than 100) between consecutive communication attempts. Let \( m \) denote the number of iterations between consecutive communication attempts. In the numerical example, we used different values of \( m \). We note that making \( m \) large or even finding a minimum of the local Lagrangians, \( L^{(i)}(x^{(i)}) \), \( i \in \mathcal{V}_p \), does not necessarily make the performance better. On the contrary, keeping \( m \) relatively small (\( m < 20 \)) often achieves a much better performance.

In Fig. 2 we compare the performance of algorithm \( (22) \), for convenience referred to as \( A_1 \), where we recall that

\[
\hat{y}_{ij}[k + 1] = (1 - a_{ij}[k])\hat{y}_{ij}[k] + a_{ij}[k]y_{ij}[k] + 2\left( \frac{\partial L^{(i)}[k + 1]}{\partial \hat{p}_{ij}} - \frac{\partial L^{(i)}[k]}{\partial \hat{p}_{ij}} \right), \quad (30a)
\]

\[
\hat{y}_{ij}[k + 1] = (1 - a_{ij}[k])\hat{y}_{ij}[k] + a_{ij}[k]y_{ij}[k] + 2\left( \frac{\partial L^{(j)}[k + 1]}{\partial \hat{p}_{ij}} - \frac{\partial L^{(j)}[k]}{\partial \hat{p}_{ij}} \right), \quad (30b)
\]

against that of algorithm \( (15)–(19) \), referred to as \( A_2 \), where

\[
\hat{y}_{ij}[k] = \frac{\partial L^{(i)}[k]}{\partial \hat{p}_{ij}}, \quad \hat{y}_{ij}[k] = \frac{\partial L^{(j)}[k]}{\partial \hat{p}_{ij}}. \quad (31)
\]

We demonstrate that updating \( \hat{y}_{ij}[k] \) and \( \hat{y}_{ij}[k] \) using \( (30) \) results in a much better performance than using \( (31) \). Figure 2 shows the trajectory of the relative cost error of the obtained solutions, namely,

\[
\frac{|f(g^{(p)}[k]) - f(g^{(p)*})|}{f(g^{(p)*})},
\]

and the evolution of the largest constraint violation. In both algorithms, each node runs \( m = 5 \) iterations between consecutive communication attempts. The results in Fig. 2 demonstrate that \( A_1 \) has geometric convergence rate, and converges significantly faster than \( A_2 \). We note that \( A_2 \) has asymptotic convergence rate and requires stepsize \( s \) to go to zero asymptotically. In the simulations, \( A_2 \) uses \( s[k] = a/(k + b) \), with \( a = 1 \), and \( b = 10 \).

**E. Time-Varying Directed Communication Graphs**

When \( G^{(c)}[k] \) is directed, the iterations in \( (22) \) fail to converge. In the following, we provide the robust extension of \( (22) \) that solves SCED over time-varying directed graphs.

In the robust extension, we let node \( i \) perform the same updates for \( g^{(p)}_i \) and \( \lambda_i \) as the ones in \( (22) \), namely,

\[
g^{(p)}_i[k + 1] = \left[ g^{(p)}_i[k] - s \frac{\partial L^{(i)}[k]}{\partial g^{(p)}_i} \right] g^{(p)\text{max}}, \quad (32a)
\]

\[
\lambda_i[k + 1] = \lambda_i[k] + s \frac{\partial L^{(i)}[k]}{\partial \lambda_i}, \quad (32b)
\]

The only difference between \( (22) \) and its robust extension is how the averaging step is performed in the updates of the flow and gradient estimates. The key idea behind the averaging step in the robust extension is to let the neighboring nodes perform averaging exactly once over possibly longer time periods. In other words, we ensure that, for any given

\[ (i, j) \in \mathcal{E}_p, \] there exists a sequence of time instants \( \{T_k\}_k=0^{\infty} \) such that nodes \( i \) and \( j \) perform averaging (using the values they received from each other) exactly once during each time interval \( (T_k, T_{k+1}) \), \( k = 0, 1, \ldots \). One of the simplest ways to implement such strategy over directed \( G^{(c)}[k] \) is to let nodes \( i \) and \( j \) perform averaging in an alternating fashion. In other words, once node \( i \) performs averaging at time instant \( t_k \) using the value received from node \( j \), it waits for node \( j \) to perform averaging (before node \( i \) performs averaging again). While in this waiting mode, node \( i \) still performs a local update and, possibly, averaging with other neighbors. To implement this strategy over directed \( G^{(c)}[k] \), nodes \( i \) and \( j \) need to maintain and communicate certain acknowledgement flags. By sending a flag, a node intends to let its neighbor know whether or not it has performed averaging; then, based on this information, the neighbor decides whether or not it should perform averaging. Below, we provide more details of this approach.

Suppose node \( i \) receives \( \hat{p}_{ij}[k] \) and the flag from node \( j \) at time instant \( t_k \). If the status of the received flag is different from the previously received one, then, node \( i \) flips the status of its own flag, stores \( \hat{p}_{ij}[k] \) and \( \hat{p}_{ij}[k] \), and performs averaging.
as follows:
\[
\hat{p}_{ij}[k+1] = \left[ (1 - a_{ij}[k])\hat{p}_{ij}[k] + a_{ij}[k]\hat{p}_{ij}[k] \right.
- s\hat{y}_{ij}[k] \right] \hat{p}_{ij}^{\max},
\] (33)

Over the next iterations, it keeps sending \( \hat{p}_{ij}[k] \), \( \bar{p}_{ij}[k] \) and its flag to node \( j \), until node \( i \) receives a different flag from node \( j \). Meanwhile, if node \( j \) receives a different flag (from the previously received one) from node \( i \) (which means that node \( i \) has performed averaging) at some time instant \( t_s > t_k \), node \( j \) flips the status of its own flag and performs averaging, but slightly differently, as follows:
\[
\hat{p}_{ij}[\tau + 1] = \left[ (1 - a_{ij}[k])\hat{p}_{ij}[k] + a_{ij}[k]\hat{p}_{ij}[k] \right.
+ (\hat{p}_{ij}[\tau] - \hat{p}_{ij}[k]) - s\hat{y}_{ij}[\tau] \right] \hat{p}_{ij}^{\max},
\] (34)

Note that, in the averaging step, node \( j \) uses the same values that node \( i \) used at time instant \( t_k \). In (34), we also have \( (\hat{p}_{ij}[\tau] - \hat{p}_{ij}[k]) \), which is the sum of all gradient terms, \( s\hat{y}_{ij}[t], t = k, k+1, \ldots, \tau \), that have been accumulated since time instant \( t_k \). In this way, we are able to mimic the corresponding iteration in algorithm (22). Also, note that if nodes \( i \) and \( j \) happen to perform averaging within the same time period \( (t_k, t_{k+1}) \), then, \( \tau = k \), and (34) exactly mimics the corresponding iteration in algorithm (22). This scheme, where nodes perform averaging in alternating fashion, is referred to as the alternating averaging protocol (see, e.g., (25), (26)). We now formally define the updates of the flow and gradient estimates as follows. For each \((i, j) \in E_p\), node \( i \) runs the following iterations:
\[
\begin{align*}
\hat{p}_{ij}[k+1] &= \left[ (1 - a_{ij}[k])\hat{p}_{ij}[k] + a_{ij}[k]\hat{p}_{ij}[k] \right.
- s\hat{y}_{ij}[k] \right] \hat{p}_{ij}^{\max},
\end{align*}
\] (35a)
\[
\begin{align*}
\hat{y}_{ij}[k+1] &= \left[ (1 - a_{ij}[k])\hat{y}_{ij}[k] + a_{ij}[k]\hat{y}_{ij}[k] \right.
+ 2\left( \frac{\partial L^{(i)}[k+1]}{\partial \hat{w}_{ij}} - \frac{\partial L^{(j)}[k]}{\partial \hat{w}_{ij}} \right),
\end{align*}
\] (35b)

while node \( j \) executes the iterations given below:
\[
\begin{align*}
\hat{p}_{ij}[k+1] &= \left[ (1 - a_{ij}[k])\hat{r}_{ij}[k] + a_{ij}[k]\hat{r}_{ij}[k] + \hat{p}_{ij}[k] - \hat{r}_{ij}[k] - s\hat{y}_{ij}[k] \right] \hat{p}_{ij}^{\max},
\end{align*}
\] (36a)
\[
\begin{align*}
\hat{y}_{ij}[k+1] &= \left[ (1 - a_{ij}[k])\hat{y}_{ij}[k] + a_{ij}[k]\hat{y}_{ij}[k] + \hat{y}_{ij}[k] - \hat{p}_{ij}[k] \right.
+ 2\left( \frac{\partial L^{(j)}[k+1]}{\partial \hat{w}_{ij}} - \frac{\partial L^{(i)}[k]}{\partial \hat{w}_{ij}} \right),
\end{align*}
\] (36b)

where \( a_{ij}[k], a_{ji}[k], \hat{r}_{ij}[k], \hat{r}_{ij}[k], \hat{p}_{ij}[k], \) and \( \hat{p}_{ij}[k] \) are updated using the alternating averaging protocol (see, e.g., (25), (26)):

node \( i \):
\[
\hat{\phi}_{ij}[k] = \begin{cases} 
\phi_{ij}[k] & \text{if } (j, i) \in E_c[k], \\
\phi_{ij}[k-1] & \text{otherwise},
\end{cases}
\]

\[
\begin{align*}
\phi_{ij}[k] &= \begin{cases} 
-\phi_{ij}[k-1] & \text{if } \hat{\phi}_{ij}[k] \neq \phi_{ij}[k-1], \\
\phi_{ij}[k-1] & \text{otherwise},
\end{cases}
\end{align*}
\]

\[
\begin{align*}
a_{ij}[k] &= \begin{cases} 
0.5 & \text{if } (j, i) \in E_c[k], \hat{\phi}_{ij}[k] \neq \phi_{ij}[k-1], \\
0 & \text{otherwise},
\end{cases}
\end{align*}
\]

node \( j \):
\[
\begin{align*}
\hat{\phi}_{ij}[k] &= \begin{cases} 
\phi_{ij}[k] & \text{if } (i, j) \in E_c[k], \\
\phi_{ij}[k-1] & \text{otherwise},
\end{cases}
\end{align*}
\]

\[
\begin{align*}
\phi_{ij}[k] &= \begin{cases} 
-\phi_{ij}[k-1] & \text{if } \hat{\phi}_{ij}[k] \neq \phi_{ij}[k-1], \\
\phi_{ij}[k-1] & \text{otherwise},
\end{cases}
\end{align*}
\]

\[
\begin{align*}
a_{ij}[k] &= \begin{cases} 
0.5 & \text{if } (i, j) \in E_c[k], \hat{\phi}_{ij}[k] \neq \phi_{ij}[k-1], \\
0 & \text{otherwise},
\end{cases}
\end{align*}
\]

\[
\begin{align*}
\hat{\rho}_{ij}[k] &= \begin{cases} 
\hat{\rho}_{ij}[k] & \text{if } (i, j) \in E_c[k], \hat{\phi}_{ij}[k] \neq \phi_{ij}[k-1], \\
\rho_{ij}[k] & \text{otherwise},
\end{cases}
\end{align*}
\]

\[
\begin{align*}
\rho_{ij}[k] &= \begin{cases} 
\hat{\rho}_{ij}[k] & \text{if } (i, j) \in E_c[k], \hat{\phi}_{ij}[k] \neq \phi_{ij}[k-1], \\
\rho_{ij}[k] & \text{otherwise},
\end{cases}
\end{align*}
\]

\[
\begin{align*}
\hat{\rho}_{ij}[k] &= \begin{cases} 
\hat{\rho}_{ij}[k] & \text{if } (i, j) \in E_c[k], \hat{\phi}_{ij}[k] \neq \phi_{ij}[k-1], \\
\rho_{ij}[k] & \text{otherwise},
\end{cases}
\end{align*}
\]

where \( \neg \) denotes the logical negation, i.e., \( \neg \xi = 1 \) if \( \xi = 0 \), and \( \neg \xi = 0 \), otherwise; \( t_k \leq k \) denotes the latest time, when node \( i \) performed averaging. In protocol (37), \( \hat{\phi}_{ij} \) and \( \phi_{ij} \) are the acknowledgement flags maintained by nodes \( i \) and \( j \), respectively. Nodes \( i \) and \( j \) store the received statuses of the flags \( \hat{\phi}_{ij} \) and \( \phi_{ij} \) in \( \hat{\phi}_{ij} \) and \( \phi_{ij} \), respectively.

Initially, \( \hat{\phi}_{ij}[0] = 1, \phi_{ij}[0] = 0, \hat{\phi}_{ij}[0] = 0, \) and \( \phi_{ij}[0] = 0 \). The reason for setting the node \( j \)'s flag, \( \hat{\phi}_{ij}[0] \), to 1 is to initiate the protocol execution. [If both flags, \( \hat{\phi}_{ij} \) and \( \phi_{ij} \), are set to zero, the protocol will never execute.] Below, we state the convergence result for the robust primal-dual algorithm (32), (35), (36), and (37), omitting the proof since it is analogous to that of a similar result established in Section IV dealing with rSOPC.

**Proposition 4.** Let Assumptions 3 and 4 hold. Then, under algorithm (32), (35), (36), and (37), we have that
\[
\|z\|_{2}^{a,K} \leq \beta,
\] (38)
for some \( a \in (0, 1), \beta > 0, \) and sufficiently small \( s > 0 \). In particular, \( a^{(i)}[k] \) converges to \( a^{(i)*} \), \( \forall i \), at a geometric rate \( O(a^{k}) \).

**IV. DISTRIBUTED OPF OVER TIME-VARYING COMMUNICATION GRAPHS**

In this section, we expand on the ideas of Section III and present the distributed primal-dual algorithms for solving rSOPC over time-varying communication graphs. We first consider the case of undirected communication graphs. Then, we tackle the general case of directed communication graphs.

**A. Time-Varying Undirected Communication Graphs**

Let \( x := [g^{(p)}, g^{(q)}, v]^{T} \) and \( \omega := [p, q, e, \ell]^{T} \), and let \( \gamma := [\lambda, \mu, \eta]^{T} \) and \( \tau \) denote the dual variables associated with the
DistFlow model constraints (13)–(15), (7), and (9) in rSOCP. Let $L(x, \gamma, \tau)$ denote the augmented Lagrangian for rSOCP given by

$$L(x, \omega, \gamma, \tau) = f(g^{(p)}) + \lambda^T b^{(p)} + \mu^T b^{(q)} + \nu^T b^{(v)} + \eta^T (e - M^T v) + \rho \|e\|_2^2 + \rho \|v - 1\|_2^2 + \tau^T (p + q - \omega) + \rho_1 \|b^{(p)}\|_2^2 + \rho_2 \|b^{(q)}\|_2^2 + \rho_3 \|b^{(v)}\|_2^2,$$

where $b^{(p)}$, $b^{(q)}$ and $b^{(v)}$ are defined as follows:

- $b^{(v)} := \varepsilon - 2R^p - 2Xq + (R^2 + X^2) \ell,$
- $b^{(p)} := g^{(p)} - f^{(p)} - Mp + N_0 R\ell,$
- $b^{(q)} := g^{(q)} - q^{(q)} - Mq + N_0 X\ell.$

The regularization terms $\rho_1 \|b^{(p)}\|_2^2, \rho_2 \|b^{(q)}\|_2^2,$ and $\rho_3 \|b^{(v)}\|_2^2$ penalize the violation of the constraints and allow us to significantly improve the convergence speed.

Our starting point to solve rSOCP is the following primal-dual algorithm:

$$x[k + 1] = \mathcal{P}_x \left( x[k] - s \frac{\partial L[k]}{\partial x} \right),$$

$$\omega[k + 1] = \mathcal{P}_\Omega \left( \omega[k] - s \frac{\partial L[k]}{\partial \omega} \right),$$

$$\gamma[k + 1] = \gamma[k] + s \frac{\partial L[k]}{\partial \gamma},$$

$$\tau[k + 1] = \left[ \tau[k] + 2s \frac{\partial L[k]}{\partial \tau} \right]^{+},$$

where $L[k] := L(x[k], \omega[k], \gamma[k], \tau[k]), \mathcal{P}_x(\cdot)$ denotes the projection onto the set

$$\mathcal{X} := \left\{ x \in [w^{\text{min}}, w^{\text{max}}], w \in \{p^{(p)}, q^{(q)}, v\} \right\},$$

$$\mathcal{P}_\Omega(\cdot)$$ denotes the projection onto the set

$$\Omega := \left\{ \omega \in [w^{\text{min}}, w^{\text{max}}], w \in \{p, q, e, \ell\}, \ell \in [0, \ell^{\text{max}}] \right\},$$

with $p^{\text{max}} := \{(p^{\text{max}})^{1/2} \circ (v^{\text{max}})^{1/2}, q^{\text{max}} := \{(q^{\text{max}})^{1/2} \circ (v^{\text{max}})^{1/2}, e^{\text{max}} := v^{\text{max}} - v^{\text{min}}, \mu^{\text{min}} := -q^{\text{max}}, \varepsilon^{\text{min}} := -e^{\text{max}}, \} \uparrow$ denotes the projection onto the interval $[0, +\infty)$. Notice that the 7-update by replacing $2 \frac{\partial L[k]}{\partial \tau}$ instead of simply using $\frac{\partial L[k]}{\partial \tau}$; this subtle change (to be clarified later when we present the convergence analysis) is due to the nonlinearity of the constraint (7). Let $x^* := [g^{(p)^*}, g^{(q)^*}, v^*]^\top$, $\omega^* := [p^*, q^*, e^*, \ell^*]^\top$, $\gamma^* := [\chi^*, \mu^*, \nu^*, \eta^*, \nu^*, \psi]^\top,$ and $\tau^* := \text{the equilibrium of (39).}$

In the proposed distributed version of (39), each node $i$ estimates the optimal values of only local primal and dual variables, denoted by $x^{(i)*} := [x^{(i)*\top}, \omega^{(i)*\top}, \gamma^{(i)*\top}, \tau^{(i)*\top}]^\top,$ with

$$x^{(i)*} := [g_i^{(p)}]^\top, [g_i^{(q)}]^\top, [v_i]^\top,$$

$$\omega^{(i)*} := \left\{ \left\{ p_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ p_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ q_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ q_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ e_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ e_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ \ell_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ \ell_i \right\}_{l, i} \in \mathcal{E}_p \right\}^\top,$$

$$\gamma^{(i)*} := \left\{ \left\{ \lambda_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ \mu_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ \nu_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ \eta_i \right\}_{l, i} \in \mathcal{E}_p \right\}^\top,$$

$$\tau^{(i)*} := \left\{ \left\{ v_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ v_i \right\}_{l, i} \in \mathcal{E}_p \right\}^\top.$$

Let $\chi^{(i)*} := \left[x^{(i)*\top}, \omega^{(i)*\top}, \gamma^{(i)*\top}, \tau^{(i)*\top} \right]^\top$ denote the vector of estimates of $x^{(i)*}$ maintained at node $i$ at time instant $k$, where

$$x_i[k] := [g_i^{(p)}]^\top, [g_i^{(q)}]^\top, [v_i]^\top,$$

$$\omega_i[k] := \left\{ \left\{ p_{ij} \right\}_{l, i} \in \mathcal{E}_p, \left\{ p_{ii} \right\}_{l, i} \in \mathcal{E}_p, \left\{ q_{ij} \right\}_{l, i} \in \mathcal{E}_p, \left\{ q_{ii} \right\}_{l, i} \in \mathcal{E}_p, \left\{ e_{ij} \right\}_{l, i} \in \mathcal{E}_p, \left\{ e_{ii} \right\}_{l, i} \in \mathcal{E}_p, \left\{ \ell_{ij} \right\}_{l, i} \in \mathcal{E}_p, \left\{ \ell_{ii} \right\}_{l, i} \in \mathcal{E}_p \right\}^\top,$$

$$\gamma_i[k] := \left\{ \left\{ \lambda_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ \mu_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ \nu_i \right\}_{l, i} \in \mathcal{E}_p, \left\{ \eta_i \right\}_{l, i} \in \mathcal{E}_p \right\}^\top.$$

Then, node $i$ performs updates using the following local Lagrangian:

$$L_i(x^{(i)}) = f_i(g_i^{(p)}) + \rho(v_i - 1)^2 + \lambda_i b_i^{(p)} + \mu_i b_i^{(q)} + \rho_1 (b_i^{(p)})^2 + \rho_2 (b_i^{(q)})^2 + \sum \frac{\rho_3}{2} \|p_{ij} - p_i\|_2^2 + \sum \frac{\rho_4}{2} \|q_{ij} - q_i\|_2^2 + \sum \frac{\rho_5}{2} \|e_{ij} - e_i\|_2^2 + \sum \frac{\rho_6}{2} \|\ell_{ij} - \ell_i\|_2^2 + \sum \frac{\rho_7}{2} \|v_i - v_i\|_2^2,$$

where

$$b_i^{(v)} := \varepsilon_i - 2R_i p_i - 2Xq_i + (R_i^2 + X_i^2) \ell_i,$$

$$b_i^{(p)} := g_i^{(p)} - f_i^{(p)} - M_i p_i + N_0 R\ell_i,$$

$$b_i^{(q)} := g_i^{(q)} - q_i^{(q)} - M_i q_i + N_0 X\ell_i.$$
\[ \hat{w}_{ij}[k+1] = \left[ (1 - a_{ij}[k]) \hat{w}_{ij}[k] + a_{ij}[k] \hat{w}_{ij}[k] \right. \\
- s_y \hat{y}_{ij}^{(w)}[k] \min_{wi}, \tag{40c} \]

\[ \hat{w}_{li}[k+1] = \left[ (1 - a_{li}[k]) \hat{w}_{li}[k] + a_{li}[k] \hat{w}_{li}[k] \right. \\
- s_y \hat{y}_{li}^{(w)}[k] \min_{wi}, \tag{40d} \]

\[ w \in \{ \ell, p, q, \varepsilon, \} , \]

\[ \hat{d}_{ij}[k+1] = \left[ (1 - a_{ij}[k]) \hat{d}_{ij}[k] + a_{ij}[k] \hat{d}_{ij}[k] \right. \\
+ s_y \hat{d}_{ij}^{(d)}[k] , \tag{40e} \]

\[ \hat{d}_{li}[k+1] = \left[ (1 - a_{li}[k]) \hat{d}_{li}[k] + a_{li}[k] \hat{d}_{li}[k] \right. \\
+ s_y \hat{d}_{li}^{(d)}[k] , \tag{40f} \]

\[ a_{ij}[k] = \begin{cases} 0.5 & \text{if } \{ i, j \} \in \mathcal{E}_p[k], \\ 0 & \text{otherwise}. \end{cases} \]

The gradients \( \hat{g}^{(w)}[k] := \{ \hat{g}^{(w)}_{ij}[k] \}_{i, j \in \mathcal{E}_p} \) and \( \hat{g}^{(w)}[k] := \{ \hat{g}^{(w)}_{li}[k] \}_{i, j \in \mathcal{E}_p} \), \( w \in \{ \ell, p, q, \varepsilon, \} \), are updated as follows:

\[ \hat{g}^{(w)}_{ij}[k+1] = \left[ (1 - a_{ij}[k]) \hat{g}^{(w)}_{ij}[k] + a_{ij}[k] \hat{g}^{(w)}_{ij}[k] \right. \\
+ 2 \left( \frac{\partial L^{(i)}[k]}{\partial \hat{w}_{ij}} \right), (i, j) \in \mathcal{E}_p, \tag{41a} \]

\[ \hat{g}^{(w)}_{li}[k+1] = \left[ (1 - a_{li}[k]) \hat{g}^{(w)}_{li}[k] + a_{li}[k] \hat{g}^{(w)}_{li}[k] \right. \\
+ 2 \left( \frac{\partial L^{(i)}[k]}{\partial \hat{w}_{li}} \right), (i, l) \in \mathcal{E}_p. \tag{41b} \]

We initialize (41) as follows:

\[ \hat{g}^{(w)}_{ij}[0] = 2 \frac{\partial L^{(i)}[0]}{\partial \hat{w}_{ij}} , (i, j) \in \mathcal{E}_p, \tag{42a} \]

\[ \hat{g}^{(w)}_{li}[0] = 2 \frac{\partial L^{(i)}[0]}{\partial \hat{w}_{li}} , (i, l) \in \mathcal{E}_p. \tag{42b} \]

We noticed from the numerical simulations that if the initial voltage magnitudes are set to 1.0 per unit, \( y^{(0)}[0] \) must be initialized differently to achieve a better performance. If (42) is used, then,

\[ \hat{g}^{(w)}[0] = \left[ \frac{\hat{g}^{(w)}_{ij}[0]}{\hat{g}^{(w)}_{li}[0]} \right] = \left[ \frac{\hat{g}^{(w)}_{ij}[0]}{\hat{g}^{(w)}_{li}[0]} \right] = 2 \left[ M_0^v v[0] \right]. \tag{43} \]

If \( V_i[0] = 1 \), \( i \in \mathcal{V}_p \), then, the second term on the right-hand side of (43) does not have any effect on the average estimate, \( \frac{1}{2} \left( \hat{g}^{(w)}[0] + \hat{g}^{(w)}[0] \right) \), since \( M_0^v v[0] + N_0^v v[0] = M v[0] = 0 \). In view of this observation and the fact that the second term in (43) can be significantly larger than the first term, it is better to neglect it during the initialization.

### B. Feedback Interconnection Representation of the Distributed Primal-Dual Algorithm

Similar to the feedback representation given in Section III-B, we show that (40)-(41) can also be represented as a feedback interconnection of a nominal system, denoted by \( \mathcal{H}_1^{opf} \), and a disturbance system, denoted by \( \mathcal{H}_2^{opf} \), which allows us to utilize the small-gain theorem for convergence analysis purposes. To this end, let

\[ \tilde{w}[k] := \left[ \{ \hat{w}_{ij}[k] \}_{i, j \in \mathcal{E}_p}, \hat{w}_{li}[k] := \left[ \{ \hat{w}_{li}[k] \}_{i, j \in \mathcal{E}_p} \right], \tag{44a} \]

\[ \tilde{\omega}[k] := \left[ \{ \hat{\omega}_{ij}[k] \}_{i, j \in \mathcal{E}_p}, \tilde{\omega}_{li}[k] := \left[ \{ \hat{\omega}_{li}[k] \}_{i, j \in \mathcal{E}_p} \right] \right], \tag{44b} \]

\[ \tilde{\gamma}_k := \left[ \{ \hat{\gamma}_k \}_{k \in \mathcal{E}_p}, \gamma_k := \left[ \{ \hat{\gamma}_k \}_{k \in \mathcal{E}_p} \right] \right] \tag{44c} \]

By using (44), we compactly write the iterations for \( x, \omega, \gamma, \), and \( \tau \), which constitute the nominal system, \( \mathcal{H}_1^{opf} \), given by:

\[ \mathcal{H}_1^{opf} : x[k+1] = P_x \left( x[k] - s \frac{\partial \mathcal{L}[k]}{\partial x} + e_x[k] \right), \tag{44a} \]

\[ \tilde{\omega}[k+1] = \frac{1}{2} P_\Omega \left( \tilde{\omega}[k] - s \frac{\partial \mathcal{L}[k]}{\partial \omega} + e_\omega[k] \right) \]

\[ \tilde{\gamma}_k[+1] = \frac{1}{2} P_\Omega \left( \tilde{\gamma}_k[+1] - s \frac{\partial \mathcal{L}[k]}{\partial \gamma} + e_\gamma[k] \right) \tag{44b} \]

\[ \tau[k+1] = \tau[k] + 2 \frac{\partial \mathcal{L}[k]}{\partial \tau} + e_\epsilon[k] \tag{44c} \]

where \( \mathcal{L}[k] := L(x, \omega[k], \gamma_k[\tau], \epsilon) \), \( e[k] := \left[ e_x[k] \right] \tag{44d} \]

\[ e[k] = A \left[ \left[ \hat{\omega}_k - \tilde{\omega}_k \right], \hat{\gamma}_k[\tau] \right] \left[ \hat{\gamma}_k - \tilde{\gamma}_k \right], \right. \]

for some constant matrix \( A \). Notice that \( e[k] \) results from \( \left( \hat{\omega}_k, \hat{\gamma}_k, \hat{\gamma}_k \right) \), and \( \left( \hat{\gamma}_k, \hat{\gamma}_k \right) \) deviating from their respective average, \( \bar{\omega}_k, \bar{\gamma}_k \), and \( \bar{\gamma}_k \); without \( e[k] \), the nominal system \( \mathcal{H}_1^{opf} \) has exactly the same form as (39). Now, we define the disturbance system, \( \mathcal{H}_2^{opf} \), as follows:

\[ \hat{w}_{ij}[k+1] = \left[ (1 - a_{ij}[k]) \hat{w}_{ij}[k] + a_{ij}[k] \hat{w}_{ij}[k] \right. \\
- s_y \hat{y}_{ij}^{(w)}[k] \min_{wi}, \tag{45a} \]

\[ \hat{w}_{li}[k+1] = \left[ (1 - a_{li}[k]) \hat{w}_{li}[k] + a_{li}[k] \hat{w}_{li}[k] \right. \\
- s_y \hat{y}_{li}^{(w)}[k] \min_{wi}. \tag{45b} \]
C. Convergence Analysis

Representing algorithm (40)–(41) in a feedback form facilitates our analysis relying on the small-gain theorem. Most of the analysis is dedicated to establishing that $\mathcal{H}_{1}^{opf}$ and $\mathcal{H}_{2}^{opf}$ are finite-gain stable. Then, the small-gain theorem is applied in a straightforward manner to show the convergence of (40)–(41).

In the next result, we establish that $\mathcal{H}_{1}^{opf}$ is finite-gain stable.

Proposition 5. Let Assumption 7 hold. Then, under (44), we have that

$$\text{R1. } \|z\|_{2}^{a} \leq \alpha_{1}\|\nu\|_{2}^{a} + \beta_{1},$$

(46)

for some positive $\alpha_{1}$ and $\beta_{1}$, $a \in (0,1)$, and sufficiently small $s > 0$, where

$$z[k] := \begin{bmatrix} x[k] - x^{*} \\ \varpi[k] - \omega^{*} \\ \gamma[k] - \gamma^{*} \\ \tau[k] - \tau^{*} \end{bmatrix},$$

Proof. Letting $L^{*} := L(x^{*}, \omega^{*}, \gamma^{*}, \tau^{*})$,

$$G[k] := \begin{bmatrix} x[k] - s \frac{\partial L[k]}{\partial x} \\ \varpi[k] - s \frac{\partial L[k]}{\partial \varpi} \\ \gamma[k] + s \frac{\partial L[k]}{\partial \gamma} \\ \tau[k] + 2s \frac{\partial L[k]}{\partial \tau} \end{bmatrix},$$

we establish the following result.

Lemma 2.

$$\|z[k+1]\| \leq \|G[k] - G^{*}\| + \|e[k]\|,$$

(47)

where $\| \cdot \|$ is a vector norm.

Proof. We note that the following relationship holds:

$$x^{*} = \mathcal{P}_{X}(x^{*} - s \frac{\partial L^{*}}{\partial x}).$$

(48)

Then, by applying the triangle inequality on multiple occasions and the Projection Theorem [24, Proposition 2.1.3], we have that

$$\|x[k+1] - x^{*}\| = \left\| \frac{1}{2} \mathcal{P}_{X}(x[k] - s \frac{\partial L[k]}{\partial x} + e[k]) ight\|$$

$$+ \frac{1}{2} \left\| \mathcal{P}_{X}(x[k] - s \frac{\partial L[k]}{\partial x} - e[k]) - x^{*} \right\|$$

$$\leq \frac{1}{2} \left\| \mathcal{P}_{X}(x[k] - s \frac{\partial L[k]}{\partial x} - e[k]) - x^{*} \right\|$$

$$+ \frac{1}{2} \left\| \mathcal{P}_{X}(x[k] - s \frac{\partial L[k]}{\partial x} - e[k]) - x^{*} \right\|$$

$$= \frac{1}{2} \left\| \mathcal{P}_{X}(x[k] - s \frac{\partial L[k]}{\partial x} + e[k]) \right\|$$

$$- \mathcal{P}_{X}\left(x^{*} - s \frac{\partial L^{*}}{\partial x}\right)$$

$$+ \frac{1}{2} \left\| \mathcal{P}_{X}(x[k] - s \frac{\partial L[k]}{\partial x} - e[k]) - x^{*} \right\|$$

$$- \mathcal{P}_{X}(x^{*} - s \frac{\partial L^{*}}{\partial x})$$

(49a)

$$\leq \frac{1}{2} \left\| x[k] - s \frac{\partial L[k]}{\partial x} - x^{*} + s \frac{\partial L^{*}}{\partial x} + e[k] \right\|$$

$$+ \frac{1}{2} \left\| x[k] - s \frac{\partial L[k]}{\partial x} - x^{*} + s \frac{\partial L^{*}}{\partial x} + e[k] \right\|$$

(49b)

$$\leq \left\| x[k] - s \frac{\partial L[k]}{\partial x} + e[k] \right\|$$

(49c)

where we used (49a) to obtain (49a), applied the triangle inequality to obtain (49b), and (49c), used (48) to obtain (49c).
and the projection theorem to obtain \(|94d\). Similarly, we find that
\[
\left\| \pi[k + 1] - \omega^* \right\| \leq \left\| \pi[k] - \left( \frac{\partial L[k]}{\partial \omega} - \omega^* + \frac{\partial L^*}{\partial \omega} \right) \right\| + \left\| e_{\omega}[k] \right\|, \tag{49f}
\]
\[
\left\| \tau[k + 1] - \gamma^* \right\| \leq \left\| \tau[k] - \left( \frac{\partial L[k]}{\partial \gamma} - \omega^* + \frac{\partial L^*}{\partial \gamma} \right) \right\| + \left\| e_{\gamma}[k] \right\|, \tag{49g}
\]
\[
\left\| \tau[k + 1] - \tau^* \right\| \leq \left\| \tau[k] - \left( \frac{\partial L[k]}{\partial \tau} - \tau^* + \frac{\partial L^*}{\partial \tau} \right) \right\| + \left\| e_{\tau}[k] \right\|, \tag{49h}
\]
yielding \(|47\).

Next, it follows from the mean value theorem \([27\, Theorem \, 5.1]\) applied to each component in \(\nabla f(g^{(p)}[k]) - \nabla f(g^{(p)*})\) that
\[
\nabla f(g^{(p)}[k]) - \nabla f(g^{(p)*}) = \nabla^2 f(v[k])(g^{(p)}[k] - g^{(p)*}), \tag{50}
\]
where \(v[k] := [v_1[k], v_2[k], \ldots, v_n[k]]^\top\), with \(v_i[k]\) lying on the line segment connecting \(g_i^{(p)}[k]\) and \(g_i^{(p)*}\), and \(\nabla^2 f(v[k])\) is the Hessian of \(f(x)\) at \(x = v[k]\). Then, by using \(|50\), we have that
\[
G[k] - G^* = F[k]z[k] + H[k]\begin{bmatrix} x^* \\ \omega^* \\ \gamma^* \\ \tau^* \end{bmatrix}, \tag{51}
\]
where
\[
H[k] := \begin{bmatrix} 0 & s(C[k]^\top - C^*[k]) \\ s(C[k]^\top - C^*[k]) & 0 \end{bmatrix},
\]
for some suitable matrices \(C^*\) and \(C[k]\) corresponding to \((x^*, \omega^*, \gamma^*, \tau^*)\) and \((\pi[k], \lambda[k], \tau[k])\), respectively, and
\[
F[k] := \begin{bmatrix} I - s(D[k]^\top + \Upsilon) & -sC[k]^\top \\ sC[k] & I \end{bmatrix},
\]
where \(\Upsilon \in \mathbb{R}^{(3n+4)\times(3n+4)}\) is a positive-semidefinite matrix, and the matrix \(D[k]\) is a block diagonal matrix given by
\[
D[k] := \begin{bmatrix} \nabla^2 f(v[k]) & 0_{n\times n} \\ 0_{n\times n} & \rho I_n \end{bmatrix},
\]
where \(0_{n\times n}\) is the all-zeros matrix, and \(I_n \in \mathbb{R}^{n\times n}\) is the identity matrix. We note that \(F[k]\) is a skew-symmetric matrix resulting from multiplying \(\frac{\partial F[k]}{\partial \Omega[k]}\) by a factor of 2 in the \(\tau\)-update in algorithms \(|40| - |41|\). Define
\[
B[k] := \begin{bmatrix} D[k] & C[k] \\ -C[k]^\top & 0 \end{bmatrix},
\]
so that \(F[k] = I - sB[k]\). The next result can be established by using some standard analysis.

**Lemma 3.** If, at time instant \(k\),
\[
B[k]\begin{bmatrix} 0_n^H \\ x^H \\ y^H \\ 0_{|\epsilon_p|}^H \\ z^H \end{bmatrix} = 0,
\]
for some \(x \in \mathbb{R}^n\), \(y \in \mathbb{R}^{3|\epsilon_p|}\), and \(z \in \mathbb{R}^{2+3|\epsilon_p|}\), then, we have that \((x, y, z) = 0\), where \(0_n\) denotes the all-zeros vector of length \(n\), and \(x^H\) denotes the Hermitian transpose of \(x\).

Next, we show that all eigenvalues of \(B[k]\) have a strictly positive real part. Suppose \(\mu\) is an eigenvalue of \(B[k]\) and \([\zeta^H, \omega^H]^H\) is an eigenvector corresponding to \(\mu\), where \(\zeta := (\zeta^{(1)}^\top, \zeta^{(2)}^\top, \zeta^{(3)}^\top)^\top\) has the same number of rows as \(D[k]\) such that \(\zeta^{(1)} \in \mathbb{C}^n\), \(\zeta^{(2)} \in \mathbb{C}^n\), \(\zeta^{(3)} \in \mathbb{C}^n\), \(\zeta^{(4)} \in \mathbb{C}^{3|\epsilon_p|}\), and \(\zeta^{(5)} \in \mathbb{C}^{2+3|\epsilon_p|}\). Then, on the one hand, we have that
\[
\Re\left(\langle \zeta^H, \omega^H \rangle B[k]\begin{bmatrix} \zeta \\ \omega \end{bmatrix}\right) = \Re\left(\mu\begin{bmatrix} \zeta^H & \omega^H \end{bmatrix}\begin{bmatrix} \zeta \\ \omega \end{bmatrix}\right) = \Re(\mu)(\|\zeta\|^2 + \|\omega\|^2).
\]
On the other hand, we have that
\[
\Re\left(\langle \zeta^H, \omega^H \rangle B[k]\begin{bmatrix} \zeta \\ \omega \end{bmatrix}\right) = \Re\left(\zeta^H D[k]\zeta + \zeta^H C[k]w - w^H C[k]^\top \zeta - \rho\|\zeta^{(3)}\|^2 + \rho\|\zeta^{(5)}\|^2 + \zeta^H \Upsilon \zeta > 0, \right.
\]
if \((\zeta^{(1)}^\top, \zeta^{(3)}^\top) \neq 0\), or \(\zeta \notin \text{null}(\Upsilon)\), denoting the null space of \(\Upsilon\). If \(\Re(\mu) = 0\), then, it follows that \(\zeta \in \text{null}(\Upsilon)\), \((\zeta^{(1)}, \zeta^{(3)}, \zeta^{(5)}) = 0\), and
\[
B[k]\begin{bmatrix} 0_n^H \\ \zeta^{(2)}^H \\ \zeta^{(4)}^H \\ 0_{|\epsilon_p|}^H \end{bmatrix} = 0. \tag{52}
\]
By Lemma \(|3| \tag{52}\) holds only if \(\zeta = 0\) and \(w = 0\), which contradicts the fact that \([\zeta^H, \omega^H]^H \neq 0\). Therefore, all eigenvalues of \(B[k]\) have a strictly positive real part, and, for some small enough \(s\), the spectral radius of \(F[k]\) denoted by \(\rho(F[k])\) is strictly less than 1. The next results can be established by using some standard analysis.

**Lemma 4.** For sufficiently small \(s\), there exists an induced matrix norm \(|\cdot|\) such that \(|F[k]| \leq b_1\), for some \(b_1 < 1\), \forall k.

**Lemma 5.**
\[
\|H[k]|x^*, \omega^*, \gamma^*, \tau^*|H| \leq sb_2\|z[k]\|, \text{ for some positive } b_2, \forall k.
\]
Taking \(|\cdot|\) on both sides of \(|51\) and applying the triangle inequality yields
\[
\|G[k] - G^*\| \leq \|F[k]\|\|z[k]\| + \|H[k]|x^*, \omega^*, \gamma^*, \tau^*|H\|, \tag{53}
\]
By applying Lemmas \(|2|\), \(|4|\), and \(|5|\) and using the inequality \(|53\), we obtain that
\[
\|z[k + 1]\| \leq \|G[k] - G^*\| + \|e[k]\| \leq \|F[k]\|\|z[k]\| + \|H[k]|x^*, \omega^*, \gamma^*, \tau^*|H\| + \|e[k]\| \leq (b_1 + sb_2)\|z[k]\| + \|e[k]\| = b\|z[k]\| + \|e[k]\|, \tag{54}
\]

where \( b := b_1 + s b_2 < 1 \) for sufficiently small \( s \). Now, by multiplying both sides of (54) by \( a^{-(k+1)} \), we obtain

\[
a^{-(k+1)}\|z[k+1]\| \leq \frac{b}{a} a^{-k} \|z[k]\| + a^{-(k+1)}\|e[k]\|. \tag{55}
\]

Then, by taking \( \max_{0 \leq k \leq K} \cdot \) on both sides of (55), we obtain

\[
\max_{0 \leq k \leq K} a^{-(k+1)}\|z[k+1]\| \leq \frac{b}{a} \max_{0 \leq k \leq K} a^{-k} \|z[k]\| + \frac{1}{a} \max_{0 \leq k \leq K+1} a^{-k} \|e[k]\|.
\tag{56}
\]

Since

\[
\max_{0 \leq k \leq K} a^{-(k+1)}\|z[k+1]\| = \max_{0 \leq k \leq K+1} a^{-k} \|z[k]\| - \|z[0]\|
\]

the relation (56) can be written as

\[
\|z\|_{a,K+1} \leq \frac{b}{a} \|z\|_{a,K} + \frac{1}{a} \|e\|_{a,K+1} + \|z[0]\|, \tag{57}
\]

where \( \|z\|_{a,K} := \max_{0 \leq k \leq K} a^{-k} \|z[k]\| \). Since \( \|z\|_{a,K+1} \geq \|z\|_{a,K} \), it follows from (57) that

\[
\|z\|_{a,K} \leq \frac{b}{a} \|z\|_{a,K} + \frac{1}{a} \|e\|_{a,K} + \|z[0]\|.
\tag{58}
\]

Then, after rearranging (58), we obtain

\[
\|z\|_{a,K} \leq \frac{b}{a-b} \|e\|_{a,K} + \frac{a}{a-b} \|z[0]\|.
\]

Because \( \|\cdot\|_2 \leq \alpha \|\cdot\| \) and \( \|\cdot\| \leq \beta \|\cdot\|_2 \) for some \( \alpha \) and \( \beta \), we have that \( \|z\|_{a,K} \leq \|z\|_{a,K}^\alpha / \alpha, \|e\|_{a,K} \leq \beta \|e\|_{a,K} \). Hence,

\[
\frac{1}{\alpha} \|z\|_{a,K}^2 \leq \beta \|e\|_{a,K} + \frac{a}{a-b} \|z[0]\|,
\]

which can be rewritten as

\[
\|z\|_{a,K}^2 \leq \alpha_1 \|e\|_{a,K} + \beta_1,
\]

where

\[
\alpha_1 = \beta \alpha / (a-b),
\]

and

\[
\beta_1 = a \alpha / (a-b) \|z[0]\|,
\]

yielding (61).

Next, we establish that \( \mathcal{H}_{2}^{opf} \) is finite-gain stable.

**Proposition 6.** Let Assumptions 7 and 2 hold. Then, under (43), we have that

\[
\|e\|_{a,K} \leq s \alpha_3 \|z\|_{a,K} + s \alpha_5 \|\hat{y}\|_{a,K}, \tag{60}
\]

for some positive \( \alpha_2 \) and \( \beta_2, a \in (0,1), \) and sufficiently small \( s > 0 \).

**Proof.** In our analysis, we need the following results stated without proofs.

**Lemma 6.** \( \|e\|_{a,K} \leq s \alpha_3 \|z\|_{a,K} + s \alpha_5 \|\hat{y}\|_{a,K}, \tag{60} \)

for some constant \( \alpha_3, \alpha_4, \) and \( \alpha_5, \) where \( \hat{z}[k] := \left[ (\hat{x}[k] - \hat{\mu}[k])^T, (\hat{x}[k] - \hat{\nu}[k])^T, (\hat{x}[k] - \hat{\xi}[k])^T, (\hat{y}[k] - \hat{\eta}[k])^T \right]^T \), \( \hat{y}[k] := \left[ (\hat{y}[k])^T, (\hat{y}^{\mu}[k])^T, (\hat{y}^{\nu}[k])^T, (\hat{y}^{\xi}[k])^T \right]^T \).

**Lemma 7.** \( \|\hat{z}\|_{a,K} \leq \beta_3 \|\hat{y}\|_{a,K} + \beta_0, \tag{61} \)

for some constant \( \beta_3 \) and \( \beta_0. \)

**Lemma 8.** For \( w \in \{ \ell, p, q, \varepsilon, \nu, \eta \}, \) let

\[
\hat{\delta}_i (w) \mid k+1 \rangle := 2 \left( \frac{\partial L_{ij}^i (w) \mid k+1 \rangle}{\partial \hat{w}_i} - \frac{\partial L_{ij}^i (w) \mid k \rangle}{\partial \hat{w}_i} \right), (i, j) \in \mathcal{E}_p,
\]

\[
\hat{\gamma}_i (w) \mid k+1 \rangle := 2 \left( \frac{\partial L_{ij}^i (w) \mid k+1 \rangle}{\partial \hat{w}_i} - \frac{\partial L_{ij}^i (w) \mid k \rangle}{\partial \hat{w}_i} \right), (i, i) \in \mathcal{E},
\]

\[
\hat{\delta}_i (w) \mid k \rangle := \left[ \delta (w) \mid k \rangle, \hat{\delta} (w) \mid k \rangle \right]^T, \delta (w) \mid k \rangle := \left[ \delta (w) \mid k \rangle, \delta (w) \mid k \rangle \right]^T.
\]

Then, for some constant \( \zeta_0 \) and \( \zeta_1, \) the following relation holds:

\[
\|\hat{y}\|_{a,K} \leq \zeta_1 \|\hat{z}\|_{a,K} + \zeta_0.
\tag{62}
\]

**Lemma 9.** \( \|\hat{\delta}_i \|_{a,K} \leq \kappa_1 \|\hat{z}\|_{a,K} + \kappa_0, \tag{63} \)

for some constant \( \kappa_0 \) and \( \kappa_1. \)

By substituting (61) for \( \|\hat{z}\|_{a,K} \) in (60), we obtain

\[
\|e\|_{a,K} \leq s \alpha_3 \|z\|_{a,K} + s \alpha_5 \|\hat{y}\|_{a,K} + \alpha_1 \|\hat{z}\|_{a,K} + s \alpha_5 \|\hat{y}\|_{a,K} + \alpha_5 \|\hat{y}\|_{a,K}
\]

\[
\leq s \alpha_3 \|z\|_{a,K} + \alpha_4 (s \alpha_5 \|\hat{y}\|_{a,K} + \beta_0) + s \alpha_5 \|\hat{y}\|_{a,K}
\]

\[
= s \alpha_4 \beta_3 + s \alpha_5 \|\hat{y}\|_{a,K} + s \alpha_5 \|\hat{y}\|_{a,K} + \alpha_4 \beta_0
\]

\[
\leq s \alpha_4 \beta_3 + s \alpha_5 \|\hat{z}\|_{a,K} + \kappa_0 + s \alpha_3 \|z\|_{a,K} + \alpha_4 \beta_0
\]

\[
\leq s \alpha_4 \beta_3 + s \alpha_5 \|\hat{z}\|_{a,K} + \kappa_0 + s \alpha_3 \|z\|_{a,K} + \alpha_4 \beta_0
\]

\[
+ \alpha_4 \beta_0 = s \zeta_2 \|\hat{z}\|_{a,K} + s \alpha_3 \|z\|_{a,K} + \kappa_3,
\tag{64}
\]

where in the last inequality we applied (62), \( \zeta_2 := (s \alpha_4 \beta_3 + s \alpha_5) \zeta_1 + s \alpha_4 \beta_0 \). By substituting (63) for \( \|\hat{\delta}_i \|_{a,K} \) in (64), we obtain

\[
\|e\|_{a,K} \leq s \zeta_2 \|\hat{z}\|_{a,K} + \kappa_0 + s \alpha_3 \|z\|_{a,K} + s \zeta_3,
\]

which can be rewritten as

\[
\|e\|_{a,K} \leq s \zeta_2 \|\hat{z}\|_{a,K} + \kappa_0 + s \alpha_3 \|z\|_{a,K} + s \zeta_3,
\]

 yielding (65).

By using Propositions 5,6, we now show that (40)–(41) converges geometrically fast.

**Proposition 7.** Let Assumptions 7 and 2 hold. Then, under algorithm (40)–(41), we have that

\[
\|z\|_{a,K} \leq \beta,
\tag{66}
\]

for some \( a \in (0,1), \) \( \beta > 0, \) and sufficiently small \( s > 0. \) In particular, \( \hat{x} := (g^{(p)}) \varepsilon, \hat{y} := (g^{(q)}) \varepsilon \) and \( \hat{x} := (g^{(p)}) \varepsilon, \hat{y} := (g^{(q)}) \varepsilon \) converge to \( x^* := (g^{(p)}) \varepsilon, p^*, p^*, q^*, \varepsilon^*, \varepsilon^* \) at a geometric rate \( O(a^k). \)
Proof. By using Propositions 5 and 6 it follows that
\[ \|z\|_2^a + \alpha_1 \|e\|_2^a + \beta_1 \leq \alpha_1 (\alpha_2 \|\hat{z}\|_2^a + \beta_2) + \beta_1, \]
which, after rearranging, results in
\[ \|z\|_2^a \leq \frac{\alpha_1 \beta_2 + \beta_1}{1 - s\alpha_1 + \alpha_2} =: \beta, \]
yielding (66). Hence, for sufficiently small s, we have that \(s\alpha_1 + \alpha_2 < 1\), which ensures that \(\beta\) is finite. Next, we show that \(\hat{x}[k]\) and \(\bar{x}[k]\) converge to \(x^*\). By substituting (63) for \(\|z\|_2^a\) in (62), we have that
\[ \|\bar{y}\|_2^a \leq \zeta_1 \|\hat{y}\|_2^a + \zeta_0 \leq \zeta_1 (\kappa_3 \|\hat{z}\|_2^a + \kappa_0) + \zeta_0 \]
for some positive \(\kappa_3\) and \(\kappa_4\). Then, by using (67) and (66) in (61), we have that
\[ \|z\|_2^a \leq s\beta_3 \|\bar{y}\|_2^a + \beta_0 \leq s\beta_3 (\kappa_4 \|\bar{z}\|_2^a + \kappa_4) + \beta_0 \leq \kappa_5, \]
for some \(\kappa_5 > 0\). Letting \(\bar{x} := (\hat{x} + \bar{x})/2\), and using the triangle inequality and the inequalities (46) and (68), we obtain that
\[ \|\bar{x} - x^*\|_2^a = \|\bar{x} - \bar{x}^* + \bar{x} - x^*\|_2^a \leq \|\bar{z}\|_2^a + \|\hat{z}\|_2^a \leq \beta + \kappa_5, \]
from where it follows that \(\bar{x}[k]\) and \(\bar{x}[k]\) converge to \(x^*\). □

Finally, we prove that \(x^*\) is the solution of rSOCP. [The proof is similar to the proof of Lemma 1]

Lemma 10. Consider \((x^*, \omega^*, \gamma^*, \tau^*)\), namely, the equilibrium of (39). Then, \(x^*\) is the solution of rSOCP.

D. Time-Varying Directed Communication Graphs

The design of the robust extension of [40–41] follows exactly the development procedure outlined in Section III-E.

In the robust extension, we let node i perform the same updates for \(\hat{y}[i]\), \(\hat{q}[i]\), \(\lambda_i\), \(\mu_i\), \(\tau_i\), \((i, i) \in \mathcal{E}_p\) as the ones in [40–41], i.e.,
\[ \chi[k + 1] = \left[ \chi[k] - s \frac{\partial L(i)[k]}{\partial \chi} \right]_{\chi_{min}} \chi \in \{\hat{y}[i], \hat{q}[i], \psi \}, \]
\[ \psi[k + 1] = \psi[k] + s \frac{\partial L(i)[k]}{\partial \psi}, \psi \in \{\lambda_i, \mu_i\}, \]
\[ \tau_i[k + 1] = \left[ \tau_i[k] + 2s \frac{\partial L(i)[k]}{\partial \tau_i} \right]_{+}, \]
(70a, 70b, 70c)
The only difference between algorithm [40–41] and its robust extension is how the averaging step is performed in the updates of the estimates of the shared quantities, namely, \(\hat{\omega}[k] := \hat{\rho}[k], \hat{\omega}[k], \bar{\omega}[k], \tilde{\omega}[k], \hat{\omega}[k], \tilde{\omega}[k], \hat{\omega}[k], \tilde{\omega}[k]\), \(\hat{\omega}[k] := \hat{\rho}[k], \hat{\omega}[k], \bar{\omega}[k], \tilde{\omega}[k], \hat{\omega}[k], \tilde{\omega}[k]\), \(\hat{\omega}[k] := \hat{\rho}[k], \hat{\omega}[k], \bar{\omega}[k], \tilde{\omega}[k], \hat{\omega}[k], \tilde{\omega}[k]\), and the gradient vectors, \(\hat{g}[w][k] := \{\hat{g}[w][k](i, j) \in \mathcal{E}_p\}\), and \(\bar{g}[w][k] := \{\bar{g}[w][k](i, j) \in \mathcal{E}_p\}\). In the robust extension, we make use of the alternating averaging protocol to execute the averaging step; namely, for each \((i, j) \in \mathcal{E}_p\), node i runs the following iterations:
\[ \hat{\omega}_{ij}[k + 1] = \left[ (1 - a_{ij}[k]) \hat{\omega}_{ij}[k] + a_{ij}[k] \hat{\omega}_{ij}[k] \right]_{\hat{\omega}_{ij}^{max}}, \]
\[ \hat{\omega}_{ij}[k + 1] = \left[ (1 - a_{ij}[k]) \hat{\omega}_{ij}[k] + a_{ij}[k] \hat{\omega}_{ij}[k] \right]_{\hat{\omega}_{ij}^{max}}, \]
\[ \hat{\omega}_{ij}[k + 1] = \left[ (1 - a_{ij}[k]) \hat{\omega}_{ij}[k] + a_{ij}[k] \hat{\omega}_{ij}[k] \right]_{\hat{\omega}_{ij}^{max}}, \]
\[ \hat{\omega}_{ij}[k + 1] = \left[ (1 - a_{ij}[k]) \hat{\omega}_{ij}[k] + a_{ij}[k] \hat{\omega}_{ij}[k] \right]_{\hat{\omega}_{ij}^{max}}, \]
(71a, 71b, 71c)
while node j executes the iterations given below:
\[ \hat{\omega}_{ij}[k + 1] = \left[ (1 - a_{ji}[k]) \hat{\omega}_{ij}[k] + a_{ji}[k] \hat{\omega}_{ij}[k] \right]_{\hat{\omega}_{ij}^{max}}, \]
\[ \hat{\omega}_{ij}[k + 1] = \left[ (1 - a_{ji}[k]) \hat{\omega}_{ij}[k] + a_{ji}[k] \hat{\omega}_{ij}[k] \right]_{\hat{\omega}_{ij}^{max}}, \]
\[ \hat{\omega}_{ij}[k + 1] = \left[ (1 - a_{ji}[k]) \hat{\omega}_{ij}[k] + a_{ji}[k] \hat{\omega}_{ij}[k] \right]_{\hat{\omega}_{ij}^{max}}, \]
\[ \hat{\omega}_{ij}[k + 1] = \left[ (1 - a_{ji}[k]) \hat{\omega}_{ij}[k] + a_{ji}[k] \hat{\omega}_{ij}[k] \right]_{\hat{\omega}_{ij}^{max}}, \]
(72a, 72b, 72c)
with \(w \in \{\ell, p, q, \varepsilon\}, d \in \{\nu, \eta\}, \) and \(\psi \in \{\ell, p, q, \varepsilon, \nu, \eta\}, \) where \(a_{ij}[k], \hat{a}_{ij}[k], \bar{a}_{ij}[k], \tilde{a}_{ij}[k], \rho_{ij}[w][k], \) and \(\bar{\rho}_{ij}[w][k], \) \((i, j) \in \mathcal{E}_p, w \in \{\ell, p, q, \varepsilon, \nu, \eta\}, \) are updated using the alternating averaging protocol provided below:

node i:
\[ \tilde{\phi}_{ij}[k] = \begin{cases} \phi_{ij}[k] & \text{if } (j, i) \in \mathcal{E}_c[k], \\ \phi_{ij}[k - 1] & \text{otherwise}, \end{cases} \]
\[ \phi_{ij}[k] = \begin{cases} \phi_{ij}[k] & \text{if } (j, i) \in \mathcal{E}_c[k], \\ \phi_{ij}[k - 1] & \text{otherwise}, \end{cases} \]
\[ a_{ij}[k] = \begin{cases} 0.5 & \text{if } (j, i) \in \mathcal{E}_c[k], \hat{\phi}_{ij}[k] \neq \bar{\phi}_{ij}[k - 1], \\ 0 & \text{otherwise}, \end{cases} \]
node j:
\[ \hat{\phi}_{ij}[k] = \begin{cases} \phi_{ij}[k] & \text{if } (j, i) \in \mathcal{E}_c[k], \\ \phi_{ij}[k - 1] & \text{otherwise}, \end{cases} \]
\[ \phi_{ij}[k] = \begin{cases} \phi_{ij}[k] & \text{if } (j, i) \in \mathcal{E}_c[k], \\ \phi_{ij}[k - 1] & \text{otherwise}, \end{cases} \]
\[ a_{ji}[k] = \begin{cases} 0.5 & \text{if } (j, i) \in \mathcal{E}_c[k], \hat{\phi}_{ij}[k] \neq \bar{\phi}_{ij}[k - 1], \\ 0 & \text{otherwise}, \end{cases} \]
\[ \hat{\rho}_{ij}[w][k] = \begin{cases} \hat{\omega}_{ij}[k] & \text{if } (j, i) \in \mathcal{E}_c[k], \hat{\phi}_{ij}[k] \neq \bar{\phi}_{ij}[k - 1], \\ 0 & \text{otherwise}, \end{cases} \]
\[ \bar{\rho}_{ij}[w][k] = \begin{cases} \hat{\omega}_{ij}[k] & \text{if } (j, i) \in \mathcal{E}_c[k], \hat{\phi}_{ij}[k] \neq \bar{\phi}_{ij}[k - 1], \\ 0 & \text{otherwise}, \end{cases} \]
\[ \hat{\rho}_{ij}[k] = \begin{cases} \tilde{y}_i^{[w]}[t_k] & \text{if } (i, j) \in \mathcal{E}_c[k], \hat{\phi}_{ij}[k] \neq \hat{\phi}_{ij}[k-1], \\ 0 & \text{otherwise.} \end{cases} \tag{73} \]

The following result can be easily established following the analysis from the proofs of Propositions 5 and 7.

**Proposition 8.** Let Assumptions 1 and 3 hold. Then, under algorithm (70)–(73), we have that

\[ \|z\|_{a,K} \leq \beta, \tag{74} \]

for some \( a \in (0, 1) \), \( \beta > 0 \), and sufficiently small \( s > 0 \). In particular, \( \hat{x} := (g^{(p)}, g^{(q)}, v, \tilde{p}, \tilde{q}, \tilde{\ell}, \tilde{\xi}) \) and \( \hat{x} := (g^{(p)}, g^{(q)}, v, \tilde{p}, \tilde{q}, \tilde{\ell}, \tilde{\xi}) \) converge to \( x^* := (g^{(p)*}, g^{(q)*}, v^*, \tilde{p}^*, \tilde{q}^*, \tilde{\ell}^*, \tilde{\xi}^*) \) at a geometric rate \( O(a^s) \).

### E. Numerical Simulations

In the following, we present numerical results to illustrate the performance of the robust distributed primal-dual algorithm (70)–(73) over directed graph \( \mathcal{G}(c)[k] \) using the IEEE 69–bus radial test system [29].

A subset of buses is designated to have a DER. For a DER at bus \( i \), we choose \( f_i(p_i) = a_i p_i^2 \), where \( a_i > 0 \) is randomly selected. It is assumed that each communication link becomes inactive with probability 0.4. The algorithm uses a constant stepsize \( s = 3 \times 10^{-2} \). For initialization, we use \( v_i[0] = 1, i \in V_p, g^{(p)}[0] = l_p[0], g^{(q)}[0] = l_q[0] \), and the initial values of the remaining variables, except for the gradients \( \hat{y}^{(w)}[0] \) and \( \hat{y}^{(w)}[0] \), \( w \in \{ \ell, p, q, \xi, \nu, \eta \} \), are set to zero. The initial values of the gradients are computed using (42), except for \( g^{(p)}[0] \) and \( g^{(q)}[0] \), which are computed by neglecting the voltages, \( v_i[0]'s \), i.e., \( g^{(p)}[0] = \tilde{v}[0] \), and \( g^{(q)}[0] = \tilde{v}[0] \) (following the suggestion in the discussion after (43)).

In Fig. 4, we compare the performance of algorithm (70)–(73) against that of the asynchronous ADMM proposed in [8]. The asynchronous ADMM has two parameters \( \rho \) and \( \alpha \). Figure 4 shows the trajectory of the relative cost error of the obtained solutions, namely,

\[ \frac{|f(g^{(p)}[k]) - f(g^{(p)*})|}{f(g^{(p)*})}, \]

and the evolution of the largest constraint violation for different values of \( \rho \) and \( \alpha \). In algorithm (70)–(73), each node runs \( m = 10 \) iterations, where we recall that \( m \) is the number of iterations between consecutive communication attempts. In ADMM, each node solves a local optimization problem (x- or z-update) at each iteration (between consecutive communication attempts). In general, the closed-form solutions of the local problems are not available, but, the local problems arising in the OPF problem (10) admit the closed-form solutions as shown in (7). The results in Fig. 4 demonstrate that algorithm (70)–(73) has geometric convergence rate, and might converge faster than the asynchronous ADMM since the latter has asymptotic convergence rate.

### V. CONCLUSION

We presented distributed algorithms for solving the OPF problem for radial distribution systems over time-varying communication networks. The algorithms have geometric convergence rate and resiliency to communication delays and random data packet losses. One interesting future direction is to extend the proposed algorithms to solve multi-period OPF problems with battery energy storage systems.

### REFERENCES

[1] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, “Semidefinite programming for optimal power flow problems,” *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6, pp. 383–392, 2008.

[2] M. Farivar and S. H. Low, “Branch flow model: Relaxations and convexification—part 1,” *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2554–2564, Aug. 2013.
[3] L. Gan, N. Li, U. Topcu, and S. H. Low, “Exact convex relaxation of optimal power flow in radial networks,” IEEE Transactions on Automatic Control, vol. 60, no. 1, pp. 72–87, Jan. 2015.

[4] E. Dall’Anese, H. Zhu, and G. B. Giannakis, “Distributed optimal power flow for smart microgrids,” IEEE Transactions on Smart Grid, vol. 4, no. 3, pp. 1464–1475, Sep. 2013.

[5] P. Sule, S. Backhaus, and M. Chertkov, “Optimal distributed control of reactive power via the alternating direction method of multipliers,” IEEE Trans. Energy Conversion, vol. 29, no. 4, pp. 968–977, Dec. 2014.

[6] S. Magnnusson, P. C. Weeraddana, and C. Fischione, “A distributed approach for the optimal power-flow problem based on admm and sequential convex approximations,” IEEE Transactions on Control of Network Systems, vol. 2, no. 3, pp. 238–253, Sep. 2015.

[7] Q. Peng and S. H. Low, “Distributed optimal power flow algorithm for radial networks, i: Balanced single phase case,” IEEE Transactions on Smart Grid, vol. 9, no. 1, pp. 111–121, Jan. 2018.

[8] J. Guo, G. Hug, and O. Tonguz, “Asynchronous ADMM for Distributed Non-Convex Optimization in Power Systems,” arXiv e-prints, p. arXiv:1710.08938, Oct. 2017.

[9] Z. Zhang and M. Y. Chow, “Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid,” IEEE Transactions on Power Systems, vol. 27, no. 4, pp. 1761–1768, Nov. 2012.

[10] A. D. Domínguez-García, S. T. Cady, and C. N. Hadjicostis, “Decentralized optimal dispatch of distributed energy resources,” in Proc. IEEE Conf. Decision and Control, Dec. 2012, pp. 3688–3693.

[11] S. Yang, S. Tan, and J. Xu, “Consensus based approach for economic dispatch problem in a smart grid,” IEEE Transactions on Power Systems, vol. 28, no. 4, pp. 4416–4426, Nov. 2013.

[12] S. Kar and G. Hug, “Distributed robust economic dispatch in power systems: A consensus + innovations approach,” in Proc. IEEE Power and Energy Soc. Gen. Meeting, July 2012, pp. 1–8.

[13] X. Zhang and A. Papachristodoulou, “Redesigning generation control in power systems: Methodology, stability and delay robustness,” in Proc. IEEE Conf. Decision and Control, Dec. 2014, pp. 953–958.

[14] S. T. Cady, A. D. Domínguez-García, and C. N. Hadjicostis, “A distributed generation control architecture for islanded ac microgrids,” IEEE Transactions on Control Systems Technology, vol. 23, no. 5, pp. 1717–1735, Sept. 2015.

[15] A. Cherukuri and J. Cortés, “Distributed generator coordination for initialization and anytime optimization in economic dispatch,” IEEE Trans. Control Network Sys., vol. 2, no. 3, pp. 226–237, Sep. 2015.

[16] J. Wu, T. Yang, D. Wu, K. Kalsi, and K. H. Johansson, “Distributed optimal dispatch of distributed energy resources over lossy communication networks,” IEEE Transactions on Smart Grid, vol. 8, no. 6, pp. 3125–3137, Nov. 2017.

[17] J. Mohammadi, G. Hug, and S. Kar, “Role of communication in the convergence rate of fully distributed dc optimal power flow,” in IEEE Int. Conf. on Smart Grid Commun., Nov. 2014, pp. 43–48.

[18] ——, “Fully distributed dc-opf approach for power flow control,” in Proc. IEEE Power Energy Society General Meeting, July 2015, pp. 1–5.

[19] M. Zholbaryssov, D. Fooladivanda, and A. D. Domínguez-García, “Resilient distributed optimal generation dispatch for lossy ac microgrids,” Systems & Control Letters, vol. 123, pp. 47 – 54, 2019.

[20] H. Khalil, Nonlinear Systems, 3rd ed. Upper Saddle River, N.J.: Prentice Hall, 2002.

[21] A. Nedić, A. Olshevsky, and W. Shi, “Achieving geometric convergence for distributed optimization over time-varying graphs,” SIAM Journal on Optimization, vol. 27, no. 4, pp. 2597–2633, 2017.

[22] M. Baran and F. F. Wu, “Optimal sizing of capacitors placed on a radial distribution system,” IEEE Transactions on Power Delivery, vol. 4, no. 4, pp. 735–743, Jan. 1989.

[23] A. Nedić and A. Ozdaglar, “Distributed subgradient methods for multiagent optimization,” IEEE Transactions on Automatic Control, vol. 54, no. 1, pp. 48–61, Jan. 2009.

[24] D. P. Bertsekas, Nonlinear Programming, 2nd ed. Athena Scientific, 1999.

[25] M. Zholbaryssov and A. D. Domínguez-García, “Microgrid distributed frequency control over time-varying communication networks,” in Proc. IEEE Conf. Decision and Control, Dec. 2018, pp. 5722–5727.

[26] ——, “Convex relaxations of the network flow problem under cycle constraints,” IEEE Trans. Control Network Syst., to appear.

[27] W. Rudin, Principles of Mathematical Analysis, 3rd ed. McGraw-Hill, 1976.

[28] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, “Matpower: Steady-state operations, planning, and analysis tools for power systems research and education,” IEEE Transactions on Power Systems, vol. 26, no. 1, pp. 12–19, Feb. 2011.