Resilient Distributed Optimal Generation Dispatch for Lossy AC Microgrids☆

Madi Zholbaryssov[a], Dariush Fooladivanda[a], Alejandro D. Domínguez-García[a]

[a]*Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

Abstract

In this paper, we consider the problem of designing a distributed algorithm, resilient against communication delays and packet drops, for coordinating the response of distributed energy resources (DERs) in AC microgrids so as to minimize electrical line losses and generation cost, and to ensure that the microgrid network constraints are satisfied. The proposed algorithm can be utilized as a secondary or tertiary controller for frequency regulation in islanded microgrids, or to coordinate DERs for providing frequency regulation services to the bulk grid when microgrid is in grid-connected mode. We validate the practical usefulness of the theoretical results through numerical simulations involving the IEEE 39–bus system.

Keywords: distributed algorithms, economic dispatch, AC microgrids, distributed energy resources

1. Introduction

It is envisioned that future electric power distribution systems will consist of several interconnected microgrids that operate in either grid-connected mode or islanded mode [1–5]. The electrical boundaries of such microgrids need to be precisely defined in order to ensure that each of them can act as a separate controllable entity with respect to the bulk grid [5]. In grid-connected mode, microgrids can optimize its operation costs by providing different types of services (e.g., frequency regulation services) to the bulk grid [1]. In islanded mode, microgrids only supply electrical power to loads using the DERs within their footprints [2]. In this paper, we focus on controlling the DERs within a microgrid when the microgrid is operating in islanded mode, and the electrical network is lossy; however, the proposed frameworks and controllers are applicable to grid-connected microgrids. To this end, we propose a distributed algorithm that is resilient against packet drops in the communication channels of the cyber network over which this algorithm is executed.

To address the problem of coordinating DERs in grid-connected and islanded microgrids under communication delays, several energy management frameworks and distributed control strategies have been recently proposed. However, most works (see, e.g., [6–19]), which take communication delays into account, neglect the network constraints, i.e., constraints on the power flow along the electrical lines, and only focus on the satisfaction of the total power demand. Although some works (see, e.g., [8, 10]) take the network constraints into consideration, they neglect the line losses. Therefore, under the existing control strategies, either the security constraints may be violated, or the system performance may not necessarily be optimal when the line losses are substantial.

In this work, we leverage multi-agent optimization techniques also applied in [7, 19] for solving the economic dispatch problem. The authors in [7] use the so-called *ratio-consensus algorithm* to determine the optimal power outputs of DERs, which minimize a quadratic cost function and satisfy the DER capacity constraints. To solve a similar problem, the authors in [19] use the so-called *subgradient-push algorithm* proposed in [20] to design an algorithm robust against communication delays and packet drops.

To compute the optimal set-points, we consider the network losses as our objective function, and formulate the problem of coordinating DERs for minimizing the losses and meeting the power demand while respecting the network constraints (we refer to this problem as **DOD1**). To solve this problem in a distributed way, we follow a three-step procedure. In the first step, we transform the constrained optimization problem **DOD1** into an unconstrained optimization problem (we call it **DOD3**) by adding some penalty terms to the objective function in **DOD1**. We then show that these penalty functions are exact, i.e., **DOD3** is equivalent to **DOD1**. In the second step, we propose a distributed algorithm for solving **DOD3**, and prove that the proposed algorithm is robust to communication delays and packet drops. Finally, to improve the convergence speed of the proposed distributed algorithm, we develop an adaptive stepsize selection strategy utilized by the algorithm.

2. Preliminaries

In this section, we present the microgrid model adopted in this work, and formulate the control objectives to be achieved.

☆This work is supported by the U.S. Department of Energy (DOE) within the GEARED Initiative, under grant DE0006341; the Advanced Research Projects Agency-Energy (ARPA-E) within the NODES program, under Award DEAR0000695; and by the DOE within the Consortium for Electric Reliability Technology Solutions (CERTS) program.

Email addresses: zholbar1@illinois.edu (M. Zholbaryssov), dfooladi@illinois.edu (D. Fooladivanda), aledan@illinois.edu (A. Domínguez-García)

2.1. Physical Layer

Consider a collection of DERs and loads interconnected by a lossy power network. The topology of this network can be described by an undirected graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $\mathcal{V} = \mathcal{V}^{(G)} \cup \mathcal{V}^{(L)}$, where $\mathcal{V}^{(G)} = \{1, \ldots, m\}$ denotes the set of nodes with a DER, and $\mathcal{V}^{(L)} = \{m+1, \ldots, n\}$ denotes the set of nodes with a load; and where $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, with $\{i, j\} \in \mathcal{E}$ if nodes i and j are electrically connected. Let $\mathcal{N}_i$ denote the set of neighbors of a node i, i.e., $\mathcal{N}_i = \{j \in \mathcal{V} : \{i, j\} \in \mathcal{E}\}$.

Define a one-to-one map, $\mathbb{I} : \mathcal{E} \to \mathbb{R}$, such that every e in the set $\{1, 2, \ldots, \ell\}$ is arbitrarily assigned to exactly one edge $\{l, j\} \in \mathcal{E}$, i.e., $\mathbb{I}(\{l, j\}) = e$. We also arbitrarily assign an orientation to each edge in $\mathcal{E}$, and denote by (l, j) the edge oriented from l to j, i.e., (l, j) is incident to j. Let $\vec{\mathcal{E}}$ denote the set of oriented edges that result from this arbitrarily chosen orientation (note that $|\vec{\mathcal{E}}| = |\mathcal{E}| =: \ell$). Let $\mathcal{N}_i^+$ and $\mathcal{N}_i^-$ respectively denote the sets of out-neighbors and in-neighbors of node i, i.e., $\mathcal{N}_i^+ = \{j \in \mathcal{V} : (i, j) \in \vec{\mathcal{E}}\}$ and $\mathcal{N}_i^- = \{j \in \mathcal{V} : (j, i) \in \vec{\mathcal{E}}\}$. Let δ denote the diameter of graph $\mathcal{G}$, which is the largest distance between any pair of vertices.

Assuming balanced three-phase operation, let $\theta_i(t)$ and $V_i(t)$ denote the phase angle and magnitude of the instantaneous voltage of a single phase at node $i \in \mathcal{V}$ in a rotating coordinate frame with rotational frequency ω^*, and define $\theta(t) = [\theta_1(t), \ldots, \theta_n(t)]^\top$ and $V(t) = [V_1(t), \ldots, V_n(t)]^\top$. Then, the net active power, P_i, injected into the network via node $i \in \mathcal{V}$ is given by

$$\begin{aligned} P_i(t) = \sum_{j \in \mathcal{V}} & V_i(t)V_j(t)[G_{ij}\cos(\theta_i(t) - \theta_j(t)) \\ & + B_{ij}\sin(\theta_i(t) - \theta_j(t))], \end{aligned} \tag{1}$$

with $G_{ii} = -\sum_{j \in \mathcal{V}} G_{ij}$, $B_{ii} = -\sum_{j \in \mathcal{V}} B_{ij}$, $G_{ij} = -g_{ij}$ and $B_{ij} = -b_{ij}$, where $g_{ij} > 0$ and $b_{ij} < 0$ are, respectively, the series conductance and series susceptance of the line connecting nodes i and j.

Here, we assume that voltage magnitudes are fixed at all buses, which is a standard assumption to make for solving the security-constrained economic dispatch problem [21]; thus, $V_i(t) = V_i^0$, where V_i^0 is some nominal voltage at node i. For brevity, we also drop the time t argument from the functions that depend on it.

To quantify line power loss, we assume that there is a certain gain associated with each line, which determines the fraction of active power sent from one end of the line, which reaches the other end; this model was used in the context of the generalized network flow problem (see, e.g., [22]), and more recently in the context of power system applications [23]. More formally, let ϕ_{ij} denote the active power flowing along line $(i, j) \in \vec{\mathcal{E}}$, which is given by

$$\phi_{ij} := (V_i^0)^2 g_{ij} + V_i^0 V_j^0 G_{ij}\cos(\theta_i - \theta_j) + V_i^0 V_j^0 B_{ij}\sin(\theta_i - \theta_j). \tag{2}$$

Also, define

$$\phi_{ji} := (V_j^0)^2 g_{ij} + V_i^0 V_j^0 G_{ij}\cos(\theta_j - \theta_i) + V_i^0 V_j^0 B_{ij}\sin(\theta_j - \theta_i). \tag{3}$$

When phase angles satisfy $|\theta_i - \theta_j| < -\tan^{-1}(B_{ij}/G_{ij})$ (which typically holds in practice), ϕ_{ij} and ϕ_{ji} are related through a scalar monotonically decreasing convex function $g_{ij}(\cdot)$ as follows [23]:

$$\phi_{ji} = g_{ij}(\phi_{ij}), \ (i, j) \in \vec{\mathcal{E}}. \tag{4}$$

Let θ^0 denote a solution of the power flow equations in (1) for the current values of active power injections denoted by P^0. We partially linearize the model in (1) and (3) such that ϕ_{ij} and ϕ_{ji} are related linearly around the operating point θ^0 as follows:

$$\phi_{ji} = -\alpha_{ij}\phi_{ij}, \ (i, j) \in \vec{\mathcal{E}}, \tag{5}$$

with $\alpha_{ij} := -\phi_{ji}^0/\phi_{ij}^0$, where the active power flow ϕ_{ij}^0 along line $(i, j) \in \vec{\mathcal{E}}$ corresponds to (θ^0, V^0). It is assumed that $\alpha_{ij} \leq 1$, which is the case if the active power flowing from node i to j is positive. In the network flow literature (see, e.g., [22]), α_{ij} is referred to as the gain of line $(i, j) \in \vec{\mathcal{E}}$, and determines the fraction of active power sent from node i, which reaches node j. Note that if $\alpha_{ij} = 1$, $\forall (i, j) \in \vec{\mathcal{E}}$, then, the microgrid network is lossless. We use α_{ij} to define a weighted node-to-edge incidence matrix, $M \in \mathbb{R}^{n \times \ell}$, as follows. For each $e = \mathbb{I}(\{i, j\})$, $M_{ie} = 1$ and $M_{je} = -\alpha_{ij}$, if $(i, j) \in \vec{\mathcal{E}}$, and $M_{ie} = 0$ and $M_{je} = 0$, if $\{i, j\} \notin \mathcal{E}$. Then, $P = M\phi$, where $\phi = [\{\phi_{ij}\}_{(i,j) \in \vec{\mathcal{E}}}]^\top$.

We also define the *generation loss factor* between nodes i and j given by $\beta_{ij} = \Delta P_j / \Delta P_i$, where ΔP_i is the required amount of change in generation at node i to satisfy the load change at node j, ΔP_j, provided that injections at all other nodes are fixed. In the remainder, we use a lower bound for β_{ij} given by $\underline{\alpha}^\delta$, where $\underline{\alpha} := \min_{(i,j) \in \vec{\mathcal{E}}} \alpha_{ij}$, which can be derived using the fact that β_{ij} depends on the gains of the lines connecting i and j. Define the minimum generation loss factor given by

$$\beta := \min_{i,j \in \mathcal{V}} \beta_{ij} \geq \underline{\alpha}^\delta. \tag{6}$$

2.2. Security-Constrained Economic Dispatch

In this paper, our goal is to design a distributed algorithm, robust to communication delays, that solves the following optimization problem for coordinating DERs:

$$\textbf{DOD1}: \min_{P \in \mathbb{R}^n, \phi \in \mathbb{R}^\ell} f(P) \tag{7}$$

$$\text{subject to } P = M\phi, \tag{8}$$

$$\underline{P} \leq P \leq \overline{P}, \tag{9}$$

$$\underline{\phi} \leq \phi \leq \overline{\phi}, \tag{10}$$

with $f(P) := \sum_{i \in \mathcal{V}^{(G)}} f_i(P_i)$, where $f_i(\cdot)$ is a non-negative convex cost function, $\underline{P} = [\underline{P}_1, \ldots, \underline{P}_n]^\top$, $\overline{P} = [\overline{P}_1, \ldots, \overline{P}_n]^\top$, $\underline{\phi} = [\{\underline{\phi}_{ij}\}_{(i,j) \in \vec{\mathcal{E}}}]^\top$, $\overline{\phi} = [\{\overline{\phi}_{ij}\}_{(i,j) \in \vec{\mathcal{E}}}]^\top$, $\overline{P}_i$ and $\underline{P}_i$ are the upper and lower capacity limits for DER i active power output, P_i, $i \in \mathcal{V}^{(G)}$, $\overline{P}_l = \underline{P}_l$, $l \in \mathcal{V}^{(L)}$, $\overline{\phi}_{ij}$ and $\underline{\phi}_{ij}$ are the upper and lower bounds for power flow along line $(i, j) \in \vec{\mathcal{E}}$, prescribed by the thermal limit considerations. Here, (8) represents the power balance constraint, whereas the box constraints in (10) capture the thermal limit constraints on the electrical lines. We note that since the total active power generation $\sum_{i\in\mathcal{V}^{(G)}} P_i$ is equal to the total line losses plus the load demand; thus, by minimizing the power generation cost, we also optimize the losses.

We assume that $f(\cdot)$ is Lipschitz on every compact set, in particular, $|f(x) - f(y)| \leq L\|x - y\|_2$, with some constant L, $\forall x, y \in \mathcal{B} := \{P \in \mathbb{R}^n : \underline{P}_i \leq P_i \leq \overline{P}_i,\ i \in \mathcal{V}\}$. Finally, $d(z)$ is a subgradient of $f(x)$ evaluated at $x = z \in \mathbb{R}^n$, when the following relation holds: $f(x) \geq f(z) + d(z)^\top (x - z)$.

2.3. Cyber Layer

The implementation of the distributed algorithms for solving DOD1 will rely on a time-varying bidirectional communication network described by a directed graph, $\mathcal{G}_k^{(c)} = (\mathcal{V}, \mathcal{E}_k^{(c)})$, where $(i, j) \in \mathcal{E}_k^{(c)}$ if node j receives information from node i during the sampling period (t_k, t_{k+1}). Here, we assume that both nodes i and j are able to transmit information to each other, but the information exchange does not necessarily happen simultaneously within the same sampling period, i.e., even if $(i, j) \in \mathcal{E}_k^{(c)}$, $(j, i) \in \mathcal{E}_k^{(c)}$ may not hold. We also let $(i, j) \in \bigcup_{l=k+1}^{k+B} \mathcal{E}_l^{(c)}$, $\forall k$, if $\{i, j\} \in \mathcal{E}$, which is equivalent to having node j receive information from node i at least once every B iterations, where B is some bounded constant; this connectivity assumption is standard in the literature [24].

3. Distributed Optimal Dispatch Over Time-Varying Communication Networks

In this section, we present a distributed algorithm for solving DOD1, which is robust to communication delays. This algorithm is designed after transforming the constrained DOD1 into an unconstrained optimization problem by using the exact penalty approach.

3.1. Exact Penalty Approach

Here, we add penalty terms to the cost in DOD1 capturing violations in the constraints of DOD1. We show that these penalty functions are exact, i.e., the unconstrained optimization problem, resulting from replacing all constraints in DOD1 with the penalty functions, is equivalent to DOD1. Define

$$b_i(P_i, \phi) := P_i - \sum_{j\in\mathcal{N}_i^+} \phi_{ij} + \sum_{l\in\mathcal{N}_i^-} \alpha_{li}\phi_{li}, \tag{11}$$

which we refer to as the *flow balance* of node i; by defining $b(P, \phi) = [b_1(P_1, \phi), b_2(P_2, \phi), \ldots, b_n(P_n, \phi)]^\top$, (11) can be written in vector form as $b(P, \phi) = P - M\phi$. Consider an optimization program obtained from DOD1 by replacing the equality constraint in (8) with a non-smooth penalty term of the form $\gamma\|b(P, \phi)\|_1 = \gamma \sum_{i\in\mathcal{V}} |b_i(P_i, \phi)|$, where $\gamma > 0$; this results in

$$\textbf{DOD2}: \min_{P\in\mathbb{R}^n, \phi\in\mathbb{R}^\ell} f(P) + \gamma\|b(P, \phi)\|_1 \tag{12}$$

$$\text{subject to } \underline{P} \leq P \leq \overline{P}, \tag{13}$$

$$\underline{\phi} \leq \phi \leq \overline{\phi}. \tag{14}$$

The next proposition shows that this penalty function is exact for sufficiently large γ, i.e., DOD2 is equivalent to DOD1.

Proposition 1. Let $\gamma > L/\beta$, where L is a Lipschitz constant, and β is given in (6). Then, (P^*, ϕ^*) is a minimum of DOD1 if it is a minimum of DOD2.

Proof. Assuming that (P^*, ϕ^*) is a minimum of DOD2, we need to show that it is a minimum of DOD1. Let $\mathcal{X}_1$ and $\mathcal{X}_2$ denote the constraint sets of DOD1 and DOD2, i.e.,

$$\mathcal{X}_1 = \{(P, \phi) \in \mathbb{R}^n \times \mathbb{R}^\ell : (P, \phi) \text{ satisfies (8) – (10)}\},$$
$$\mathcal{X}_2 = \{(P, \phi) \in \mathbb{R}^n \times \mathbb{R}^\ell : (P, \phi) \text{ satisfies (13) – (14)}\}.$$

First, note that $\mathcal{X}_1 \subset \mathcal{X}_2$. Let $\hat{x} := (\hat{P}, \hat{\phi})$ denote the projection of $x := (P, \phi) \in \mathcal{X}_2$ onto $\mathcal{X}_1$. We first show that, for any $(P, \phi) \in \mathcal{X}_2$, $\|P - \hat{P}\|_1 \leq \|b(P, \phi)\|_1/\beta$. Note that $b(P, \phi)$ can be viewed as a load change balanced with the change in generation, $\hat{P} - P$. Since $b_i(P, \phi)$'s can be of different signs, the total load change is not greater than $\|b(P, \phi)\|_1$. It follows from the definition of β_{ij} and the fact that $\beta_{ij} \geq \beta$ that the total change in generation is not greater than $\|b(P, \phi)\|_1/\beta$. Then, for any $x \notin \mathcal{X}_1$,

$$\begin{aligned} f(P) + \gamma\|b(P, \phi)\|_1 &\geq f(P) + \gamma\beta\|P - \hat{P}\|_2 = f(\hat{P}) + f(P) \\ &\quad - f(\hat{P}) + \gamma\beta\|P - \hat{P}\|_2 > f(\hat{P}) + f(P) \\ &\quad - f(\hat{P}) + L\|P - \hat{P}\|_2 > f(\hat{P}), \end{aligned}$$

where, to arrive at the last inequality, we used the fact that $\gamma > L/\beta$, and also the fact that $f(\cdot)$ is Lipschitz continuous over $\mathcal{B}$ with constant L. Because $b(\hat{P}, \hat{\phi}) = 0$, $f(\hat{P})$ is the DOD2 cost at $(\hat{P}, \hat{\phi})$, which is lower than that at (P, ϕ). Then, it follows that a minimum of DOD2 belongs to $\mathcal{X}_1$ and satisfies the balance constraints in (8). Therefore, if (P^*, ϕ^*) is a minimum of DOD2, then, it is also a minimum of DOD1. □

Now, we formulate an unconstrained optimization problem, referred to as DOD3, which is equivalent to DOD2. To this end, we augment the cost of DOD2 with additional penalty terms for violating the box constraints in (13) – (14); this results in

$$\begin{aligned} \textbf{DOD3}: \min_{P\in\mathbb{R}^n, \phi\in\mathbb{R}^\ell} & f(P) + \gamma\|b(P, \phi)\|_1 + \mu \sum_{i\in\mathcal{V}} \Big([P_i - \overline{P}_i]^+ \\ & + [\underline{P}_i - P_i]^+\Big) + \lambda \sum_{(i,j)\in\vec{\mathcal{E}}} \Big([\phi_{ij} - \overline{\phi}_{ij}]^+ + [\underline{\phi}_{ij} - \phi_{ij}]^+\Big), \end{aligned}$$

where $[x]^+ = \max(x, 0)$, μ and λ are positive constants. If μ and λ are large enough, the penalty functions in DOD3 are exact as established next.

Proposition 2. Define $\mathcal{B}_i := \{x \in \mathbb{R} : \underline{P}_i - \nu_i \leq x \leq \overline{P}_i + \nu_i\}$, $i \in \mathcal{V}^{(G)}$, for some $\nu_i > 0$, so that

$$f_i(\overline{P}_i + \nu_i) > f(\overline{P}),\ f_i(\underline{P}_i - \nu_i) > f(\overline{P}),\ \forall i \in \mathcal{V}^{(G)}. \tag{15}$$

Let $f_i(\cdot)$ be Lipschitz on $\mathcal{B}_i$ with constant L_ν, $\forall i \in \mathcal{V}^{(G)}$, that is, $|f_i(x) - f_i(y)| \leq L_\nu|x - y|$, $\forall i \in \mathcal{V}^{(G)}$, $x, y \in \mathcal{B}_i$. Let $\gamma > L/\beta$, $\lambda > \gamma$, and $\mu > \gamma + L_\nu$. Then, (P^*, ϕ^*) is a minimum of DOD1 if it is a minimum of DOD3.

Proof. Since, by Proposition 1, DOD1 is equivalent to DOD2, it is enough to show that DOD2 is equivalent to DOD3. Thus, assuming that (P^*, ϕ^*) is a minimum of DOD3, we need to show that it is a minimum of DOD2. We recall that $\mathcal{X}_2$ denotes the constraint set of DOD2, i.e.,

$$\mathcal{X}_2 = \{(P, \phi) \in \mathbb{R}^n \times \mathbb{R}^\ell : (P, \phi) \text{ satisfies (13) – (14)}\},$$

and let $(\hat{P}, \hat{\phi})$ denote the projection of (P^*, ϕ^*) onto $\mathcal{X}_2$. For brevity, define $x^* = (P^*, \phi^*)$ and $\hat{x} = (\hat{P}, \hat{\phi})$. Since $\|b(x)\|_1 := \|b(P, \phi)\|_1$ is convex,

$$\|b(x^*)\|_1 \geq \|b(\hat{x})\|_1 + d(\hat{x})^\top (x^* - \hat{x}), \tag{16}$$

where $d(\hat{x})$ is a subgradient of $\|b(x)\|_1$ at $x = \hat{x}$. Since $\|d(x)\|_\infty \leq 1, \forall x$, it follows from (16) that

$$\begin{aligned}\|b(\hat{x})\|_1 - \|b(x^*)\|_1 &\leq d(\hat{x})^\top (\hat{x} - x^*) \leq \|\hat{x} - x^*\|_1 \\ &= \sum_{i \in \mathcal{V}} \left([P_i^* - \overline{P}_i]^+ + [\underline{P}_i - P_i^*]^+\right) \\ &\quad + \sum_{(i,j) \in \vec{\mathcal{E}}} \left([\phi_{ij}^* - \overline{\phi}_{ij}]^+ + [\underline{\phi}_{ij} - \phi_{ij}^*]^+\right).\end{aligned}$$

We recall that DOD3 is unconstrained and (P^*, ϕ^*) is a minimum of DOD3. Then, if (P^*, ϕ^*) satisfies the box constraints of DOD2 in (13) – (14), then, (P^*, ϕ^*) is a minimum of DOD2 because the costs of DOD2 and DOD3 are equal at any point of the constraint set of DOD2. By assuming that (P^*, ϕ^*) does not satisfy the box constraints of DOD2 in (13) – (14), one can show the existence of $(\hat{P}, \hat{\phi})$ that yields a smaller cost, which contradicts the fact that (P^*, ϕ^*) is a minimum of DOD3. □

3.2. Resilient Distributed Optimal Dispatch

Previously, we introduced the unconstrained optimization problem DOD3, which is equivalent to the original problem DOD1. Next, we solve DOD3 using a distributed algorithm, which is robust to communication delays. Let

$$\begin{aligned}F_i(P, \phi) :=\, & f_i(P_i) + \gamma |b_i(P_i, \phi)| + \mu([P_i - \overline{P}_i]^+ + [\underline{P}_i - P_i]^+) \\ & + \lambda \sum_{j \in \mathcal{N}_i^+} ([\phi_{ij} - \overline{\phi}_{ij}]^+ + [\underline{\phi}_{ij} - \phi_{ij}]^+) \\ & + \lambda \sum_{j \in \mathcal{N}_i^-} ([\phi_{ji} - \overline{\phi}_{ji}]^+ + [\underline{\phi}_{ji} - \phi_{ji}]^+)\end{aligned}$$

denote the local cost function at node $i \in \mathcal{V}$, where $f_l(P_l) \equiv 0$ for any load $l \in \mathcal{V}^{(L)}$. Then, by Proposition 2, clearly,

$$\textbf{DOD}: \min_{P \in \mathbb{R}^n, \phi \in \mathbb{R}^\ell} \sum_{i=1}^n F_i(P, \phi), \tag{17}$$

is equivalent to DOD3. DOD can be viewed as an unconstrained multi-agent optimization problem that can be effectively solved using the following algorithm (see, e.g., [24]):

$$x^{(i)}[k+1] = \sum_{j=1}^n a_j^i[k] x^{(j)}[k] - s[k] d_i[k], \tag{18}$$

where $x^{(i)}[k] \in \mathbb{R}^{m\ell}$ is an estimate of a minimum of DOD, (P^*, ϕ^*), maintained at node i, $d_i[k]$ is a subgradient of the cost function $F_i(P, \phi)$ evaluated at $x^{(i)}[k]$, $a^i[k] = [a_1^i[k], \ldots, a_m^i[k]]^\top$ is a vector of some weights, and $s[k]$ is a stepsize. Later, we propose a strategy for tuning the stepsize adaptively based on the maximum flow imbalance of the network, namely, $\max_{i \in \mathcal{V}} |b_i(P_i, \phi)|$. As demonstrated in the numerical simulations, such adaptive stepsize allows to significantly improve convergence speed.

It is worth to note that, in order to solve DOD, each node does not have to estimate all entries of (P^*, ϕ^*), but rather those components of (P^*, ϕ^*) that affect its local cost, that is, node i only needs to estimate P_i^*, ϕ_{ij}^*, $(i, j) \in \vec{\mathcal{E}}$, and ϕ_{li}^*, $(l, i) \in \vec{\mathcal{E}}$. In this case, ϕ_{ij}^*, $(i, j) \in \vec{\mathcal{E}}$, is estimated only by nodes i and j. Then, the algorithm in (18) can be simplified so that each node i executes the following iterations:

$$P_i[k+1] = P_i[k] - s[k] g_i^{(i)}[k], \tag{19}$$

$$\phi_{ij}^{(i)}[k+1] = a_{ij}^{(i)}[k]\phi_{ij}^{(i)}[k] + a_{ij}^{(j)}[k]\phi_{ij}^{(j)}[k] - s[k] g_{ij}^{(i)}[k], \tag{20}$$

$$\phi_{li}^{(i)}[k+1] = a_{il}^{(i)}[k]\phi_{li}^{(i)}[k] + a_{il}^{(l)}[k]\phi_{li}^{(l)}[k] - s[k] g_{li}^{(i)}[k], \tag{21}$$

where $P_i[k]$, $\phi_{ij}^{(i)}[k]$, $(i, j) \in \vec{\mathcal{E}}$, and $\phi_{li}^{(i)}[k]$, $(l, i) \in \vec{\mathcal{E}}$, are the estimates of P_i^*, ϕ_{ij}^*, and ϕ_{li}^*, respectively, maintained at node i,

$$a_{pq}^{(q)}[k] = \begin{cases} 1/2 & \text{if } (q, p) \in \mathcal{E}_k^{(c)}, \\ 0 & \text{else,} \end{cases}$$

$a_{pq}^{(p)}[k] = 1 - a_{pq}^{(q)}[k], \forall \{p, q\} \in \mathcal{E}$, $g_i^{(i)}$, $g_{ij}^{(i)}$, and $g_{li}^{(i)}$ are the subgradients of $F_i(P, \phi)$ with respect to P_i, ϕ_{ij} and ϕ_{li}, respectively, given by the following expressions:

$$\begin{aligned}g_i^{(i)}[k] &= \nabla f(P_i[k]) + \gamma \operatorname{sgn}(b_i[k]) + \mu[\eta(P_i[k], \overline{P}_i) \\ &\quad - \eta(\underline{P}_i, P_i[k])], \\ g_{ij}^{(i)}[k] &= -\gamma \operatorname{sgn}(b_i[k]) + \lambda[\eta(\phi_{ij}^{(i)}[k], \overline{\phi}_{ij}) - \eta(\underline{\phi}_{ij}, \phi_{ij}^{(i)}[k])], \\ g_{li}^{(i)}[k] &= \alpha_{li}\gamma \operatorname{sgn}(b_i[k]) + \lambda[\eta(\phi_{li}^{(i)}[k], \overline{\phi}_{li}) - \eta(\underline{\phi}_{li}, \phi_{li}^{(i)}[k])],\end{aligned}$$

where $\eta(x, y) = \operatorname{sgn}([x - y]^+)$, $\operatorname{sgn}(x) := x/|x|$, if $x \neq 0$, and 0, otherwise, and $b_i[k] := P_i[k] - \sum_{(i,j) \in \vec{\mathcal{E}}} \phi_{ij}^{(i)}[k] + \sum_{(l,i) \in \vec{\mathcal{E}}} \alpha_{li}\phi_{li}^{(i)}[k]$. Besides implementing a subgradient method, when $a_{ij}^{(j)} = 1/2$, (20) performs averaging of the flow estimates $\phi_{ij}^{(i)}[k]$ and $\phi_{ij}^{(j)}[k]$ to ensure that both estimates converge to the same value, ϕ_{ij}^*.

Alternatively, instead of using the penalty $\gamma|b_i(P_i, \phi)|$ in $F_i(P, \phi)$, we can use $\gamma[b_i(P_i, \phi)]^-$, where $[x]^- := -\min(x, 0)$, which, as we observed in the numerical simulations, allows to achieve a faster convergence to the optimal set-points. To explain this, we note that penalizing the flow balances ensures that the set-points meet the power demand and respect the power balance constraints. The penalty $\gamma|b_i(P_i, \phi)|$ prioritizes immediate satisfaction of the power balance constraints, rather than minimization of the generation cost (although generation cost is also minimized asymptotically as shown in Proposition 2). In contrast, the penalty $\gamma[b_i(P_i, \phi)]^-$ penalizes the power balance constraints only if $b_i(P_i, \phi) < 0$, and when $b_i(P_i, \phi) > 0$, the priority falls on the generation cost minimization.

Before we prove that using the penalty $\gamma[b_i(P_i,\phi)]^-$, instead of $\gamma|b_i(P_i,\phi)|$, in $F_i(P,\phi)$ allows to obtain a minimum of DOD, we make the following assumption:

Assumption 1. We assume that (a) $\sum_{i\in\mathcal{V}} \underline{P}_i < 0$; and (b) $\sum_{i\in\mathcal{V}_1} \underline{P}_i - \sum_{\{i,j\}\in\mathcal{E}_1} \overline{\phi}_{ij} < 0$, for any connected subgraph with the set of nodes $\mathcal{V}_1$, where $\mathcal{E}_1 := \{\{i,j\}\in\mathcal{E} : i\in\mathcal{V}_1, j\in\mathcal{V}\setminus\mathcal{V}_1\}$.

Proposition 3. (P^*,ϕ^*) is a minimum of DOD if (P^*,ϕ^*) is a minimum of DOD′ given below:

$$\textbf{DOD}' : \min_{P\in\mathbb{R}^n,\phi\in\mathbb{R}^\ell} \sum_{i=1}^{n} h_i(P,\phi), \tag{22}$$

where

$$\begin{aligned} h_i(P,\phi) :=& f_i(P_i) + \gamma[b_i(P_i,\phi)]^- + \mu([P_i-\overline{P}_i]^+ + [\underline{P}_i - P_i]^+) \\ &+ \lambda \sum_{j\in\mathcal{N}_i^-} ([\phi_{ij}-\overline{\phi}_{ij}]^+ + [\underline{\phi}_{ij}-\phi_{ij}]^+) \\ &+ \lambda \sum_{j\in\mathcal{N}_i^+} ([\phi_{ji}-\overline{\phi}_{ji}]^+ + [\underline{\phi}_{ji}-\phi_{ji}]^+). \end{aligned} \tag{23}$$

Proof. We first show that if there is an i such that $b_i(P_i^*,\phi^*) > 0$, then, there is $(P^{(1)},\phi^{(1)})$ such that

$$\sum_{i=1}^{n} h_i(P^{(1)},\phi^{(1)}) < \sum_{i=1}^{n} h_i(P^*,\phi^*), \tag{24}$$

which will contradict the fact that (P^*,ϕ^*) is a minimum of DOD′. Choose a maximal connected subgraph $\mathcal{S} = (\mathcal{V}^S,\mathcal{E}^S)$ such that every node i has $b_i(P_i^*,\phi^*) > 0$. If there exists a node i in $\mathcal{S}$ such that $P_i^* > \underline{P}_i$, then, choosing $P_i^{(1)} = P_i^* - \epsilon \geq \underline{P}_i$ with strictly positive $\epsilon\in(0,b_i(P_i^*,\phi^*)]$, $P_j^{(1)} = P_j^*$, $\forall j\neq i$, and $\phi^{(1)} = \phi^*$ yields (24). Therefore, $P_i^* = \underline{P}_i$, $\forall i\in\mathcal{V}^S$. If $\mathcal{S}$ contains all n nodes of $\mathcal{G}$, then, since $\phi_{ij}^* + \phi_{ji}^* > 0$, we have that

$$0 < \sum_{i\in\mathcal{V}^S} b_i(P_i^*,\phi^*) = \sum_{i\in\mathcal{V}^S}(\underline{P}_i - \sum_{j\in\mathcal{N}_i}\phi_{ij}^*) \leq \sum_{i\in\mathcal{V}^S}\underline{P}_i, \tag{25}$$

which contradicts Assumption 1(a). Thus, $\mathcal{S}$ does not contain all n nodes of $\mathcal{G}$. Denote $\mathcal{U} := \{\{i,j\}\in\mathcal{E} : i\in\mathcal{V}^S, j\in\mathcal{V}\setminus\mathcal{V}^S\}$, and $\mathcal{W} := \{\{i,j\}\in\mathcal{U} : \phi_{ij}^* < \overline{\phi}_{ij}\}$. For each $\{i,j\}\in\mathcal{W}$, choose the largest $\epsilon_{ij} > 0$ such that $\phi_{ij}^* + \epsilon_{ij} \leq \overline{\phi}_{ij}$ and $b_i(P_i^{(1)},\phi^{(1)}) \geq 0$, where $P^{(1)} = P^*$, $\phi_{ij}^{(1)} = \phi_{ij}^* + \epsilon_{ij}$, $\{i,j\}\in\mathcal{W}$, and $\phi_{ab}^{(1)} = \phi_{ab}^*$, $\{a,b\}\notin\mathcal{W}$. If there exists $\{i,j\}\in\mathcal{W}$ such that $b_j(P_j^*,\phi^*) < 0$, (24) follows from the fact that $b_j(P_j^{(1)},\phi^{(1)}) > b_j(P_j^*,\phi^*)$. Therefore, for all $\{i,j\}\in\mathcal{W}$, we have that $b_j(P_j^*,\phi^*) = 0$ and $b_j(P_j^{(1)},\phi^{(1)}) > 0$. If we redefine all variables including $\mathcal{S}$, $\mathcal{U}$, $\mathcal{W}$ and repeat the above analysis for $(P^{(1)},\phi^{(1)})$, we arrive at the conclusion that either $\mathcal{S}$ contains all n nodes of $\mathcal{G}$, which was shown to contradict Assumption 1(a), or $\phi_{ij}^* = \overline{\phi}_{ij}$ for all $\{i,j\}\in\mathcal{U}$. Now, we show that the latter statement contradicts Assumption 1(b). Indeed, by using the fact that $\phi_{ij}^* + \phi_{ji}^* > 0$, we have that

$$0 < \sum_{i\in\mathcal{V}^S} b_i(P_i^*,\phi^*) = \sum_{i\in\mathcal{V}^S}(\underline{P}_i - \sum_{j\in\mathcal{N}_i}\phi_{ij}^*) \leq \sum_{i\in\mathcal{V}^S}\underline{P}_i - \sum_{\{i,j\}\in\mathcal{U}}\overline{\phi}_{ij}),$$

which contradicts Assumption 1(b). Therefore, $b_i(P_i^*,\phi^*) \leq 0$, $\forall i$. Suppose (P^*,ϕ^*) is not a minimum of DOD, and let (P^0,ϕ^0) denote a minimum of DOD. Then, there is an i such that $b_i(P_i^*,\phi^*) < 0$. By using the same reasoning as in the proof of Proposition 2, it follows that $P^0\in\mathcal{B}$. Since $\gamma > L/\beta$, P^0 and $P^*\in\mathcal{B}$, and $f(P)$ is Lipschitz continuous over $\mathcal{B}$ with constant L, we have that $f(P^0) - f(P^*) < \gamma\beta\|P^0 - P^*\|_2 \leq \gamma\|b(P^*,\phi^*)\|_1 = \gamma\sum_{i=1}^{n}[b_i(P_i^*,\phi^*)]^-$, and, hence, $\sum_{i=1}^{n} h_i(P^0,\phi^0) < \sum_{i=1}^{n} h_i(P^*,\phi^*)$, which contradicts the fact that (P^*,ϕ^*) is a minimum of DOD′. □

To solve DOD′, each node i executes the following iterations:

$$P_i[k+1] = P_i[k] - s[k]d_i^{(i)}[k], \tag{26}$$

$$\phi_{ij}^{(i)}[k+1] = a_{ij}^{(i)}[k]\phi_{ij}^{(i)}[k] + a_{ij}^{(j)}[k]\phi_{ij}^{(j)}[k] - s[k]d_{ij}^{(i)}[k], \tag{27}$$

$$\phi_{li}^{(i)}[k+1] = a_{il}^{(i)}[k]\phi_{li}^{(i)}[k] + a_{il}^{(l)}[k]\phi_{li}^{(l)}[k] - s[k]d_{li}^{(i)}[k], \tag{28}$$

where $d_i^{(i)}$, $d_{ij}^{(i)}$, and $d_{li}^{(i)}$ are the subgradients of $h_i(P,\phi)$ with respect to P_i, ϕ_{ij}, $(i,j)\in\vec{\mathcal{E}}$, and ϕ_{li}, $(l,i)\in\vec{\mathcal{E}}$, respectively.

If the communication network is not time-varying, i.e., neighbors communicate at each iteration bidirectionally, then, the proposed algorithms in (19) – (21) and (26) – (28) can be written as $x[k+1] = x[k] - s[k]g[k]$, where $x[k] := [P[k]^\top,\phi[k]^\top]^\top$, $g[k]$ is the subgradient of $F[k] := \sum_{i=1}^{n} F_i(P[k],\phi[k])$ evaluated at $x[k]$. Then, by using a basic analysis, one can show that

$$F^{[k]} - F^* \leq \left(\|x[1]-x^*\|_2^2 + C^2\sum_{i=1}^{k} s^2[i]\right)/\left(2\sum_{i=1}^{k} s[i]\right), \tag{29}$$

where $C \geq \|g[k]\|_2$, $\forall k$. If $\sum_{i=1}^{\infty} s^2[i] < \infty$ and $\sum_{i=1}^{k} s[i] \to \infty$ as $k\to\infty$, it follows from (29) that the time it takes for the objective to be within ϵ of its optimal value scales linearly with the number of nodes in the network.

3.3. Asynchronous Information Exchange

The convergence analysis tools in [24, Proposition 3] can be leveraged to show that both algorithms in (19) – (21) and (26) – (28) converge to a minimum of DOD, if $\sum_{k=1}^{\infty} s^2[k] < \infty$ and $\sum_{k=1}^{\infty} s[k] = +\infty$. However, the proof of [24, Proposition 3] builds on the assumption that neighboring nodes exchange information at the same time. To overcome this limitation, we utilize a scheme in [25] that allows iterations in (19) – (21) and (26) – (28) to converge to the minimum of DOD without assuming simultaneous information exchange. The main idea behind this scheme is to let neighboring nodes perform averaging in alternating fashion, in which node i performs averaging of its estimate with the estimate of node j, and before it performs another averaging with node j, it waits for node j to perform averaging. More formally, in the scheme, for each $\{i, j\} \in \mathcal{E}$, node i maintains $\nu_{ij}^{(i)}[k]$ and $\nu_{ji}^{(i)}[k]$, and additional variables $\mu_{ij}^{(i)}[k]$ and $\mu_{ji}^{(j)}[k]$ if $(j, i) \in \vec{\mathcal{E}}$, and updates them as follows:

$$\nu_{ji}^{(i)}[k] = \begin{cases} \nu_{ji}^{(j)}[k] & \text{if } (j, i) \in \mathcal{E}_k^{(c)}, \\ \nu_{ji}^{(i)}[k-1] & \text{otherwise}, \end{cases} \tag{30}$$

$$\nu_{ij}^{(i)}[k] = \begin{cases} \nu_{ij}^{(i)}[k-1] & \text{if } \nu_{ji}^{(i)}[k] = \nu_{ji}^{(i)}[k-1], \\ 1 - \nu_{ij}^{(i)}[k-1] & \text{otherwise}, \end{cases} \tag{31}$$

and, if $(j, i) \in \vec{\mathcal{E}}$,

$$\mu_{ij}^{(i)}[k] = \begin{cases} \mu_{ij}^{(i)}[k-1] & \text{if } \nu_{ji}^{(i)}[k] = \nu_{ji}^{(i)}[k-1], \\ \phi_{ji}^{(i)}[k] & \text{otherwise}, \end{cases} \tag{32}$$

$$\mu_{ji}^{(j)}[k] = \begin{cases} \mu_{ji}^{(j)}[k-1] & \text{if } \nu_{ji}^{(i)}[k] = \nu_{ji}^{(i)}[k-1], \\ \phi_{ji}^{(j)}[k] & \text{otherwise}. \end{cases} \tag{33}$$

If $(j, i) \in \vec{\mathcal{E}}$, then, $\nu_{ij}^{(i)}[k]$, $\mu_{ij}^{(i)}[k]$ and $\mu_{ji}^{(j)}[k]$ comprise the information sent by node i to node j during time interval (t_k, t_{k+1}). If $(i, j) \in \vec{\mathcal{E}}$, $\nu_{ij}^{(i)}[k]$ and $z_{ij}[k]$ are sent by node i to node j. Assuming $(i, j) \in \vec{\mathcal{E}}$, the initial values of $\nu_{ij}^{(i)}[k]$, $\nu_{ji}^{(i)}[k]$, $\nu_{ji}^{(j)}[k]$, and $\nu_{ij}^{(j)}[k]$ are set to $\nu_{ij}^{(i)}[0] = 1$, $\nu_{ji}^{(i)}[0] = 0$, $\nu_{ji}^{(j)}[0] = 0$, and $\nu_{ij}^{(j)}[0] = 0$. Also, node j initializes $\mu_{ij}^{(i)}[0]$ and $\mu_{ji}^{(j)}[0]$ to zero. The reason for initializing $\nu_{ij}^{(i)}[0] = 1$ and other variables to zero is that, for every $(i, j) \in \vec{\mathcal{E}}$, i is chosen to be the node that initiates the scheme execution at $k = 0$. The algorithm in (26) – (28) is redefined as follows:

$$P_i[k+1] = P_i[k] - s[k] d_i^{(i)}[k], \tag{34}$$

$$\begin{aligned} \phi_{ij}^{(i)}[k+1] = {} & a_{ij}^{(i)}[k]\mu_{ij}^{(i)}[k] + a_{ij}^{(j)}[k]\mu_{ji}^{(j)}[k] + \phi_{ij}^{(i)}[k] - \mu_{ij}^{(i)}[k] \\ & - s[k] d_{ij}^{(i)}[k], \ \forall (i, j) \in \vec{\mathcal{E}}, \end{aligned} \tag{35}$$

$$\phi_{li}^{(i)}[k+1] = a_{il}^{(i)}[k]\phi_{li}^{(i)}[k] + a_{il}^{(l)}[k]\phi_{li}^{(l)}[k] - s[k] d_{li}^{(i)}[k], \ \forall (l, i) \in \vec{\mathcal{E}}, \tag{36}$$

where, $\forall \{p, q\} \in \mathcal{E}$, $a_{pq}^{(p)}[k] = 1 - a_{pq}^{(q)}[k]$, and

$$a_{pq}^{(q)}[k] = \begin{cases} 1/2 & \text{if } (q, p) \in \mathcal{E}_k^{(c)}, \ \nu_{qp}^{(p)}[k] \neq \nu_{qp}^{(p)}[k-1], \\ 0 & \text{else}. \end{cases} \tag{37}$$

The scheme (30) – (33) allows flow averaging between neighboring nodes to happen in alternating fashion. In particular, $\nu_{ij}^{(i)}[k]$ acts as an acknowledgement flag sent by node i to node j to indicate that information from node j has been received and used to perform flow averaging in (35). Node i initiates information exchange if $(i, j) \in \vec{\mathcal{E}}$ by sending the most recent value of $\phi_{ij}^{(i)}$. If acknowledgment flag sent by node i is different from its previous value, node j updates $\mu_{ji}^{(j)}$, which is set to its local flow estimate, $\phi_{ij}^{(j)}$, and also updates its acknowledgement flag, $\nu_{ji}^{(j)}$, and $\mu_{ij}^{(i)}$ and sends them to node i; node j also uses received $\phi_{ij}^{(i)}$ to carry out flow averaging and update its flow estimate $\phi_{ij}^{(j)}$. Once a new acknowledgement flag arrives from node j, node i uses $\mu_{ij}^{(i)}$ and $\mu_{ji}^{(j)}$ to perform averaging in (35) to update $\phi_{ij}^{(i)}$.

By following the steps in the proof of [24, Proposition 3], the algorithm in (34) – (36) based on the acknowledgement scheme in (30) – (33) converges to the minimum of DOD, if $\sum_{k=1}^{\infty} s^2[k] < \infty$ and $\sum_{k=1}^{\infty} s[k] = +\infty$. Regarding the proof of [24, Proposition 3], we note that it assumes that the subgradient of the cost function $h_i(P, \phi)$ is bounded; it can be easily shown that this boundedness condition for the subgradient holds for the proposed algorithm in (34) – (36). As a final remark, if $s[k]$ is constant, then, (34) – (36) converges to the minimum of DOD within some error bounded by the stepsize $s[k]$. Later through a numerical example, we demonstrate that the constant stepsizes are not practical because of the resulting slow convergence. To this end, we propose a scheme for computing the stepsizes to ensure faster convergence.

3.4. Imbalance-Driven Adaptive Stepsize

As noted previously, a broad class of stepsizes allow to achieve convergence. One such stepsize, which is diminishing with k, is of the form $s[k] = \frac{a}{k+b}$, where a and b are parameters to be chosen. Finding the right values for these parameters so as to ensure fast convergence for all scenarios and initial conditions, might be impossible. For example, when initial flows are far from the solution, setting a and b to some large values is needed in order to avoid slow convergence when k is large. However, these values cause an oscillatory behavior at the initial iterations or when initial flows are close to the solution, which results in deviating from the optimal solution.

To address these issues, we introduce a stepsize that adapts based on the maximum flow imbalance, namely, $\overline{b}[k] := \max_{i \in \mathcal{V}} |b_i[k]|$. Such scheme has certain parameters that are tuned on-line based on the evolution of the maximum flow imbalance. We give the following intuition behind its operation. Let $\mathcal{B}_\epsilon := \{x \in \mathbb{R} : |x| \leq \epsilon\}$. For any given ϵ, if $s[k]$ is too large, $\overline{b}[k]$ will never enter the ball $\mathcal{B}_\epsilon$. This suggests that $s[k]$ be small enough to avoid oscillation of the flow balances $|b_i[k]|$ around $\mathcal{B}_\epsilon$, without entering it, but large enough to avoid slow convergence.

In Algorithm 1, we provide the pseudocode for the proposed adaptive stepsize scheme, which consists of two parts. In the first part, if $\overline{b}[k]$ is inside $\mathcal{B}_\epsilon$, the radius of $\mathcal{B}_\epsilon$, $\epsilon[k]$, is decreased, and the stepsize $s[k]$ is also decreased to make sure $\overline{b}[k]$ enters a smaller ball $\mathcal{B}_\epsilon$. Note that $s[k] = \xi[k]\epsilon[k]$ depends on two parameters $\epsilon[k]$ and $\xi[k]$. Here, $\epsilon[k]$ determines the radius of the ball $\mathcal{B}_\epsilon$. The value of $\xi[k]$ is decreased to reduce the stepsize if, after a certain amount of time, $\overline{b}[k]$ is not inside the ball $\mathcal{B}_\epsilon$, which is exactly what is implemented in the second part of the proposed algorithm. Here, $\overline{b}[k]$ is compared to $\overline{b}[k-w]$ and $\tilde{b}[k]$ denoting an average of $\overline{b}[k]$ over w iterations, i.e., $\tilde{b}[k] := \frac{1}{w}\sum_{\tau=k-w}^{k-1} \overline{b}[\tau]$.

Finally, we note that $\rho \in (0, 1)$ and $\kappa \in (0, 1)$ are the constant parameters which determine the rate of reduction of $\epsilon[k]$ and

Algorithm 1: Imbalance-driven adaptive stepsize.

▷ Part 1

if $\overline{b}[k] < \epsilon[k]$ **then**
  $\epsilon[k] := \rho\epsilon[k-1]$ ▷ Decrease radius of ball $\mathcal{B}_\epsilon$
  $\xi[k] := \xi[k-1]$ ▷ Fix $\xi[k]$
  $s[k] := \xi[k]\epsilon[k]$ ▷ Shrink stepsize
end if

▷ Part 2

if no improvement after w iterations, i.e., $\overline{b}[k] > \overline{b}[k-w], \tilde{b}[k]$ **then**
  $\epsilon[k] := \epsilon[k-1]$ ▷ Fix radius of ball $\mathcal{B}_\epsilon$
  $\xi[k] := \kappa\xi[k-1]$ ▷ Decrease only $\xi[k]$
  $s[k] := \xi[k]\epsilon[k]$ ▷ Shrink stepsize
end if

Table 1: Stepsize parameters.

Adaptive stepsize	$\kappa = 0.93, \rho = 0.9, w = 10,$ $\epsilon(0) = \overline{b}(0)/2, \xi(0) = 0.2/\gamma$
Constant stepsize	$\alpha = 0.3/\gamma$
Diminishing stepsize	$a = 1/\gamma, b = 1$

$\xi[k]$, and $\epsilon[k]$ and $\xi[k]$ are initialized such that $\epsilon[0] = \overline{b}[0]/2$ or less, and $\xi[0] = 1/\gamma$ or less.

4. Numerical Example

In this section, we illustrate the usefulness of our theoretical results via simulations using the IEEE 39–bus test system with 10 generators; the topology of this system is shown in Fig. 1. All model parameters are taken from [26] with a few modifications. To make the system lossier, values of all series conductances, g_{ij}'s, have been increased by a factor of 10, and the base voltage has been decreased by a factor of 500; this turns the test system into a low voltage network with substantial losses.

Here, we assume that the test system is connected to the bulk grid through a single tie-line at bus 22. The goal here is to coordinate DERs to supply power to loads within the microgrid and required amount of power through the tie-line to the bulk grid as specified by an independent system operator (ISO) [27, 28]. In particular, the ISO sends a regulation signal every 2 to 4 s, which specifies the amount of power that needs to be supplied through the tie-line. This problem can be cast as DOD$'$, solved every time when a new regulation signal is received. We consider a sequence of regulation signal values, namely, the PJM RegD signal from December 2016 [29]; see Fig. 2.

To ensure that a minimum of DOD$'$ is a minimum of DOD1, we choose $\gamma = 12.5 > L/\beta$, $\lambda = 13 > \gamma$, and $\mu = 40 > \gamma + L_v$, as suggested by Propositions 1 – 3. We execute the algorithm in (34) – (36) and test its performance for three different stepsize rules: (i) the adaptive stepsize in Algorithm 1, (ii) the constant stepsize $s[k] = \alpha$, $\alpha > 0$, and (iii) the diminishing static stepsize $s[k] = \frac{a}{k+b}$, with positive constant parameters a and b. Values of the parameters of the stepsize rules are given in Table 1. For practical purposes, the number of the instances, when two neighbors i and j exchange and average their estimates of the flow ϕ^*_{ij}, is limited to 50 per every 2 s. Note that, between such averaging instances, nodes i and j update their local estimates using only local information, and, thus, several iterations can take place before communication is established between these two nodes. Figure 2a shows the numerical results for the adaptive stepsize, where we provided plots of the regulation signal sent by ISO and active power flowing through the tie-line into the bus 22; the tie-line power is able to closely track the regulation signal, under communication delays. Obviously, the tracking quality can be improved if the number of the averaging instances between neighboring nodes is increased.

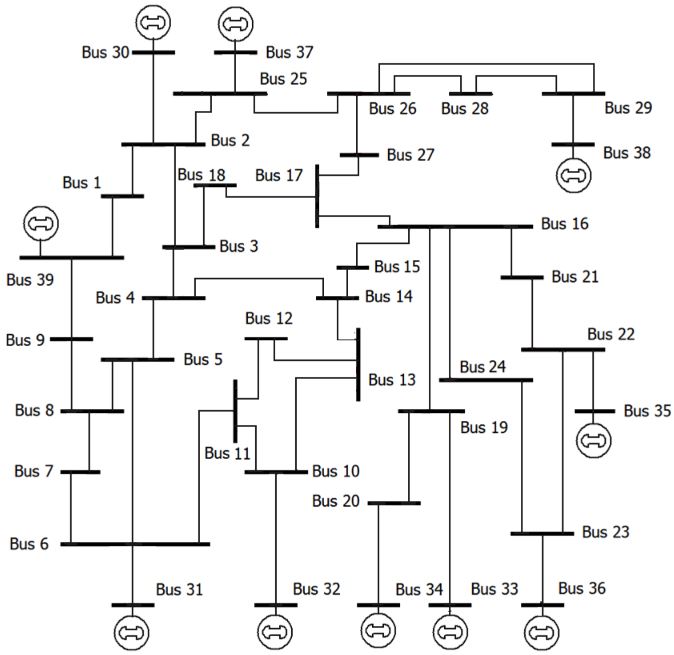

Figure 1: IEEE 39-bus test system.

We conducted similar numerical experiments for other stepsizes, the constant and diminishing static stepsizes. Figures 2b – 2c demonstrate that with these stepsizes the tie-line power is not able to track the regulation signal well, and many more iterations are required in order to achieve satisfactory tracking. We note that no pre-determined values for a and b ensure satisfactory convergence for the given regulation signal and random load variations, which motivates the adaptive way of choosing the stepsizes $s[k]$'s based on the flow balances, $b_i[k]$'s. As seen in Figure 2, the adaptive stepsize scheme significantly improves convergence since by using the maximum flow imbalance, $\overline{b}[k]$, it is able to better adapt to fluctuations in load and the regulation signal.

Figure 2 illustrates the fact that by solving DOD$'$ using the proposed algorithm in (34) – (36) with the adaptive stepsize we are able to satisfy all constraints in DOD1, i.e., the power balance constraint in (8), capacity constraints on the power outputs of the generators and the line constraints on the power flows. We also closely track the optimal solution, where the resulting cost is within 5-7% of its optimal value. The optimal solution can be tracked even better if we increase the execution time.

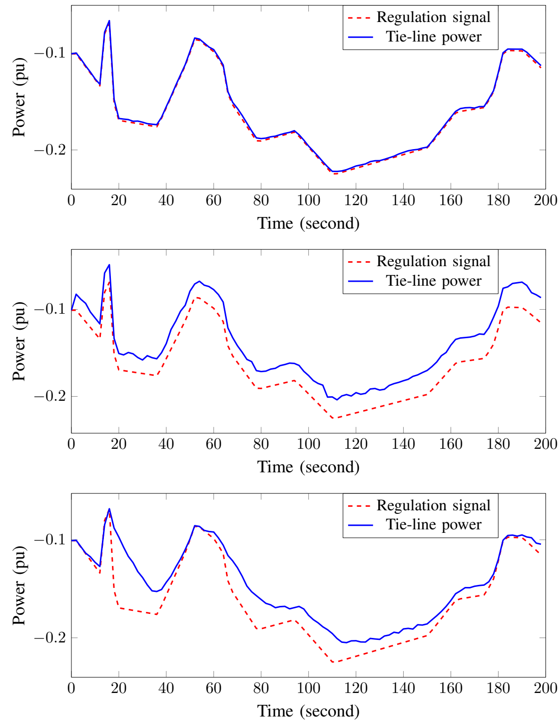

Figure 2: Tie-line power and regulation signal as a function of time for different stepsize rules (from top to bottom): (a) adaptive, (b) constant, and (c) diminishing static stepsizes.

Theoretically, the adaptive stepsize allows to achieve convergence to the optimal solution within any given error if the parameter w is taken to be large enough. But since in practice it is more preferable to limit the execution time and work with some good approximation of the optimal solution, we choose w to be smaller than the conservative value provided by the theory. If a very close convergence to the optimal solution is needed, then, the algorithm may as well switch from using the adaptive stepsize to using the diminishing static stepsize after a certain number of iterations, because the diminishing static stepsize always guarantees convergence to the optimal solution.

5. Conclusion

We have proposed a distributed algorithm that allows us to determine, under bounded time-varying communication delays, the optimal power outputs of a set of DERs in a power distribution network. The proposed approach minimizes the line losses and takes into account the network constraints. To improve the convergence rate, the algorithm uses the adaptive stepsize computed based on the maximum flow imbalance of the network. We have demonstrated through numerical simulations that the adaptive stepsize significantly outperforms diminishing static or constant stepsizes.

References

[1] D. Olivares, et al., Trends in microgrid control, IEEE Trans. Smart Grid 5 (4) (2014) 1905–1919.

[2] F. Katiraei, R. Iravani, N. Hatziargyriou, A. Dimeas, Microgrids management, IEEE Power & Energy Mag. 6 (3) (2008) 54–65.

[3] N. Hatziargyriou, H. Asano, R. Iravani, C. Marnay, Microgrids, IEEE Power & Energy Mag. 5 (4) (2007) 78–94.

[4] A. G. Tsikalakis, et al., Centralized control for optimizing microgrids operation, IEEE Trans. Energy Convers. 23 (1) (2008) 241–248.

[5] M. Smith, D. Ton, Key connections: The U.S. Department of Energy's microgrid initiative, IEEE Power & Energy Mag. 11 (4) (2013) 22–27.

[6] Z. Zhang, M. Y. Chow, Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid, IEEE Trans. Power Syst. 27 (4) (2012) 1761–1768.

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

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

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

[10] J. Mohammadi, G. Hug, S. Kar, Fully distributed dc-opf approach for power flow control, in: Proc. IEEE Power & Energy Soc. Gen. Meeting, 2015, pp. 1–5.

[11] S. Liu, X. Wang, P. X. Liu, Impact of communication delays on secondary frequency control in an islanded microgrid, IEEE Trans. Ind. Electron. 62 (4) (2015) 2021–2031.

[12] C. Ahumada, et al., Secondary control strategies for frequency restoration in islanded microgrids with consideration of communication delays, IEEE Trans. Smart Grid 7 (3) (2016) 1430–1441.

[13] J. Lai, H. Zhou, X. Lu, X. Yu, W. Hu, Droop-based distributed cooperative control for microgrids with time-varying delays, IEEE Trans. Smart Grid 7 (4) (2016) 1775–1789.

[14] J. Lai, H. Zhou, X. Lu, Z. Liu, Distributed power control for ders based on networked multiagent systems with communication delays, Neurocomputing 179 (Supplement C) (2016) 135 – 143.

[15] G. Chen, Z. Zhao, Distributed optimal active power control in microgrid with communication delays, in: Proc. Chinese Control Conf., 2016, pp. 7515–7520.

[16] W. Liu, et al., Pinning-based distributed cooperative control for autonomous microgrids under uncertain communication topologies, IEEE Trans. Power Syst. 31 (2) (2016) 1320–1329.

[17] C. Li, et al., Convergence analysis of distributed control for operation cost minimization of droop controlled dc microgrid based on multiagent, in: IEEE Appl. Power Electron. Conf. & Expo., 2016, pp. 3459–3464.

[18] H. Cai, G. Hu, Distributed nonlinear hierarchical control of ac microgrid via unreliable communication, IEEE Trans. Smart Grid (to appear).

[19] J. Wu, et al., Distributed optimal dispatch of distributed energy resources over lossy communication networks, IEEE Trans. Smart Grid 8 (6) (2017) 3125–3137.

[20] A. Nedić, A. Olshevsky, Distributed optimization over time-varying directed graphs, IEEE Trans. Autom. Control 60 (3) (2015) 601–615.

[21] J. Zhu, Optimization of Power System Operation, 2nd Edition, Wiley-IEEE Press, 2015.

[22] W. Jewells, Optimal flows through networks with gains, Operations Research 10 (4) (1962) 476–499.

[23] S. Sojoudi, J. Lavaei, Convexification of generalized network flow problem with application to power systems, in: Proc. IEEE Conf. Decision and Control, 2013, pp. 7552–7559.

[24] A. Nedić, A. Ozdaglar, Distributed subgradient methods for multi-agent optimization, IEEE Trans. Autom. Control 54 (1) (2009) 48–61.

[25] M. Zholbaryssov, A. D. Domínguez-García, Microgrid distributed frequency control over time-varying communication networks, in: IEEE Conf. Decision and Control, 2018, accepted.

[26] R. D. Zimmerman, C. E. Murillo-Sanchez, R. J. Thomas, Matpower: Steady-state operations, planning, and analysis tools for power systems research and education, IEEE Trans. Power Syst. 26 (1) (2011) 12–19.

[27] D. Fooladivanda, M. Zholbaryssov, A. D. Dominguez-Garcia, Control of networked distributed energy resources in grid-connected ac microgrids, IEEE Trans. Control Network Syst. (to appear).

[28] Performance, Mileage and the Mileage Ratio., `http://www.pjm.com.`

[29] RTO, Regulation Signal Data., `http://www.pjm.com/markets-and-operations/ancillary-services.aspx.`