Initialization-free distributed coordination for economic dispatch under varying loads and generator commitment ⋆

Ashish Cherukuri [a] Jorge Cortés [a]

[a] *Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, 92093, USA*

Abstract

This paper considers the economic dispatch problem for a network of power generating units communicating over a strongly connected, weight-balanced digraph. The collective aim is to meet a power demand while respecting individual generator constraints and minimizing the total generation cost. We design a distributed coordination algorithm consisting of two interconnected dynamical systems. One block uses dynamic average consensus to estimate the evolving mismatch in load satisfaction given the generation levels of the units. The other block adjusts the generation levels based on the optimization objective and the estimate of the load mismatch. Our convergence analysis shows that the resulting strategy provably converges to the solution of the dispatch problem starting from any initial power allocation, and therefore does not require any specific procedure for initialization. We also characterize the algorithm robustness properties against the addition and deletion of units (capturing scenarios with intermittent power generation) and its ability to track time-varying loads. Our technical approach employs a novel refinement of the LaSalle Invariance Principle for differential inclusions, that we also establish and is of independent interest. Several simulations illustrate our results.

Key words: distributed optimization; economic dispatch; power networks; dynamic average consensus; invariance principles

1 Introduction

The advent of renewable energy sources and their integration into electricity grids is making power generation and distribution an increasingly decentralized problem. The large-scale and highly dynamic nature of the resulting grid optimization problems makes traditional centralized, top-down approaches impractical because they rely on the assumption of a fixed, limited number of generation units. To solve these problems efficiently, there is a need to design distributed algorithms that can handle dynamic loads, provide plug-and-play capabilities, are robust against transmission and generation failures, and adequately preserve the privacy of the entities involved. These considerations motivate us to consider here the design of distributed algorithmic solutions to the economic dispatch (ED) problem, where a group of power generators aims to meet a power demand while minimizing the total generation cost (the summation of individual costs) and respecting the individual generators' capacity constraints. We are interested in the synthesis of strategies that solve the ED problem starting from any initial power allocation, can handle time-varying loads, and are robust against intermittent power generation caused by unit addition and deletion.

Literature review: The ED problem has been traditionally solved in a centralized manner, see e.g. [Chowdhury and Rahman, 1990] and references therein. Since distributed algorithmic solutions to grid optimization problems are envisioned as part of the future smart grid [Farhangi, 2010], this has motivated the appearance of a number of distributed strategies for the ED problem in the literature. While there exists a broad variety in the assumptions made, a majority of the works rely on the specific form of the solutions of the optimization problem and propose consensus-based algorithms. Many works consider convex, quadratic cost functions for the power generators and perform consensus over their incremental costs under undirected [Zhang and Chow, 2012, Kar and Hug, 2012] or directed [Dominguez-Garcia et al., 2012, Binetti et al., 2014a] communication topologies. Some works consider general convex cost functions, like we do here, but either do not consider capacity constraints on the generators [Mudumbai et al., 2012], assume the initial power allocation to meet the total load [Cherukuri and Cortés, 2015, Pantoja et al., 2014], or require feedback on the power mismatch

⋆ A preliminary version of this work appeared as [Cherukuri and Cortés, 2014] at the 2014 Allerton Conference on Communication, Control, and Computing, Monticello, Illinois.

Email addresses: acheruku@ucsd.edu (Ashish Cherukuri), cortes@ucsd.edu (Jorge Cortés).

from the shift in frequency due to primary droop control [Zhang et al., 2014]. Along with load and capacity constraints, [Binetti et al., 2014a, Loia and Vaccaro, 2013] consider transmission losses, and [Binetti et al., 2014b] additionally take into account valve-point loading effects and prohibited operating zones. However, these constraints make the problem nonconvex and prevent these works from obtaining theoretical guarantees on the algorithm convergence properties. In [Du et al., 2012], the authors propose best-response dynamics for a potential-game formulation of the nonconvex ED problem, but the implementation requires all-to-all communication among the generators. [Xiao and Boyd, 2006, Johansson and Johansson, 2009] introduce distributed methods to solve resource allocation problem very similar to the ED problem, but without taking into account individual agent constraints. Instead, these are incorporated in the formulation of [Simonetto et al., 2012], but the proposed algorithm arrives at suboptimal solutions of the optimization problem. Our work is also related with the literature on distributed optimal frequency control in power networks [Zhao et al., 2014, Mallada et al., 2015, Li et al., 2015] dealing with primary and secondary control (tightly connected with the physical network and its dynamics over short time horizons), while we focus here on tertiary control (which allows for longer time horizons and more flexibility in planning). Our algorithm design and analysis rely on dynamic average consensus and differential inclusions. In dynamic average consensus, see e.g. [Freeman et al., 2006, Kia et al., 2015] and references therein, each agent has access to a time-varying input signal and interacts with its neighbors in order to track the average of the input signals across the network. We build on our previous work [Cherukuri and Cortés, 2015], which requires a proper algorithm initialization, and employ tools from dynamic average consensus to synthesize a coordination strategy that converges from any initial condition. Regarding analysis, our technical approach builds on Lyapunov stability tools for differential inclusions and nonsmooth systems, see e.g. [Bacciotti and Ceragioli, 1999, Cortés, 2008] and references therein. Of particular importance is the work [Arsie and Ebenbauer, 2010] for differential equations, that provides a way to further refine the description of omega-limit sets of trajectories by employing more than one LaSalle-type function.

Statement of contributions: We start with the formal definition of the ED problem for a network of power generators communicating over a strongly connected, weight-balanced digraph. The optimization problem is convex as the individual cost functions are smooth and convex, the load satisfaction is a linear constraint, and the individual generators' capacities prescribe convex inequality constraints. Our formulation is a simplification of the ED problem in its full generality, which in practice may have additional constraints that make it nonconvex. However, our developments show that obtaining a provably correct algorithmic solution for the formulation here of the ED problem given the performance requirements (distributed, convergent irrespective of initial condition, able to handle time-varying loads, and robust to intermittent power generation) is challenging. Our first contribution is the design of a centralized algorithm, termed "load mismatch + Laplacian-nonsmooth-gradient" dynamics, that solves the ED problem starting from any initial power allocation. This strategy has two components: one component seeks to optimize the network generation cost while keeping constant the total power generated; the other component is a feedback correction term driven by the error between the desired total load and the network generation. This latter term is responsible for ensuring that the algorithm trajectories asymptotically satisfy the load satisfaction constraint irrespective of the initial power allocation. These observations set the basis for our second contribution, which is the synthesis of a distributed coordination algorithm, termed "dynamic average consensus + Laplacian-nonsmooth-gradient" dynamics, with the same convergence guarantees. Our design consists of two coupled dynamical systems: a dynamic average consensus algorithm to estimate the mismatch between generation and desired load in a distributed fashion and distributed Laplacian-nonsmooth-gradient that employ these estimates to dynamically allocate the unit generation levels. The convergence analysis of both the centralized and distributed algorithms relies on a combination of tools from algebraic graph theory, nonsmooth analysis, set-valued dynamical systems, and dynamic average consensus, and most notably on a refined version of the LaSalle Invariance Principle for differential inclusions, which constitutes our third contribution. Roughly speaking, the application of the known LaSalle Invariance Principle would only establish asymptotic convergence towards the network satisfaction of the total load. Instead, the use of the refined version allows us, for each algorithm, to establish global convergence of the trajectories to the solutions of the ED problem. Our final contribution is the formal characterization of the robustness properties of the distributed algorithm. Building on the observation that the mismatch dynamics between network generation and total load are exponentially convergent and input-to-state stable, we establish the algorithm ability to track time-varying loads and its robustness in scenarios with intermittent power generation. The generality of our design and analysis makes the proposed algorithm applicable to other resource allocation problems with constraints beyond the area of power systems.

2 Preliminaries

This section introduces basic concepts and preliminaries. We begin with some notational conventions. Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, $\mathbb{Z}_{\geq 1}$ denote the real, nonnegative real, positive real, and positive integer numbers, resp. For $r \in \mathbb{R}$ we denote $\mathcal{H}_r = \{x \in \mathbb{R}^n \mid \mathbf{1}_n^\top x = r\}$. The 2- and ∞-norms on $\mathbb{R}^n$ and their respective induced norms on $\mathbb{R}^{n\times n}$ are denoted with $\|\cdot\|$ and $\|\cdot\|_\infty$, resp. We let $B(x,\delta) = \{y \in \mathbb{R}^n \mid \|y - x\| < \delta\}$. For $D \subset \mathbb{R}^n$, $\overline{D}$ denotes its closure. For $x \in \mathbb{R}^n$,

$x_i \in \mathbb{R}$ denotes its i-th component. Given vectors $x, y \in \mathbb{R}^n$, $x \leq y$ if and only if $x_i \leq y_i$ for all $i \in \{1, \dots, n\}$. We denote $\mathbf{1}_n = (1, \dots, 1) \in \mathbb{R}^n$.A set-valued map $f : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ associates to each point in $\mathbb{R}^n$ a set in $\mathbb{R}^m$. For a symmetric matrix $A \in \mathbb{R}^{n\times n}$, $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the minimum and maximum eigenvalues of A. Finally, we let $[u]^+ = \max\{0, u\}$ for $u \in \mathbb{R}$.

2.1 Graph theory

We present basic notions from algebraic graph theory following [Bullo et al., 2009]. A *directed graph* (or *digraph*) is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $\mathcal{V} = \{1, \dots, n\}$ the vertex set and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ the edge set. A path is a sequence of vertices connected by edges. A digraph is *strongly connected* if there is a path between any pair of vertices. The sets of out- and in-neighbors of v are, resp., $N^{\text{out}}(v) = \{w \in \mathcal{V} \mid (v, w) \in \mathcal{E}\}$ and $N^{\text{in}}(v) = \{w \in \mathcal{V} \mid (w, v) \in \mathcal{E}\}$. A *weighted digraph* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathsf{A})$ is composed of a digraph $(\mathcal{V}, \mathcal{E})$ and an *adjacency matrix* $\mathsf{A} \in \mathbb{R}^{n\times n}_{\geq 0}$ with $a_{ij} > 0$ iff $(i, j) \in \mathcal{E}$. The weighted out- and in-degree of i are, resp., $d^{\text{out}}(i) = \sum_{j=1}^n a_{ij}$ and $d^{\text{in}}(i) = \sum_{j=1}^n a_{ji}$. The *Laplacian* matrix is $\mathsf{L} = \mathsf{D}_{\text{out}} - \mathsf{A}$, where D_{out} is the diagonal matrix with $(\mathsf{D}_{\text{out}})_{ii} = d^{\text{out}}(i)$, for all $i \in \{1, \dots, n\}$. Note that $\mathsf{L}\mathbf{1}_n = 0$. If $\mathcal{G}$ is strongly connected, then 0 is a simple eigenvalue of L. $\mathcal{G}$ is undirected if $\mathsf{L} = \mathsf{L}^\top$. $\mathcal{G}$ is *weight-balanced* if $d^{\text{out}}(v) = d^{\text{in}}(v)$, for all $v \in \mathcal{V}$ iff $\mathbf{1}_n^\top \mathsf{L} = 0$ iff $\mathsf{L} + \mathsf{L}^\top \geq 0$. Note that any undirected graph is weight-balanced. If $\mathcal{G}$ is weight-balanced and strongly connected, then 0 is a simple eigenvalue of $\mathsf{L} + \mathsf{L}^\top$. In such case, one has for $x \in \mathbb{R}^n$,

$$x^\top(\mathsf{L} + \mathsf{L}^\top)x \geq \lambda_2(\mathsf{L} + \mathsf{L}^\top)\left\|x - \frac{1}{n}(\mathbf{1}_n^\top x)\mathbf{1}_n\right\|^2, \quad (1)$$

with $\lambda_2(\mathsf{L}+\mathsf{L}^\top)$ the smallest non-zero eigenvalue of $\mathsf{L}+\mathsf{L}^\top$.

2.2 Dynamic average consensus

Here, we introduce notions on dynamic average consensus following [Kia et al., 2015]. Consider $n \in \mathbb{Z}_{\geq 1}$ agents communicating over a strongly connected, weight-balanced digraph $\mathcal{G}$ whose Laplacian is denoted as L. Each agent is associated with a state $x_i \in \mathbb{R}$ and an input signal $t \mapsto u_i(t) \subset \mathbb{R}$ that is measurable and locally essentially bounded. The aim is to provide distributed dynamics such that the state of each agent $x_i(t)$ tracks the average signal $\frac{1}{n}\sum_{i=1}^n u_i(t)$ asymptotically. This can be achieved via the dynamics $X_{\mathsf{dac}} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$,

$$\begin{aligned}\dot{x} &= -\alpha x - \beta \mathsf{L} x - v + \nu u,\\ \dot{v} &= \alpha\beta \mathsf{L} x,\end{aligned}$$

where $\alpha, \beta, \nu > 0$ are design parameters and $v \in \mathbb{R}^n$ is an auxiliary state. If the initial condition satisfies $\mathbf{1}_n^\top v(0) = 0$ and the time-derivatives of the input signals are bounded, then one can show, cf. [Kia et al., 2015, Corollary 4.1], that the error signal $t \mapsto \left|x_i(t) - \frac{1}{n}\sum_{i=1}^n u_i(t)\right|$ is ultimately bounded for each $i \in \{1, \dots, n\}$. Moreover, this error vanishes if the input signal converges to a constant value.

2.3 Nonsmooth analysis and differential inclusions

We review here some notions from nonsmooth analysis and differential inclusions following [Cortés, 2008]. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *locally Lipschitz* at $x \in \mathbb{R}^n$ if there exist $L_x, \epsilon \in (0, \infty)$ such that $\|f(y) - f(y')\| \leq L_x\|y - y'\|$, for all $y, y' \in B(x, \epsilon)$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *regular* at $x \in \mathbb{R}^n$ if, for all $v \in \mathbb{R}^n$, the right and generalized directional derivatives of f at x in the direction of v coincide, see [Cortés, 2008] for definitions of these notions. A function that is continuously differentiable at x is regular at x. Also, a convex function is regular. A set-valued map $\mathcal{H} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *upper semicontinuous* at $x \in \mathbb{R}^n$ if, for all $\epsilon \in (0, \infty)$, there exists $\delta \in (0, \infty)$ such that $\mathcal{H}(y) \subset \mathcal{H}(x) + B(0, \epsilon)$ for all $y \in B(x, \delta)$. Also, $\mathcal{H}$ is *locally bounded* at $x \in \mathbb{R}^n$ if there exist $\epsilon_x, \delta_x \in (0, \infty)$ such that $\|z\| \leq \epsilon_x$ for all $z \in \mathcal{H}(y)$ and $y \in B(x, \delta_x)$. $\mathcal{H}$ is locally bounded if it is so at each point in $\mathbb{R}^n$.

Given a locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$, let Ω_f be the set (of measure zero) of points where f is not differentiable. The *generalized gradient* $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is

$$\partial f(x) = \text{co}\{\lim_{i\to\infty} \nabla f(x_i) \mid x = \lim_{i\to\infty} x_i \,\text{and}\, x_i \notin S \cup \Omega_f\},$$

where co denotes convex hull and $S \subset \mathbb{R}^n$ can be any set of measure zero. The map ∂f is locally bounded, upper semicontinuous, and takes non-empty, compact, and convex values. A *critical point* x of f satisfies $0 \in \partial f(x)$.

Given a set-valued map $\mathcal{H} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, a differential inclusion on $\mathbb{R}^n$ is

$$\dot{x} \in \mathcal{H}(x). \quad (3)$$

A solution of (3) on $[0, T] \subset \mathbb{R}$ is an absolutely continuous map $x : [0, T] \to \mathbb{R}^n$ that satisfies (3) for almost all $t \in [0, T]$. If $\mathcal{H}$ is locally bounded, upper semicontinuous, and takes non-empty, compact, and convex values, then existence of solutions is guaranteed. The set of equilibria of (3) is $\text{Eq}(\mathcal{H}) = \{x \in \mathbb{R}^n \mid 0 \in \mathcal{H}(x)\}$. A set $S \subset \mathbb{R}^n$ is *weakly (resp., strongly) positively invariant* under (3) if, for each $x \in S$, at least a solution (resp., all solutions) starting from x is (resp., are) entirely contained in S. For dynamics with uniqueness of solution, both notions coincide and are referred as *positively invariant*. Given a locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$, the *set-valued Lie derivative* $\mathcal{L}_\mathcal{H} f : \mathbb{R}^n \rightrightarrows \mathbb{R}$ of f with respect to (3) is

$$\mathcal{L}_\mathcal{H} f(x) = \{a \in \mathbb{R} \mid \exists v \in \mathcal{H}(x) \text{ s.t. } \zeta^\top v = a \text{ for all } \zeta \in \partial f(x)\}.$$

For a trajectory $t \mapsto \varphi(t), \varphi(0) \in \mathbb{R}^n$ of (3), the evolution of f along it satisfies

$$\frac{d}{dt} f(\varphi(t)) \in \mathcal{L}_\mathcal{H} f(\varphi(t))$$

for almost all $t \geq 0$. The omega-limit set of the trajec-

tory, denoted $\Omega(\varphi)$, is the set of all points $y \in \mathbb{R}^n$ for which there exists a sequence $\{t_k\}_{k=1}^{\infty}$ with $t_k \to \infty$ and $\lim_{k\to\infty} \varphi(t_k) = y$. If the trajectory is bounded, then the omega-limit set is nonempty, compact, connected, and weakly invariant. These tools allow us to characterize the asymptotic behavior of solutions of differential inclusions. In Appendix A we develop a novel refinement of the LaSalle Invariance Principle for differential inclusions, see e.g., [Cortés, 2008], which is suitable for the analysis of the coordination algorithms.

3 Problem statement

This section presents the network model and the economic dispatch problem we set out to solve in a distributed and robust fashion. Consider $n \in \mathbb{Z}_{\geq 1}$ power generators communicating over a strongly connected and weight-balanced digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathsf{A})$. Each generator corresponds to a vertex in the digraph and an edge (i, j) represents the ability of generator j to send information to generator i. The cost of power generation for unit i is measured by $f_i : \mathbb{R} \to \mathbb{R}_{\geq 0}$, assumed to be convex and continuously differentiable. Representing the power generated by unit i by $P_i \in \mathbb{R}$, the total cost incurred by the network with the power allocation $P = (P_1, \dots, P_n) \in \mathbb{R}^n$ is measured by $f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ as

$$f(P) = \sum_{i=1}^{n} f_i(P_i).$$

Note that f is convex and continuously differentiable. The generators aim to minimize the total cost $f(P)$ while meeting the total power load $P_l \in \mathbb{R}_{>0}$, i.e., $\sum_{i=1}^{n} P_i = P_l$. Each generator has an upper and a lower limit on the power it can produce, $P_i^m \leq P_i \leq P_i^M$ for $i \in \{1, \dots, n\}$. Formally, the economic dispatch (ED) problem is

$$\begin{aligned} &\text{minimize} && f(P), && \text{(4a)}\\ &\text{subject to} && \mathbf{1}_n^\top P = P_l, && \text{(4b)}\\ & && P^m \leq P \leq P^M. && \text{(4c)} \end{aligned}$$

The constraint (4b) is the *load condition* and (4c) are the *box constraints*. The set of allocations satisfying the box constraints is $\mathcal{F}_B = \{P \in \mathbb{R}^n \mid P^m \leq P \leq P^M\}$. Further, we denote the feasibility set of (4) as $\mathcal{F}_{\text{ED}} = \mathcal{F}_B \cap \mathcal{H}_{P_l} = \{P \in \mathbb{R}^n \mid P^m \leq P \leq P^M \text{ and } \mathbf{1}_n^\top P = P_l\}$ and the set of solutions as $\mathcal{F}_{\text{ED}}^*$. Since $\mathcal{F}_{\text{ED}}$ is compact, $\mathcal{F}_{\text{ED}}^*$ is compact. Note that $P^M \in \mathcal{F}_{\text{ED}}$ implies $\mathcal{F}_{\text{ED}} = \{P^M\}$. Similarly $P^m \in \mathcal{F}_{\text{ED}}$ implies $\mathcal{F}_{\text{ED}} = \{P^m\}$. Therefore, we assume P^M and P^m are not feasible.

Our objective is to design a distributed coordination algorithm that allows the team of generators to solve the ED problem (4) starting from any initial condition, can handle time-varying loads, and is robust to intermittent power generation.

Remark 3.1 *(Additional practical constraints):* We do not consider here, for simplicity, other constraints on the ED problem such as transmission losses, transmission line capacities, valve-point loading effects, ramp rate limits, and prohibited operating zones. As our forthcoming treatment will show, the design and analysis of algorithmic solutions to the ED problem without these additional constraints is already quite challenging given our performance requirements. Nevertheless, Remark 5.5 later comments on how to adapt our algorithm to deal with more general scenarios. •

Our design strategy relies on the following reformulation of the ED problem without inequality constraints. Consider the modified ED problem

$$\begin{aligned} &\text{minimize} && f^\epsilon(P), && \text{(5a)}\\ &\text{subject to} && \mathbf{1}_n^\top P = P_l, && \text{(5b)} \end{aligned}$$

where the objective function is

$$f^\epsilon(P) = \sum_{i=1}^{n} f_i(P_i) + \frac{1}{\epsilon}\Big(\sum_{i=1}^{n}([P_i - P_i^M]^+ + [P_i^m - P_i]^+)\Big).$$

This corresponds to each generator $i \in \{1, \dots, n\}$ having the modified local cost

$$f_i^\epsilon(P_i) = f_i(P_i) + \frac{1}{\epsilon}([P_i - P_i^M]^+ + [P_i^m - P_i]^+).$$

Note that f_i^ϵ is convex, locally Lipschitz, and continuously differentiable on $\mathbb{R}$ except at $P_i = P_i^m$ and $P_i = P_i^M$. Moreover, the total cost f^ϵ is convex, locally Lipschitz, and regular. According to our previous work [Cherukuri and Cortés, 2015, Proposition 5.2], the solutions to the original (4) and the modified (5) ED problems coincide for $\epsilon \in \mathbb{R}_{>0}$ such that

$$\epsilon < \frac{1}{2\max_{P\in\mathcal{F}_{\text{ED}}} \|\nabla f(P)\|_\infty}. \tag{6}$$

Throughout the paper, we assume the parameter ϵ satisfies this condition. A useful fact is that $P^* \in \mathbb{R}^n$ is a solution of (5) if and only if there exists $\mu \in \mathbb{R}$ such that

$$\mu \mathbf{1}_n \in \partial f^\epsilon(P^*) \quad \text{and} \quad \mathbf{1}_n^\top P^* = P_l. \tag{7}$$

4 Robust centralized algorithmic solution

This section presents a robust strategy to make the network power allocation converge to the solution set of the ED problem starting from any initial condition. Even though this algorithm is centralized, its design provides enough insight to tackle later the design of a distributed algorithmic solution. Consider the "load mismatch + Laplacian-nonsmooth-gradient" (abbreviated `lm+L`∂) dynamics, represented by the set-valued

map $X_{\mathtt{lm+L}\partial} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$,

$$\dot{P} \in -\mathsf{L}\partial f^\epsilon(P) + \frac{1}{n}(P_l - \mathbf{1}_n^\top P)\mathbf{1}_n, \tag{8}$$

where L is the Laplacian associated to the strongly connected and weight-balanced communication digraph $\mathcal{G}$. For each generator, the first term seeks to minimize the total cost while leaving unchanged the total generated power. The second term is a feedback element that seeks to drive the units towards the satisfaction of the load. The first term is computable using information from its neighbors but the second term requires them to know the aggregated state of the whole network, which makes it not directly implementable in a distributed manner. The next result shows that the trajectories of (8) converge to the set of solutions of the ED problem.

Theorem 4.1 *(Convergence of the trajectories of $X_{\mathtt{lm+L}\partial}$ to the solutions of the* ED *problem): The trajectories of* (8) *starting from any point in $\mathbb{R}^n$ converge to the set of solutions of* (4)*.*

PROOF. Our proof strategy proceeds by applying the refined LaSalle Invariance Principle for differential inclusions established in Appendix A, cf. Proposition A.1. Consider the following function $V_1 : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$,

$$V_1(P) = \frac{1}{2}(P_l - \mathbf{1}_n^\top P)^2.$$

The set-valued Lie derivative of V_1 along $X_{\mathtt{lm+L}\partial}$ is

$$\mathcal{L}_{X_{\mathtt{lm+L}\partial}} V_1(P) = \{-(P_l - \mathbf{1}_n^\top P)^2\} = \{-2V_1(P)\}.$$

Thus, starting at any $P(0) \in \mathbb{R}^n$, the trajectory of $X_{\mathtt{lm+L}\partial}$ satisfies $V_1(P(t)) = V_1(P(0))e^{-2t}$ and its omega-limit set (provided the trajectory is bounded, a fact that we assume is true for now and establish later) is contained in $\mathcal{H}_{P_l}$. In the notation of Proposition A.1, $\mathcal{H}_{P_l}$ plays the role of the closed submanifold $\mathcal{S}$ of $\mathbb{R}^n$. We next show that the hypotheses of this result hold. In the notation of the Lemma A.1, the function f^ϵ, the map $(P, \zeta) \mapsto -\mathsf{L}\zeta$, and the set-valued map $P \rightrightarrows -\mathsf{L}\partial f^\epsilon(P)$ play the role of W, g, and F, respectively (our choice of F is because $X_{\mathtt{lm+L}\partial}$ takes the form $\dot{P} \in -\mathsf{L}\partial f^\epsilon(P)$ on $\mathcal{S} = \mathcal{H}_{P_l}$). Notice that $\zeta \mapsto -\mathsf{L}\zeta$ is a continuous map and, since $\mathcal{G}$ is strongly connected and weight-balanced, we have $\zeta^\top(-\mathsf{L}\zeta) = -\frac{1}{2}\zeta^\top(\mathsf{L} + \mathsf{L}^\top)\zeta \leq 0$ for any $\zeta \in \partial f^\epsilon(P)$. Therefore, Lemma A.1(i) is satisfied. Moreover, if $\zeta^\top(-\mathsf{L}\zeta) = 0$ for some $\zeta \in \partial f^\epsilon(P)$, then $\zeta \in \operatorname{span}\{\mathbf{1}_n\}$. Since for $P \in \mathcal{H}_{P_l}$, we have

$$\mathcal{L}_{X_{\mathtt{lm+L}\partial}} f^\epsilon(P) = \{-\zeta^\top \mathsf{L}\zeta \mid \zeta \in \partial f^\epsilon(P)\},$$

we deduce $0 \in \mathcal{L}_{X_{\mathtt{lm+L}\partial}} f^\epsilon(P)$, i.e., Lemma A.1(ii) holds. The application of Lemma A.1 then yields that Proposition A.1(ii) holds too. In addition, from the above analysis, note that if $0 \in \mathcal{L}_{X_{\mathtt{lm+L}\partial}} f^\epsilon(P)$ for some $P \in \mathcal{H}_{P_l}$, then there exists $\mu \in \mathbb{R}$ such that $\mu\mathbf{1}_n \in \partial f^\epsilon(P)$ and, from (7), P is a solution of (4). Therefore, $\{P \in \mathcal{H}_{P_l} \mid 0 \in \mathcal{L}_{X_{\mathtt{lm+L}\partial}} f^\epsilon(P)\}$ is the set of solutions of the ED problem and belongs to a level set of f^ϵ, which establishes that Proposition A.1(i) also holds.

To be able to apply Proposition A.1 and conclude the proof, it remains to show that the trajectories of $X_{\mathtt{lm+L}\partial}$ are bounded. We reason by contradiction, i.e., assume there exists a trajectory $t \mapsto P(t)$, $P(0) \in \mathbb{R}^n$ of $X_{\mathtt{lm+L}\partial}$ such that $\|P(t)\| \to \infty$. From the analysis above, we know that along this trajectory $\mathbf{1}_n^\top P(t) \to P_l$ and $f^\epsilon(P(t)) \to \infty$ (as f^ϵ is radially unbounded). Therefore, there exist a sequence of times $\{t_k\}_{k=1}^\infty$ with $t_k \to \infty$ such that for all $k \in \mathbb{Z}_{\geq 1}$,

$$\left|\mathbf{1}_n^\top P(t_k) - P_l\right| < \frac{1}{k} \text{ and } \max \mathcal{L}_{X_{\mathtt{lm+L}\partial}} f^\epsilon(P(t_k)) > 0. \tag{9}$$

This implies that there exists a sequence $\{\zeta_k\}_{k=1}^\infty$ with $\zeta_k \in \partial f^\epsilon(P(t_k))$ such that, for all $k \in \mathbb{Z}_{\geq 1}$,

$$\begin{aligned} &-\zeta_k^\top \mathsf{L}\zeta_k + \frac{1}{n}(P_l - \mathbf{1}_n^\top P(t_k))(\mathbf{1}_n^\top \zeta_k) > 0 \\ &\Rightarrow -\zeta_k^\top \Big(\frac{\mathsf{L} + \mathsf{L}^\top}{2}\Big)\zeta_k + \frac{1}{nk}\left|\mathbf{1}_n^\top \zeta_k\right| > 0 \\ &\Rightarrow -\frac{\lambda_2(\mathsf{L} + \mathsf{L}^\top)}{2}\left\|\zeta_k - \frac{1}{n}(\mathbf{1}_n^\top \zeta_k)\mathbf{1}_n\right\|^2 + \frac{1}{nk}\left|\mathbf{1}_n^\top \zeta_k\right| > 0, \end{aligned} \tag{10}$$

where we have used (9) in the first implication and (1) in the second. Next, we consider two cases depending on whether (a) $\left|\mathbf{1}_n^\top \zeta_k\right|$ is bounded or (b) $\left|\mathbf{1}_n^\top \zeta_k\right| \to \infty$. In case (a), taking the limit $k \to \infty$ in the last inequality of (10), we get

$$\lim_{k\to\infty} \left\|\zeta_k - \frac{1}{n}(\mathbf{1}_n^\top \zeta_k)\mathbf{1}_n\right\| = 0. \tag{11}$$

Since, $\|P(t)\| \to \infty$ and $\mathbf{1}_n^\top P(t) \to P_l$, there exist $i, j \in \{1, \ldots, n\}$ such that $P_i(t_k) \to \infty$ and $P_j(t_k) \to -\infty$. Let $P^* \in \mathcal{F}_{\text{ED}}^*$ and $\mu\mathbf{1}_n \in \partial f^\epsilon(P^*)$ for some $\mu \in \mathbb{R}$. Then, without loss of generality, we assume that $P_i^* \leq P_i(t_k) \leq P_i(t_{k+1})$ and $P_j^* \geq P_j(t_k) \geq P_j(t_{k+1})$ for all k. This fact along with the expression of $\partial f_i^\epsilon : \mathbb{R} \rightrightarrows \mathbb{R}$,

$$\partial f_i^\epsilon(P_i) = \begin{cases} \{\nabla f_i(P_i) - \frac{1}{\epsilon}\} & \text{if } P_i < P_i^m, \\ [\nabla f_i(P_i) - \frac{1}{\epsilon}, \nabla f_i(P_i)] & \text{if } P_i = P_i^m, \\ \{\nabla f_i(P_i)\} & \text{if } P_i^m < P_i < P_i^M, \\ [\nabla f_i(P_i), \nabla f_i(P_i) + \frac{1}{\epsilon}] & \text{if } P_i = P_i^M, \\ \{\nabla f_i(P_i) + \frac{1}{\epsilon}\} & \text{if } P_i > P_i^M. \end{cases}$$

gives us the following property for all $k \in \mathbb{Z}_{\geq 1}$,

$$\min \partial f_i^\epsilon(P_i(t_k)) \geq \mu, \quad \max \partial f_j^\epsilon(P_j(t_k)) \leq \mu, \tag{12a}$$
$$\min \partial f_i^\epsilon(P_i(t_{k+1})) \geq \max \partial f_i^\epsilon(P_i(t_k)), \tag{12b}$$
$$\max \partial f_j^\epsilon(P_j(t_{k+1})) \leq \min \partial f_j^\epsilon(P_j(t_k)). \tag{12c}$$

Note that the limit (11) yields $\lim_{k\to\infty} |(\zeta_k)_i - (\zeta_k)_j| = 0$. On the other hand, from (12b)-(12c), we obtain $|(\zeta_k)_i - (\zeta_k)_j| \le |(\zeta_{k+1})_i - (\zeta_{k+1})_j|$ for all k. Therefore, we obtain $(\zeta_k)_i = (\zeta_k)_j$ for all k and from (12a), we get $\mu = (\zeta_k)_i = (\zeta_k)_j$ for all k. From (12b)-(12c), this further implies that $\mu \in \partial f_i^\epsilon(x)$ for all $x \in [P_i^*, \infty)$ and that $\mu \in \partial f_j^\epsilon(x)$ for all $x \in (-\infty, P_j^*]$. Using this fact, one can construct an unbounded set of solutions to the ED problem in the following manner. First, fix all the components of P^* except i and j. Now pick any $x \in \mathbb{R}_{\ge 0}$ and consider $P_i^* + x$ and $P_j^* - x$. From what we have reasoned so far, all such points that we obtain by varying x are solutions to the ED problem as they satisfy (7). This contradicts the fact that $\mathcal{F}_{\text{ED}}^*$ is bounded.

In case (b), assume without loss of generality that $\mathbf{1}_n^\top \zeta_k \to \infty$ (the argument for $\mathbf{1}_n^\top \zeta_k \to -\infty$ follows similarly). As reasoned above, there exists $j \in \{1, \dots, n\}$ such that $P_j(t_k) \to -\infty$ and there exists $\mu \in \mathbb{R}$ such that $(\zeta_k)_j \le \mu$ for all $k \in \mathbb{Z}_{\ge 1}$. Using this fact, we upper bound the left hand side of the inequality (10) by

$$\begin{aligned}
&-\frac{\lambda_2(\mathsf{L}+\mathsf{L}^\top)}{2}\Big\|\zeta_k - \frac{1}{n}(\mathbf{1}_n^\top\zeta_k)\mathbf{1}_n\Big\|^2 + \frac{1}{nk}(\mathbf{1}_n^\top\zeta_k)\\
&\le -\frac{\lambda_2(\mathsf{L}+\mathsf{L}^\top)}{2}\Big((\zeta_k)_j - \frac{1}{n}(\mathbf{1}_n^\top\zeta_k)\Big)^2 + \frac{1}{nk}(\mathbf{1}_n^\top\zeta_k)\\
&\le -\frac{\lambda_2(\mathsf{L}+\mathsf{L}^\top)}{2}\Big(\mu - \frac{1}{n}(\mathbf{1}_n^\top\zeta_k)\Big)^2 + \frac{1}{nk}(\mathbf{1}_n^\top\zeta_k), \quad (13)
\end{aligned}$$

where the last inequality is valid for all but a finite number of k. Hence, as $\mathbf{1}_n^\top \zeta_k \to \infty$, there is $\bar{k} \in \mathbb{Z}_{\ge 1}$ such that the expression in (13) is negative for $k \ge \bar{k}$, contradicting (10). Thus, we conclude the trajectories are bounded. □

From the proof above, it is interesting to note that the feedback term (8) drives the mismatch between generation and load to zero at an exponential rate, no matter what the initial power allocation. This is a good indication of its robustness properties: time-varying loads or scenarios with generators going down and coming back online can be handled as long as the rate of these changes is lower than the exponential rate of convergence associated to the load satisfaction. We provide a formal characterization of these properties for the distributed implementation of this strategy in the next section.

5 Robust distributed algorithmic solution

This section presents a distributed strategy to solve the ED problem starting from any initial power allocation. We build on the centralized design presented in Section 4. We also formally characterize the robustness properties against addition and deletion of generators and time-varying loads.

Given the discussion on the centralized nature of the dynamics (8), the core idea of our design is to employ a dynamic average consensus algorithm that allows each unit in the network to estimate the mismatch in load satisfaction. To this end, we assume the total load P_l is only known to one generator $r \in \{1, \dots, n\}$ (its specific identity is arbitrary). Following Section 2.2, consider the dynamics,

$$\begin{aligned}
\dot{z} &= -\alpha z - \beta \mathsf{L} z - v + \nu_2(P_l e_r - P),\\
\dot{v} &= \alpha\beta \mathsf{L} z,
\end{aligned}$$

where $e_r \in \mathbb{R}^n$ is the unit vector along the r-th direction and $\alpha, \beta, \nu_2 > 0$ are design parameters. Note that this algorithm is distributed over the communication graph $\mathcal{G}$. For each $i \in \{1, \dots, n\}$, z_i plays the role of an estimator associated to i which aims to track the average signal $t \mapsto \frac{1}{n}(P_l - \mathbf{1}_n^\top P(t))$. This observation justifies substituting the feedback term in (8) by $z \in \mathbb{R}^n$, giving rise to the "dynamic average consensus + Laplacian-nonsmooth-gradient" dynamics, abbreviated $\texttt{dac+L}\partial$ for convenience, mathematically represented by the set-valued map $X_{\texttt{dac+L}\partial} : \mathbb{R}^{3n} \rightrightarrows \mathbb{R}^{3n}$,

$$\dot{P} \in -\mathsf{L}\partial f^\epsilon(P) + \nu_1 z, \quad (15a)$$
$$\dot{z} = -\alpha z - \beta \mathsf{L} z - v + \nu_2(P_l e_r - P), \quad (15b)$$
$$\dot{v} = \alpha\beta \mathsf{L} z, \quad (15c)$$

where $\nu_1 > 0$ is a design parameter. Unlike (8), this algorithm is distributed, as each agent only needs to interact with its neighbors to implement it.

Remark 5.1 *(Comparison with finite-time initialization approach):* Our previous distributed solution for the ED problem [Cherukuri and Cortés, 2015] involves initializing the generation levels inside the feasibility set $\mathcal{F}_{\text{ED}}$ using a finite-time message passing algorithm. This finite-time initialization approach is best suited for scenarios with a fixed set of participating generators. In the presence of intermittent generation, every time a generator joins or leaves the network, the generators have to stop the dynamics, execute the finite-time algorithm, and then re-run the dynamics. This approach, however, cannot deliver perfect tracking of continuously time-varying loads. In contrast, the $\texttt{dac+L}\partial$ dynamics does not suffer from these limitations, as discussed later in Section 5.2. •

5.1 Convergence analysis

Here we characterize the asymptotic convergence properties of the $\texttt{dac+L}\partial$ dynamics. We start by establishing an important fact on the omega-limit set of any trajectory of (15) with initial condition in $\mathbb{R}^n \times \mathbb{R}^n \times \mathcal{H}_0$.

Lemma 5.2 *(Characterizing the omega-limit set of the trajectories of the $\texttt{dac+L}\partial$ dynamics): The omega-limit set of any trajectory of* (15) *with initial condition* $(P_0, z_0, v_0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{H}_0$ *is contained in* $\mathcal{H}_{P_l} \times \mathcal{H}_0 \times \mathcal{H}_0$.

PROOF. From (15c), note that $\mathbf{1}_n^\top \dot{v} = 0$. Since $v_0 \in \mathcal{H}_0$, this implies that $\mathbf{1}_n^\top v(t) = 0$ for all $t \ge 0$. Now,

define $\zeta(t) = \mathbf{1}_n^\top P(t) - P_l$ and note that

$$\dot{\zeta}(t) = \mathbf{1}_n^\top \dot{P}(t) = \nu_1 \mathbf{1}_n^\top z(t),$$

where we have used (15a), and

$$\begin{aligned}\ddot{\zeta}(t) &= \nu_1 \mathbf{1}_n^\top \dot{z}(t) \\ &= \nu_1 \mathbf{1}_n^\top (-\alpha z(t) - \beta \mathsf{L} z(t) - v(t) + \nu_2(P_l e_k - P(t)) \\ &= -\nu_1 \alpha(\mathbf{1}_n^\top z(t)) - \nu_1\nu_2 \zeta(t) = -\alpha\dot{\zeta}(t) - \nu_1\nu_2\zeta(t),\end{aligned}$$

where we have used (15b). We write this system as a first-order one by defining $x_1 = \zeta$ and $x_2 = \dot{\zeta}$ to get

$$\begin{bmatrix}\dot{x}_1 \\ \dot{x}_2\end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\nu_1\nu_2 & -\alpha\end{bmatrix}\begin{bmatrix}x_1 \\ x_2\end{bmatrix}. \tag{16}$$

Evaluating the Lie derivative of the positive definite, radially unbounded function $V_2(x_1, x_2) = \nu_1\nu_2 x_1^2 + x_2^2$ along the above dynamics and applying the LaSalle Invariance Principle [Khalil, 2002], we deduce that $x_1(t) \to 0$ and $x_2(t) \to 0$ as $t \to \infty$, that is, $\mathbf{1}_n^\top P(t) \to P_l$ and $\mathbf{1}_n^\top z(t) \to 0$. Since the system (16) is linear, the convergence is exponential. □

The next result builds on this fact and Proposition A.1 to establish that the trajectory of power allocations under (15) converges to the solution set of the ED problem.

Theorem 5.3 *(Convergence of the* $\mathsf{dac+L\partial}$ *dynamics to the solutions of* ED *problem): For* $\alpha, \beta, \nu_1, \nu_2 > 0$ *with*

$$\frac{\nu_1}{\beta\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)} + \frac{\nu_2^2\lambda_{\max}(\mathsf{L}^\top\mathsf{L})}{2\alpha} < \lambda_2(\mathsf{L}+\mathsf{L}^\top), \tag{17}$$

the trajectories of (15) *starting from any point in* $\mathbb{R}^n \times \mathbb{R}^n \times \mathcal{H}_0$ *converge to the set* $\mathcal{F}^*_{\text{aug}} = \{(P, z, v) \in \mathcal{F}^*_{\text{ED}} \times \{0\} \times \mathbb{R}^n \mid v = \nu_2(P_l e_r - P)\}$.

PROOF. Our proof strategy is based on the refined LaSalle Invariance Principle for differential inclusions established in Appendix A, cf. Proposition A.1. Before justifying that all its hypotheses are satisfied, we reformulate the expression for the dynamics to help simplify the analysis. Consider first the change of coordinates, $(P, z, v) \mapsto (P, z, \bar{v})$, with $\bar{v} = v - \nu_2(P_l e_r - P)$. The set-valued map $X_{\mathsf{dac+L\partial}}$ then takes the form

$$\begin{aligned}X_{\mathsf{dac+L\partial}}(P, z, \bar{v}) = \{&(-\mathsf{L}\zeta + \nu_1 z, -(\alpha + \beta\mathsf{L})z - \bar{v}, \\ &(\alpha\beta\mathsf{L} + \nu_1\nu_2)z - \nu_2\mathsf{L}\zeta) \in \mathbb{R}^{3n} \mid \zeta \in \partial f^\epsilon(P)\}.\end{aligned}$$

The change of coordinates shifts the equilibrium of the consensus dynamics to the origin. Under the additional change of coordinates $(P, z, \bar{v}) \mapsto (P, \xi_1, \xi_2)$, with

$$\begin{bmatrix}\xi_1 \\ \xi_2\end{bmatrix} = \begin{bmatrix} I & 0 \\ \alpha I & I\end{bmatrix}\begin{bmatrix} z \\ \bar{v}\end{bmatrix}, \tag{18}$$

the set-valued map $X_{\mathsf{dac+L\partial}}$ takes the form

$$\begin{aligned}X_{\mathsf{dac+L\partial}}(P, \xi_1, \xi_2) = \{&(-\mathsf{L}\zeta + \nu_1\xi_1, -\beta\mathsf{L}\xi_1 - \xi_2, \\ &\nu_1\nu_2\xi_1 - \alpha\xi_2 - \nu_2\mathsf{L}\zeta) \in \mathbb{R}^{3n} \mid \zeta \in \partial f^\epsilon(P)\}.\end{aligned} \tag{19}$$

This extra change of coordinates makes it easier to identify the candidate Lyapunov function $V_3 : \mathbb{R}^{3n} \to \mathbb{R}_{\geq 0}$,

$$V_3(P, \xi_1, \xi_2) = f^\epsilon(P) + \frac{1}{2}(\nu_1\nu_2\|\xi_1\|^2 + \|\xi_2\|^2).$$

For convenience, denote the overall change of coordinates by $D : \mathbb{R}^{3n} \to \mathbb{R}^{3n}$,

$$(P, \xi_1, \xi_2) = D(P, z, v) = (P, z, v + \alpha z - \nu_2(P_l e_r - P)).$$

Our analysis now focuses on proving that, in the new coordinates, the trajectories of (15) converge to the set

$$\overline{\mathcal{F}}_{\text{aug}} = D(\mathcal{F}^*_{\text{aug}}) = \mathcal{F}^*_{\text{ED}} \times \{0\} \times \{0\}.$$

Note that $D(\mathcal{H}_{P_l} \times \mathcal{H}_0 \times \mathcal{H}_0) = \mathcal{H}_{P_l} \times \mathcal{H}_0 \times \mathcal{H}_0$ and therefore, from Lemma 5.2, the omega-limit set of a trajectory $t \mapsto (P(t), \xi_1(t), \xi_2(t))$ starting in $D(\mathbb{R}^n \times \mathbb{R}^n \times \mathcal{H}_0)$ belongs to $\mathcal{H}_{P_l} \times \mathcal{H}_0 \times \mathcal{H}_0$.

Our next step is to show that the hypotheses of Proposition A.1 are satisfied where $\mathcal{H}_{P_l} \times \mathcal{H}_0 \times \mathcal{H}_0$ and V_3 play the role of the closed submanifold $\mathcal{S}$ of $\mathbb{R}^{3n}$ and the function W, respectively. To do so, we resort to Lemma A.1. Define the continuous function $g : \mathbb{R}^{3n} \times \mathbb{R}^{3n} \to \mathbb{R}^{3n}$ by

$$\begin{aligned}g(P, \xi_1, \xi_2, \hat{\zeta}) = (&-\mathsf{L}\hat{\zeta}_1 + \nu_1\xi_1, -\beta\mathsf{L}\xi_1 - \xi_2, \\ &\nu_1\nu_2\xi_1 - \alpha\xi_2 - \nu_2\mathsf{L}\hat{\zeta}_1),\end{aligned}$$

and note that the dynamics (19) can be expressed as $X_{\mathsf{dac+L\partial}}(P, \xi_1, \xi_2) = \{g(P, \xi_1, \xi_2, \hat{\zeta}) \mid \hat{\zeta} \in \partial V_3(P, \xi_1, \xi_2)\}$. For $(P, \xi_1, \xi_2) \in \mathcal{H}_{P_l} \times \mathcal{H}_0 \times \mathcal{H}_0$ and $\hat{\zeta} \in \partial V_3(P, \xi_1, \xi_2)$,

$$\begin{aligned}\hat{\zeta}^\top g(P, \xi_1, \xi_2, \hat{\zeta}) = &-\zeta^\top\mathsf{L}\zeta + \nu_1\zeta^\top\xi_1 - \beta\nu_1\nu_2\xi_1^\top L\xi_1 \\ &- \alpha\|\xi_2\|^2 - \nu_2\xi_2^\top L\zeta,\end{aligned} \tag{20}$$

where we have used that $\zeta = \hat{\zeta}_1 \in \partial f^\epsilon(P)$, $\hat{\zeta}_2 = \nu_1\nu_2\xi_1$, and $\hat{\zeta}_3 = \xi_2$. Since the digraph $\mathcal{G}$ is strongly connected and weight-balanced, we apply (1) and the fact that

$\mathbf{1}_n^\top \xi_1 = 0$ to bound the above expression as

$$-\frac{1}{2}\lambda_2(\mathsf{L}+\mathsf{L}^\top)\|\eta\|^2+\nu_1\eta^\top\xi_1-\frac{1}{2}\beta\nu_1\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)\|\xi_1\|^2 - \alpha\|\xi_2\|^2 - \nu_2\xi_2^\top L\eta = \gamma^\top M\gamma,$$

where $\eta = \zeta - \frac{1}{n}(\mathbf{1}_n^\top\zeta)\mathbf{1}_n$, $\gamma^\top = [\eta^\top, \xi_1^\top, \xi_2^\top]$, and

$$M = \begin{bmatrix} -\frac{1}{2}\lambda_2(\mathsf{L}+\mathsf{L}^\top)I & \frac{1}{2}\nu_1 I & -\frac{1}{2}\nu_2\mathsf{L}^\top \\ \frac{1}{2}\nu_1 I & -\frac{1}{2}\beta\nu_1\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)I & 0 \\ -\frac{1}{2}\nu_2\mathsf{L} & 0 & -\alpha I \end{bmatrix}.$$

Reasoning with the Schur complement [Boyd and Vandenberghe, 2009], $M \in \mathbb{R}^{3n\times 3n}$ is negative definite if

$$-\frac{1}{2}\lambda_2(\mathsf{L}+\mathsf{L}^\top)I- \begin{bmatrix}\frac{1}{2}\nu_1 I & -\frac{1}{2}\nu_2\mathsf{L}^\top\end{bmatrix}\begin{bmatrix}-\frac{1}{2}\beta\nu_1\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)I & 0 \\ 0 & -\alpha I\end{bmatrix}^{-1}\begin{bmatrix}\frac{1}{2}\nu_1 I \\ -\frac{1}{2}\nu_2\mathsf{L}\end{bmatrix}$$
$$= -\frac{1}{2}\lambda_2(\mathsf{L}+\mathsf{L}^\top)I + \frac{\nu_1}{2\beta\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)}I + \frac{\nu_2^2}{4\alpha}\mathsf{L}^\top\mathsf{L},$$

is negative definite. This latter fact is implied by (17). As a consequence, $\hat{\zeta}^\top g(P,\xi_1,\xi_2,\hat{\zeta}) \leq 0$ and so, Lemma A.1(i) holds. Moreover, $\hat{\zeta}^\top g(P,\xi_1,\xi_2,\hat{\zeta}) = 0$ if and only if $\eta = \xi_1 = \xi_2 = 0$, which means $\zeta \in \operatorname{span}\{\mathbf{1}_n\}$. Using this fact along with the definition of the set-valued Lie derivative and the characterization of optimizers (7), we deduce that $\hat{\zeta}^\top g(P,\xi_1,\xi_2,\hat{\zeta}) = 0$ if and only if (a) $0 \in \mathcal{L}_{X_{\mathsf{dac+L}\partial}} V_3(P,\xi_1,\xi_2)$ and (b) P is a solution of the ED problem. Fact (a) implies that Lemma A.1(ii) holds and hence, Proposition A.1(ii) holds too. Fact (b) implies that over the set $\mathcal{H}_{P_l}\times\mathcal{H}_0\times\mathcal{H}_0$, we have $0 \in \mathcal{L}_{X_{\mathsf{dac+L}\partial}} V_3(P,\xi_1,\xi_2)$ if and only if $(P,\xi_1,\xi_2) \in \overline{\mathcal{F}}_{\text{aug}}$. Since, $\overline{\mathcal{F}}_{\text{aug}}$ belongs to a level set of V_3, we conclude that Proposition A.1(i) holds too.

To be able to apply Proposition A.1 and conclude the proof, it remains to show that the trajectories starting from $D(\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{H}_0)$ are bounded. We reason by contradiction, i.e., assume there exists a trajectory $t \mapsto (P(t),\xi_1(t),\xi_2(t))$, with initial condition $(P(0),\xi_1(0),\xi_2(0)) \in D(\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{H}_0)$ of $X_{\mathsf{dac+L}\partial}$ such that $\|(P(t),\xi_1(t),\xi_2(t)\| \to \infty$. Since V_3 is radially unbounded, this implies $V_3(P(t),\xi_1(t),\xi_2(t)) \to \infty$. Additionally, from Lemma 5.2, we know that $\mathbf{1}_n^\top P(t) \to P_l$ and $\mathbf{1}_n^\top\xi_1(t) \to 0$. Thus, there exists a sequence of times $\{t_k\}_{k=1}^\infty$ with $t_k \to \infty$ such that for all $k \in \mathbb{Z}_{\geq 1}$,

$$\left|\mathbf{1}_n^\top\xi_1(t_k)\right| < 1/k, \tag{21a}$$
$$\max \mathcal{L}_{X_{\mathsf{dac+L}\partial}} V_3(P(t_k),\xi_1(t_k),\xi_2(t_k)) > 0. \tag{21b}$$

Note that (21b) implies that there exists a sequence $\{\zeta_k\}_{k=1}^\infty$ with $\zeta_k \in \partial f^\epsilon(P(t_k))$ such that

$$-\zeta_k^\top\mathsf{L}\zeta_k + \nu_1\zeta_k^\top\xi_1(t_k) - \beta\nu_1\nu_2\xi_1(t_k)^\top\mathsf{L}\xi_1(t_k) - \alpha\|\xi_1(t_k)\|^2 - \nu_2\xi_2(t_k)^\top\mathsf{L}\zeta_k > 0,$$

for all $k \in \mathbb{Z}_{\geq 1}$, where we have used the fact that an element of $\mathcal{L}_{X_{\mathsf{dac+L}\partial}} V_3(P,\xi_1,\xi_2)$ has the form given in (20). Letting $\eta_k = \zeta_k - \frac{1}{n}(\mathbf{1}_n^\top\zeta_k)\mathbf{1}_n$, we use (1) to deduce from the above inequality that

$$-\frac{1}{2}\lambda_2(\mathsf{L}+\mathsf{L}^\top)\|\eta_k\|^2 + \nu_1\eta_k^\top\xi_1(t_k) + \frac{1}{n}\nu_1(\mathbf{1}_n^\top\zeta_k)(\mathbf{1}_n^\top\xi_1(t_k)) - \frac{1}{2}\beta\nu_1\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)\|\xi_1(t_k) - \frac{1}{n}(\mathbf{1}_n^\top\xi_1(t_k))\mathbf{1}_n\|^2 - \alpha\|\xi_1(t_k)\|^2 - \nu_2\xi_2(t_k)^\top\mathsf{L}\eta_k > 0.$$

Further, using the expression

$$\|\xi_1(t_k) - \frac{1}{n}(\mathbf{1}_n^\top\xi_1(t_k))\mathbf{1}_n\|^2 = \|\xi_1(t_k)\|^2 - \frac{1}{n}(\mathbf{1}_n^\top\xi_1(t_k))^2,$$

the inequality can be rewritten as

$$\gamma_k^\top M\gamma_k + \frac{1}{n}\nu_1(\mathbf{1}_n^\top\zeta_k)(\mathbf{1}_n^\top\xi_1(t_k)) + \frac{\beta\nu_1\nu_2}{2n}\lambda_2(\mathsf{L}+\mathsf{L}^\top)(\mathbf{1}_n^\top\xi_1(t_k))^2 > 0,$$

where $\gamma_k^\top = [\eta_k^\top, \xi_1(t_k)^\top, \xi_2(t_k)^\top]$. Using now the bound (21a), we arrive at the inequality,

$$\gamma_k^\top M\gamma_k + \frac{\nu_1}{nk}\left|\mathbf{1}_n^\top\zeta_k\right| + \frac{\beta\nu_1\nu_2}{2nk^2}\lambda_2(\mathsf{L}+\mathsf{L}^\top) > 0. \tag{22}$$

Next, we consider two cases, depending on whether the sequence $\{P(t_k)\}$ is (a) bounded or (b) unbounded. In case (a), the sequence $\{(\xi_1(t_k),\xi_2(t_k))\}$ must be unbounded. Since M is negative definite, we have $\gamma_k^\top M\gamma_k \leq \lambda_{\max}(M)\|(\xi_1(t_k),\xi_2(t_k))\|^2$. Thus, (22) implies that

$$\lambda_{\max}(M)\|(\xi_1(t_k),\xi_2(t_k))\|^2 + \frac{\nu_1}{nk}\left|\mathbf{1}_n^\top\zeta_k\right| + \frac{\beta\nu_1\nu_2}{2nk^2}\lambda_2(\mathsf{L}+\mathsf{L}^\top) > 0.$$

Now, from the expression of ∂f^ϵ, since $\{P(t_k)\}$ is bounded, the sequence $\{\zeta_k\}$ must be bounded. Combining these facts with $\lambda_{\max}(M) < 0$, one can find $\bar{k} \in \mathbb{Z}_{\geq 1}$ such that the above inequality is violated for all $k \geq \bar{k}$, which is a contradiction. For case (b), we use the bound $\gamma_k^\top M\gamma_k \leq \lambda_{\max}(M)\|\eta_k\|^2$ to deduce from (22) that

$$\lambda_{\max}(M)\|\eta_k\|^2 + \frac{\nu_1}{nk}\left|\mathbf{1}_n^\top\zeta_k\right| + \frac{\beta\nu_1\nu_2}{2nk^2}\lambda_2(\mathsf{L}+\mathsf{L}^\top) > 0.$$

One can then use a similar argument as laid out in the proof of Theorem 4.1, considering the two cases of $\left|\mathbf{1}_n^\top \zeta_k\right|$ being bounded or unbounded, arriving in both cases at similar contradictions. This concludes the proof. □

Note that as a consequence of the above result, the dac+L∂ dynamics do not require any specific preprocessing for the initialization of the power allocations. Each generator can select any generation level, independent of the other units, and the algorithm guarantees convergence to the solutions of the ED problem.

Remark 5.4 *(Distributed selection of algorithm design parameters):* The convergence of the dac+L∂ dynamics relies on a selection of the parameters α, β, ν_1 and $\nu_2 \in \mathbb{R}_{>0}$ that satisfy (17). Checking this inequality requires knowledge of the spectrum of matrices related to the Laplacian matrix, and hence the entire network structure. Here, we provide an alternative condition that implies (17) and can be checked by the units in a distributed way. Let $n_{\max}$ be an upper bound on the number of units, $d^{\text{out}}_{\max}$ be an upper bound on the out-degree of all units, and $a_{\min}$ be a lower bound on the edge weights,

$$n \leq n_{\max}, \ \max_{i\in\mathcal{V}} d^{\text{out}}(i) \leq d^{\text{out}}_{\max}, \ \min_{(i,j)\in\mathcal{E}} a_{ij} \geq a_{\min}. \quad (23)$$

A straightforward generalization of [Mohar, 1991, Theorem 4.2] for weighted graphs gives rise to the following lower bound on $\lambda_2(\mathsf{L}+\mathsf{L}^\top)$,

$$\frac{4a_{\min}}{n^2_{\max}} \leq \lambda_2(\mathsf{L}+\mathsf{L}^\top). \quad (24)$$

On the other hand, using properties of matrix norms [Bernstein, 2005, Chapter 9], one can deduce

$$\begin{aligned}\lambda_{\max}(\mathsf{L}^\top\mathsf{L}) &= \|\mathsf{L}\|^2 \leq (\sqrt{n}\|\mathsf{L}\|_\infty)^2 \\ &\leq (2\sqrt{n}d^{\text{out}}_{\max})^2 \leq 4n_{\max}(d^{\text{out}}_{\max})^2. \quad (25)\end{aligned}$$

Using (24)-(25), the left-hand side of (17) can be upper bounded by

$$\begin{aligned}&\frac{\nu_1}{\beta\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)} + \frac{\nu_2^2\lambda_{\max}(L^\top L)}{2\alpha} \\ &\qquad \leq \frac{\nu_1 n^2_{\max}}{4a_{\min}\beta\nu_2} + \frac{2\nu_2^2 n_{\max}(d^{\text{out}}_{\max})^2}{\alpha}.\end{aligned}$$

Further, the right-hand side of (17) can be lower bounded using (24). Putting the two together, we obtain the new condition

$$\frac{\nu_1 n^2_{\max}}{4a_{\min}\beta\nu_2} + \frac{2\nu_2^2 n_{\max}(d^{\text{out}}_{\max})^2}{\alpha} < \frac{4a_{\min}}{n^2_{\max}}, \quad (26)$$

which implies (17). The network can ensure that this condition is met in various ways. For instance, if the bounds $n_{\max}$, $d^{\text{out}}_{\max}$, and $a_{\min}$ are not available, the network can implement distributed algorithms for max- and min-consensus [Ren and Beard, 2008] to compute them in finite time. Once known, any generator can select α, β, ν_1 and ν_2 satisfying (26) and broadcast its choice. Alternatively, the computation of the design parameters can be implemented concurrently with the determination of the bounds via consensus by specifying a specific formula to select them that is guaranteed to satisfy (26). Note that the units necessarily need to agree on the parameters, otherwise if each unit selects a different set of parameters, the dynamic average consensus would not track the average input signal. •

Remark 5.5 *(Distributed loads and transmission losses):* Here we expand on our observations in Remark 3.1 regarding the inclusion of additional constraints on the ED problem. Our algorithmic solution can be easily modified to deal with the alternative scenarios studied in [Zhang et al., 2011, Kar and Hug, 2012, Binetti et al., 2014a, Loia and Vaccaro, 2013], where each generator has the knowledge of the load at the corresponding bus that it is connected to and the total load is the aggregate of these individual loads. Mathematically, denoting the load demanded at generator bus i by $P_i^L \in \mathbb{R}$, the total load is given by $P_l = \sum_{i=1}^n P_i^L$. For this case, replacing the vector $P_l e_r$ by P^L in the dac+L∂ dynamics (15b) gives an algorithm that solves the ED problem for the load P_l. Our solution strategy can also handle transmission losses as modeled in [Binetti et al., 2014a], where it is assumed that each generator i can estimate the power loss in the transmission lines adjacent to it. With those values available, the generator could add them to the quantity P_i^L, which would make the network find a power allocation that takes care of the transmission losses. •

5.2 Robustness analysis

In this section, we study the robustness properties of the dac+L∂ dynamics in the presence of time-varying load signals and intermittent power unit generation. Our analysis relies on the exponential stability of the mismatch dynamics between total generation and load established in Lemma 5.2, which implies that (16) is input-to-state stable (ISS) [Khalil, 2002, Lemma 4.6], and consequently robust against arbitrary bounded perturbations. The following result provides an explicit, exponentially decaying, bound for the evolution of any trajectory of (16). While the rate of decay can also be determined by computing the eigenvalues of matrix defining the dynamics, here we employ a Lyapunov argument to obtain also the value of the gain associated to the rate.

Lemma 5.6 *(Convergence rate of the mismatch dynamics* (16)*): Let $R \in \mathbb{R}^{2\times 2}$ be defined by*

$$R = \frac{1}{2\alpha\nu_1\nu_2}\begin{bmatrix}\alpha^2+\nu_1\nu_2+(\nu_1\nu_2)^2 & \alpha \\ \alpha & 1+\nu_1\nu_2\end{bmatrix}.$$

Then $R \succ 0$ and any trajectory $t \mapsto x(t)$ of the dynamics (16) satisfies $\|x(t)\| \leq c_1 e^{-c_2 t}\|x(0)\|$, where $c_1 = \sqrt{\lambda_{\max}(R)/\lambda_{\min}(R)}$ and $c_2 = 1/2\lambda_{\max}(R)$.

PROOF. Let $A \in \mathbb{R}^{2\times 2}$ be the system matrix of (16). Then, one can see that $A^\top R + RA = -I$, i.e., $V_4(x) = x^\top Rx$ is a Lyapunov function for (16). Note that

$$\lambda_{\min}(R)\|x\|^2 \leq V_4(x) \leq \lambda_{\max}(R)\|x\|^2. \qquad (27)$$

From the Lyapunov equation, we have $\mathcal{L}_{Ax}V_4(x) = -\|x\|^2 \leq -\frac{1}{\lambda_{\max}(R)}V_4(x)$, which implies $V_4(x(t)) \leq e^{-1/\lambda_{\max}(R)}V_4(x(0))$ along any trajectory $t \mapsto x(t)$ of (16). Again using (27), we get

$$\|x(t)\|^2 \leq \frac{\lambda_{\max}(R)}{\lambda_{\min}(R)} e^{-1/\lambda_{\max}(R)}\|x(0)\|^2,$$

which concludes the claim. □

In the above result, it is interesting to note that the convergence rate is independent of the specific communication digraph (as long as it is weight-balanced). We use next the exponentially decaying bound obtained above to illustrate the extent to which the network can collectively track a dynamic load (which corresponds to a time-varying perturbation in the mismatch dynamics) and is robust to intermittent power generation (which corresponds to perturbations in the state of the mismatch dynamics).

5.2.1 Tracking dynamic loads

Here we consider a time-varying total load given by a twice continuously differentiable trajectory $\mathbb{R}_{\geq 0} \ni t \mapsto P_l(t)$ and show how the total generation of the network under the dac+L∂ dynamics tracks it. We assume the signal is known to an arbitrary unit $r \in \{1, \dots, n\}$. In this case, the dynamics (16) take the following form

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ -\nu_1\nu_2 & -\alpha\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ -\alpha\dot{P}_l - \ddot{P}_l\end{bmatrix}.$$

Using Lemma 5.6, one can compute the following bound on any trajectory of the above system

$$\|x(t)\| \leq c_1 e^{-c_2 t}\|x(0)\| + \frac{c_1}{c_2}\sup_{s\in[0,t]}\left|\alpha\dot{P}_l(s) + \ddot{P}_l(s)\right|.$$

In particular, for a signal with bounded $\dot{P}_l$ and $\ddot{P}_l$, the mismatch between generation and load, i.e., $x_1(t)$ is bounded. Also, the mismatch has an ultimate bound as $t \to \infty$. The following result summarizes this notion formally. The proof is straightforward application of Lemma 5.6 following the exposition of input-to-state stability in [Khalil, 2002].

Proposition 5.7 *(Power mismatch is ultimately bounded for dynamic load under the dac+L∂ dynamics): Let $\mathbb{R}_{\geq 0} \ni t \mapsto P_l(t)$ be twice continuously differentiable and such that*

$$\sup_{t\geq 0}\left|\dot{P}_l(t)\right| \leq d_1, \quad \sup_{t\geq 0}\left|\ddot{P}_l(t)\right| \leq d_2,$$

for some $d_1, d_2 > 0$. Then, the mismatch $\mathbf{1}_n^\top P(t) - P_l(t)$ between load and generation is bounded along the trajectories of (15) and has ultimate bound $\frac{c_1}{c_2}(\alpha d_1 + d_2)$, with c_1, c_2 given in Lemma 5.6. Moreover, if $\dot{P}_l(t) \to 0$ and $\ddot{P}_l(t) \to 0$ as $t \to \infty$, then $\mathbf{1}_n^\top P(t) \to P_l(t)$ as $t \to \infty$.

5.2.2 Robustness to intermittent power generation

Here, we characterize the algorithm robustness against unit addition and deletion to capture scenarios with intermittent power generation. Addition and deletion events are modeled via a time-varying communication digraph, which we assume remains strongly connected and weight-balanced at all times. When a unit stops generating power (deletion event), the corresponding vertex and its adjacent edges are removed. When a unit starts providing power (addition event), the corresponding node is added to the digraph along with a set of edges. Given the intricacies of the convergence analysis for the dac+L∂ dynamics, cf. Theorem 5.3, it is important to make sure that the state v remains in the set $\mathcal{H}_0$, irrespectively of the discontinuities caused by the events. The following routine makes sure that this is the case.

> TRAJECTORY INVARIANCE: When a unit i joins the network at time t, it starts with $v_i(t) = 0$. When a unit i leaves the network at time t, it passes a token with value $v_i(t)$ to one of its in-neighbors $j \in \mathcal{N}^{\text{in}}(i)$, who resets its value to $v_j(t) + v_i(t)$.

The TRAJECTORY INVARIANCE routine ensures that the dynamics (16) are the appropriate description for the evolution of the load satisfaction mismatch. This, together with the ISS property established in Lemma 5.6, implies that the mismatch effect in power generation caused by addition/deletion events vanishes exponentially fast. In particular, if the number of addition/deletion events is finite, then the set of generators converge to the solution of the ED problem. We formalize this next.

Proposition 5.8 *(Convergence of the dac+L∂ dynamics under intermittent power generation): Let n_{max} be the maximum number of generators that can contribute to the power generation at any time. Let $\Sigma_{n_{max}}$ be the set of digraphs that are strongly connected and weight-balanced and whose vertex set is included in $\{1, \dots, n_{max}\}$. Let $\sigma : [0, \infty) \to \Sigma_{n_{max}}$ be a piecewise constant, right-continuous switching signal described by the set of switching times $\{t_1, t_2, \dots\} \subset \mathbb{R}_{\geq 0}$, with $t_k \leq t_{k+1}$, each corresponding to either an addition or a deletion event. Denote by $X^\sigma_{\texttt{dac+L}\partial}$ the switching dac+L∂ dynamics corresponding to σ, defined by (15) with L replaced by L$(\sigma(t))$ for all $t \geq 0$, and assume agents execute the* TRAJECTORY IN-

VARIANCE *routine when they leave or join the network. Then,*

(i) at any time $t \in \{0\} \cup \{t_1, t_2, \dots\}$, if the variables $(P(t), z(t))$ for the generators in $\sigma(t)$ satisfy $\left|\mathbf{1}_n^\top P(t) - P_l\right| \leq M_1$ and $\left|\mathbf{1}_n^\top z(t)\right| \leq M_2$ for some $M_1, M_2 > 0$, then the magnitude of the mismatch between generation and load becomes less than or equal to $\rho > 0$ in time

$$t_\rho = \frac{1}{c_2} \ln\Big(\frac{c_1(M_1 + \nu_1 M_2)}{\rho}\Big),$$

provided no event occurs in the interval $(t, t + t_\rho)$;

(ii) if the number of events is finite, say N, then the trajectories of $X^\sigma_{\texttt{dac+L}\partial}$ converge to the set of solutions of the ED *problem for the group of generators in $\sigma(t_N)$ provided* (17) *is met for $\sigma(t_N)$.*

Note that the generators can ensure that the condition (17), required for the convergence of the dac+L∂ dynamics, holds at all times even under addition and deletion events, if they rely on verifying that (26) holds and the bounds (23) are valid for all the topologies in $\Sigma_{n_{\max}}$.

6 Simulations in a IEEE 118 bus system

This section illustrates the convergence of the dac+L∂ dynamics to the solutions of the ED problem (4) starting from any initial power allocation and its robustness properties. We consider the IEEE 118 bus system [IEEE 118 bus], that consists of 54 generators. The cost function of each generator i is quadratic, $f_i(P_i) = a_i + b_i P_i + c_i P_i^2$, with coefficients belonging to the ranges $a_i \in [6.78, 74.33]$, $b_i \in [8.3391, 37.6968]$, and $c_i \in [0.0024, 0.0697]$. The communication topology is the digraph $\mathcal{G}$ described in Table 1. We choose the design parameters as $\nu_1 = 1, \nu_2 = 1.3, \alpha = 10, \beta = 40, \epsilon = 0.0086$, which satisfy the conditions (6) and (17) for $\mathcal{G}$. The total load is 4600 for the first 150 seconds and 4200 for the next 150 seconds, and is known to unit 3. Figure 1(a)-(c) depicts the evolution of the power allocation, total cost, and the mismatch between the total generation and load under the dac+L∂ dynamics starting at the initial condition $(P(0), z(0), v(0)) = (0.5 * (P^m + P^M), 0, 0)$. Note that the generators initially converge to a power allocation that meets the load 4600 and minimizes the total cost of generation. Later, with the decrease in desired load to 4200, the network decreases the total generation while minimizing the total cost.

Next, we consider a time-varying total load given by a constant plus a sinusoid, $P_l(t) = 4300 + 100 \sin(0.05t)$. With the same communication topology, design parameters, and initial condition as above, Figure 1 (d)-(f) illustrates the behavior of the network under the dac+L∂ dynamics. As established in Proposition 5.7, the total generation tracks the time-varying load signal and the mismatch between these values has an ultimate bound. Additionally, to illustrate how that the mismatch vanishes if the load becomes constant, we show in Figure 2 a load signal that consists of short bursts of sinusoidal

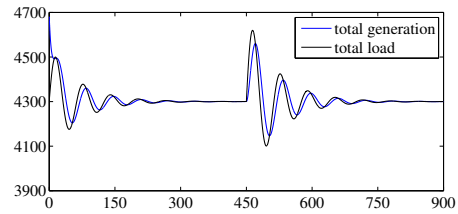


Fig. 2. Evolution of the total power generation for the IEEE 118 bus example under the dac+L∂ dynamics for the communication digraph $\mathcal{G}$, design parameters $\nu_1 = 1$, $\nu_2 = 1.3$, $\alpha = 10$, $\beta = 40$ and $\epsilon = 0.0086$, and time-varying total load. The example depicts the input-to-state stability of the mismatch dynamics.

variation that decay exponentially. The difference between generation and load becomes smaller and smaller as the load tends towards a constant signal.

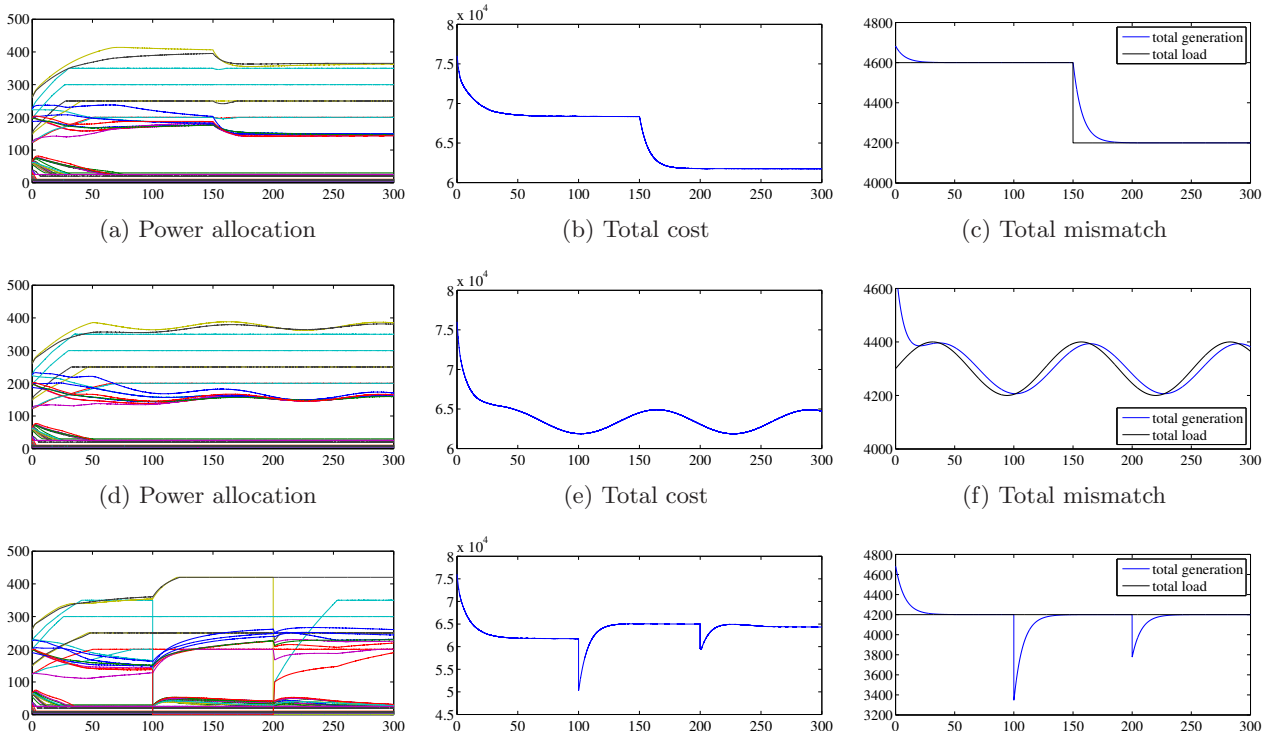
Our final scenario considers addition and deletion of generators. The initial communication topology is the undirected graph $\hat{\mathcal{G}}$ described in Table 1. The design parameters and the initial condition are the same as above. The total load is 4200 and is same at all times. For the first 100 seconds, the power allocations converge to a neighborhood of a solution of the ED problem for the set of generators in $\hat{\mathcal{G}}$. At time $t = 100s$, the units $\{4, 11, 25, 45\}$ stop generating power and leave the network. We select these generators because of their substantial impact in the total power generation. After this event, the resulting communication graph is $\hat{\mathcal{G}}_{\setminus\{4,11,25,45\}}$, cf. Table 1. The generators implement the TRAJECTORY INVARIANCE routine, after which the dac+L∂ algorithm drives the mismatch to zero and minimizes the total cost. At $t_2 = 200s$, another event occurs, the units $\{11, 45\}$ get added to the network while the generator 27 leaves. The resulting communication topology is $\hat{\mathcal{G}}_{\setminus\{4,25,27\}}$, cf. Table 1. After executing the TRAJECTORY INVARIANCE routine, the algorithm converges eventually to the optimizers of the ED problem for the set of generators in $\hat{\mathcal{G}}_{\setminus\{4,25,27\}}$, as shown in Figure 1(g)-(i). This example illustrates the robustness of the dac+L∂ dynamics against intermittent generation by the units, as formally established in Proposition 5.8. In addition to the presented examples, we also successfully simulated scenarios of the kind described in Remark 5.5, where the total load is not known to a single generator and is instead the aggregate of the local loads connected to each of the generator buses, but we do not report them here for space reasons. We have also observed in multiple simulations that the dac+L∂ dynamics respects the box constraints along its trajectories if they are satisfied at the initial condition.

7 Conclusions

We have designed a novel provably-correct distributed strategy that allows a group of generators to solve the economic dispatch problem starting from any initial power allocation. Our algorithm design combines elements from average consensus to dynamically estimate the mismatch between generation and desired load

$\mathcal{G}$	digraph over 54 vertices consisting of a directed cycle through vertices $1, \ldots, 54$ and bi-directional edges $\{(i, \mathrm{id}_{54}(i+5)), (i, \mathrm{id}_{54}(i+10)), (i, \mathrm{id}_{54}(i+15)), (i, \mathrm{id}_{54}(i+20))\}$ for each $i \in \{1, \ldots, 54\}$, where $\mathrm{id}_{54}(x) = x$ if $x \in \{1, \ldots, 54\}$ and $x - 54$ otherwise. All edge weights are 0.1.
$\hat{\mathcal{G}}$	obtained from $\mathcal{G}$ by replacing the directed cycle with an undirected one keeping the edge weights same
$\hat{\mathcal{G}}_{\setminus\{4,11,25,45\}}$	obtained from $\hat{\mathcal{G}}$ by removing the vertices $\{4, 11, 25, 45\}$ and the edges adjacent to them
$\hat{\mathcal{G}}_{\setminus\{4,25,27\}}$	obtained from $\hat{\mathcal{G}}$ by removing the vertices $\{4, 25, 27\}$ and the edges adjacent to them

Table 1
Definition of the digraphs $\mathcal{G}$, $\hat{\mathcal{G}}$, $\hat{\mathcal{G}}_{\setminus\{4,11,25,45\}}$, and $\hat{\mathcal{G}}_{\setminus\{4,25,27\}}$.



(a) Power allocation (b) Total cost (c) Total mismatch

(d) Power allocation (e) Total cost (f) Total mismatch

(g) Power allocation (h) Total cost (i) Total mismatch

Fig. 1. Evolution of the power allocation, the total cost, and the total mismatch between generation and load under the `dac+L`∂ dynamics for the IEEE 118 bus example in different scenarios. In the first case (a)-(c), the communication topology is $\mathcal{G}$, the load is initially 4600 and later 4200, and parameters are $\nu_1 = 1$, $\nu_2 = 1.3$, $\alpha = 10$, $\beta = 40$, and $\epsilon = 0.0086$. In the second scenario (d)-(f), the digraph and the parameters remain the same but the load is time-varying given by $P_l(t) = 4300 + 100\sin(0.05t)$. In the last case (g)-(i), the parameters remain the same, the communication graph is initially the graph $\hat{\mathcal{G}}$. At $t = 100s$, units $\{4, 11, 25, 45\}$ leave the network, resulting in the communication topology $\hat{\mathcal{G}}_{\setminus\{4,11,25,45\}}$, and the remaining agents run the TRAJECTORY INVARIANCE routine. Later, at $t = 200s$, units $\{11, 45\}$ join the network while unit 27 leaves it, resulting in the communication topology $\hat{\mathcal{G}}_{\setminus\{4,25,27\}}$. After implementing the TRAJECTORY INVARIANCE routine, the `dac+L`∂ algorithm eventually converges to an optimizer of the ED problem for the network $\hat{\mathcal{G}}_{\setminus\{4,25,27\}}$.

and ideas from distributed optimization to dynamically allocate the unit generation levels. Our analysis has shown that the mismatch dynamics between total generation and load are input-to-state stable and, as a consequence, the coordination algorithm is robust to initialization errors, time-varying load signals, and intermittent power generation. Our technical approach relies on tools from algebraic graph theory, dynamic average consensus, set-valued dynamical systems, and nonsmooth analysis, including a novel refinement of the LaSalle Invariance Principle for differential inclusions that we have stated and proved. We see the results obtained here as a step upon which one can build towards synthezising solutions for increasingly complex and realistic scenarios. Future work will explore the study of the preservation of the generator box constraints under the proposed coordination strategy, the extension to scenarios that involve additional constraints, such as transmission losses, transmission line capacity constraints, ramp rate limits, prohibited operating zones, and valve-point loading effects, and the study of the stability and convergence properties of algorithm de-

signs that combine our approach here with traditional primary and secondary generator controllers.

Acknowledgments

Research supported by NSF award CMMI-1300272.

References

A. Arsie and C. Ebenbauer. Locating omega-limit sets using height functions. *Journal of Differential Equations*, 248(10): 2458–2469, 2010.

J. P. Aubin and A. Cellina. *Differential Inclusions*, volume 264 of *Grundlehren der mathematischen Wissenschaften*. Springer, New York, 1984.

A. Bacciotti and F. Ceragioli. Stability and stabilization of discontinuous systems and nonsmooth Lyapunov functions. *ESAIM: Control, Optimisation & Calculus of Variations*, 4:361–376, 1999.

D. S. Bernstein. *Matrix Mathematics*. Princeton University Press, Princeton, NJ, 2005.

G. Binetti, A. Davoudi, F. L. Lewis, D. Naso, and B. Turchiano. Distributed consensus-based economic dispatch with transmission losses. *IEEE Transactions on Power Systems*, 29(4):1711–1720, 2014a.

G. Binetti, A. Davoudi, D. Naso, B. Turchiano, and F. L. Lewis. A distributed auction-based algorithm for the nonconvex economic dispatch problem. *IEEE Transaction on Industrial Informatics*, 10(2):1124–1132, 2014b.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2009. ISBN 0521833787.

F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton University Press, 2009. ISBN 978-0-691-14195-4. Electronically available at http://coordinationbook.info.

A. Cherukuri and J. Cortés. Distributed coordination for economic dispatch with varying load and generator commitment. In *Allerton Conf. on Communications, Control and Computing*, pages 475–482, Monticello, IL, October 2014.

A. Cherukuri and J. Cortés. Distributed generator coordination for initialization and anytime optimization in economic dispatch. *IEEE Transactions on Control of Network Systems*, 2 (3), 2015.

B. H. Chowdhury and S. Rahman. A review of recent advances in economic dispatch. *IEEE Transactions on Power Systems*, 5(4):1248–1259, November 1990.

J. Cortés. Discontinuous dynamical systems - a tutorial on solutions, nonsmooth analysis, and stability. *IEEE Control Systems Magazine*, 28(3):36–73, 2008.

A. D. Dominguez-Garcia, S. T. Cady, and C. N. Hadjicostis. Decentralized optimal dispatch of distributed energy resources. In *IEEE Conf. on Decision and Control*, pages 3688–3693, Hawaii, USA, December 2012.

L. Du, S. Grijalva, and R. G. Harley. Potential-game theoretical formulation of optimal power flow problems. In *IEEE Power and Energy Society General Meeting*, San Diego, CA, July 2012. Electronic Proceedings.

H. Farhangi. The path of the smart grid. *IEEE Power and Energy Magazine*, 8(1):18–28, 2010.

R. A. Freeman, P. Yang, and K. M. Lynch. Stability and convergence properties of dynamic average consensus estimators. In *IEEE Conf. on Decision and Control*, pages 398–403, San Diego, CA, 2006.

IEEE 118 bus. http://motor.ece.iit.edu/data/JEAS_IEEE118.doc.

B. Johansson and M. Johansson. Distributed non-smooth resource allocation over a network. In *IEEE Conf. on Decision and Control*, pages 1678–1683, Shanghai, China, December 2009.

S. Kar and G. Hug. Distributed robust economic dispatch in power systems: A consensus + innovations approach. In *IEEE Power and Energy Society General Meeting*, San Diego, CA, July 2012. Electronic proceedings.

H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 3 edition, 2002. ISBN 0130673897.

S. S. Kia, J. Cortés, and S. Martínez. Dynamic average consensus under limited control authority and privacy requirements. *International Journal on Robust and Nonlinear Control*, 25(13): 1941–1966, 2015.

N. Li, C. Zhao, and L. Chen. Connecting automatic generation control and economic dispatch from an optimization view. *IEEE Transactions on Control of Network Systems*, 2015. To appear.

V. Loia and A. Vaccaro. Decentralized economic dispatch in smart grids by self-organizing dynamic agents. *IEEE Transactions on Systems, Man & Cybernetics: Systems*, 2013. To appear.

E. Mallada, C. Zhao, and S. Low. Optimal load-side control for frequency regulation in smart grids. *IEEE Transactions on Automatic Control*, 2015. To appear.

B. Mohar. Eigenvalues, diameter, and mean distance in graphs. *Graphs and Combinatorics*, 7(1):53–64, 1991.

R. Mudumbai, S. Dasgupta, and B. B. Cho. Distributed control for optimal economic dispatch of a network of heterogeneous power generators. *IEEE Transactions on Power Systems*, 27 (4):1750–1760, 2012.

A. Pantoja, N. Quijano, and K. M. Passino. Dispatch of distributed generators under local-information constraints. In *American Control Conference*, pages 2682–2687, Portland, OR, June 2014.

W. Ren and R. W. Beard. *Distributed Consensus in Multi-Vehicle Cooperative Control*. Communications and Control Engineering. Springer, 2008. ISBN 978-1-84800-014-8.

A. Simonetto, T. Keviczky, and M. Johansson. A regularized saddle-point algorithm for networked optimization with resource allocation constraints. In *IEEE Conf. on Decision and Control*, pages 7476–7481, Hawaii, USA, December 2012.

L. Xiao and S. Boyd. Optimal scaling of a gradient method for distributed resource allocation. *Journal of Optimization Theory & Applications*, 129(3):469–488, 2006.

W. Zhang, W. Liu, X. Wang, L. Liu, and F. Ferrese. Online optimal generation control based on constrained distributed gradient algorithm. *IEEE Transactions on Power Systems*, 2014. To appear.

Z. Zhang and M. Chow. Convergence analysis of the incremental cost consensus algorithm under different communication network topologies. *IEEE Transactions on Power Systems*, 27(4): 1761–1768, 2012.

Z. Zhang, X. Ying, and M. Chow. Decentralizing the economic dispatch problem using a two-level incremental cost consensus algorithm in a smart grid environment. In *North American Power Symposium*, Boston, MA, August 2011. Electronic Proceedings.

C. Zhao, U. Topcu, N. Li, and S. Low. Design and stability of load-side primary frequency control in power systems. *IEEE Transactions on Automatic Control*, 59(5):1177–1189, 2014.

A Refined LaSalle Invariance Principle for differential inclusions

In this section we provide a refinement of the LaSalle Invariance Principle for differential inclusions, see e.g., [Cortés, 2008], by extending the results of [Arsie and Ebenbauer, 2010] for differential equations. Our motivation for developing this refinement comes from the need to provide the necessary tools to tackle the convergence analysis of the coordination algorithms presented in Sections 4 and 5. Nevertheless, the results stated here are of independent interest.

Proposition A.1 *(Refined LaSalle Invariance Principle for differential inclusions): Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be upper semicontinuous, taking nonempty, convex, and compact values at every point $x \in \mathbb{R}^n$. Consider the differential inclusion $\dot{x} \in F(x)$ and let $t \mapsto \varphi(t)$ be a bounded solution whose omega-limit set $\Omega(\varphi)$ is contained in $\mathcal{S} \subset \mathbb{R}^n$, a closed embedded submanifold of $\mathbb{R}^n$. Let $\mathcal{O}$ be an open neighborhood of $\mathcal{S}$ where a locally Lipschitz, regular function $W : \mathcal{O} \to \mathbb{R}$ is defined. Assume the following holds,*

(i) the set $\mathcal{E} = \{x \in \mathcal{S} \mid 0 \in \mathcal{L}_F W(x)\}$ belongs to a level set of W,

(ii) for any compact set $\mathcal{M} \subset \mathcal{S}$ with $\mathcal{M} \cap \mathcal{E} = \emptyset$, there exists a compact neighborhood $\mathcal{M}_c$ of $\mathcal{M}$ in $\mathbb{R}^n$ and $\delta < 0$ such that $\sup_{x \in \mathcal{M}_c} \max \mathcal{L}_F W(x) \leq \delta$.

Then, $\Omega(\varphi) \subset \mathcal{E}$.

Before proceeding with the proof of the result, we establish an auxiliary result.

Lemma A.2 *Under the hypotheses of Proposition A.1, the sets $\Omega(\varphi)$ and $\mathcal{E}$ have nonempty intersection.*

PROOF. By contradiction, assume $\Omega(\varphi) \cap \mathcal{E} = \emptyset$. Then, using the hypothesis (ii) in Proposition A.1, there exists $\delta < 0$ such that $\sup_{x \in \Omega(\varphi)} \max \mathcal{L}_F W(x) \leq \delta$. Let $x \in \Omega(\varphi)$. Since this set is weakly positively invariant, there exists a trajectory $t \mapsto \tilde{\varphi}(t)$ of the differential inclusion with $\tilde{\varphi}(0) = x$ such that $\tilde{\varphi}(t) \in \Omega(\varphi)$. Since $\frac{d}{dt} W(\tilde{\varphi}(t)) \in \mathcal{L}_F W(\tilde{\varphi}(t))$ for almost all $t \geq 0$, we get $W(\tilde{\varphi}(t)) - W(x) \leq \delta t$. This is in contradiction with the fact that $t \mapsto \tilde{\varphi}(t)$ belongs to the compact set $\Omega(\varphi)$, where W is lower bounded. □

We are now ready to prove Proposition A.1.

PROOF. [Proof of Proposition A.1] We consider two cases, depending on whether the set $\Omega(\varphi)$ (a) is or (b) is not contained in a level set of W. In case (a), given any $x \in \Omega(\varphi)$, there exists a trajectory of F starting at x that remains in $\Omega(\varphi)$ (because of the weak positive invariance of the omega-limit set). If $x \notin \mathcal{E}$, then by the hypotheses (ii), there exists a compact neighborhood $\mathcal{M}_x$ of x in $\mathbb{R}^n$ and $\delta < 0$ such that $\sup_{y \in \mathcal{M}_x} \mathcal{L}_F W(y) \leq \delta$. Since $\Omega(\varphi) \subset \mathcal{S}$, the trajectory of F starting at x remains in the set $\mathcal{M}_x \cap \mathcal{S}$ for a finite time, say t_1. Over the time interval $[0, t_1]$, we have $W(t) - W(0) \leq \delta t$. This, however, is in contradiction with the fact that the trajectory belongs to $\Omega(\varphi)$ which is contained in a level set of W. Therefore, $x \in \mathcal{E}$, and since this point is generic, we conclude $\Omega(\varphi) \subset \mathcal{E}$.

Next, we consider case (b) and reason by contradiction, i.e., assume that $\Omega(\varphi)$ is not contained in $\mathcal{E}$ (see Figure A.1). Given $\epsilon > 0$, let $\mathcal{B}_\epsilon \subset \mathcal{O}$ be a compact neighborhood of $\Omega(\varphi)$ in $\mathbb{R}^n$ such that $d(\mathcal{B}_\epsilon, \Omega(\varphi)) \leq \epsilon$. Let $\mathcal{U}$ be an open neighborhood of $\mathcal{E}$ in $\mathbb{R}^n$ and define $\mathcal{U}_\epsilon = \mathcal{U} \cap \mathcal{B}_\epsilon$. Note that $\mathcal{U}_\epsilon$ is nonempty because $\Omega(\varphi) \cap \mathcal{E}$ is nonempty by Lemma A.2. Since $\Omega(\varphi)$ is not contained in a level set of W but $\mathcal{E}$ is by hypotheses (i), we can

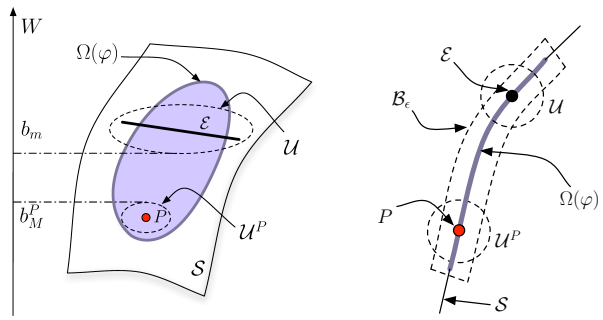


Fig. A.1. Illustration (adapted from [Arsie and Ebenbauer, 2010, Figure 1]) depicting various elements involved in the case (b) of the proof of Proposition A.1.

choose $P \in \Omega(\varphi) \setminus \mathcal{E}$ such that $W(P) \neq W(\mathcal{E})$. Without loss of generality, assume $W(P) < W(\mathcal{E})$ (the reasoning is analogous for the other case). Select an open neighborhood $\mathcal{U}^P$ of P in $\mathbb{R}^n$ and define $\mathcal{U}_\epsilon^P = \mathcal{U}^P \cap \mathcal{B}_\epsilon$. Define the following quantities

$$b_m = \inf_{x \in \mathcal{U}_\epsilon} W(x), \quad b_M^P = \sup_{x \in \mathcal{U}_\epsilon^P} W(x).$$

Note that the neighborhoods $\mathcal{U}$ and $\mathcal{U}^P$ can be chosen such that the set $\Omega(\varphi) \setminus (\mathcal{U} \cup \mathcal{U}^P)$ is nonempty, compact, and its intersection with $\mathcal{E}$ is empty. Along with this, one can select ϵ in such a way that $b_m > b_M^P$ and from assumption (ii) we get

$$\sup_{x \in \mathcal{B}_\epsilon \setminus (\mathcal{U} \cup \mathcal{U}^P)} \max \mathcal{L}_F W(x) \leq \delta < 0, \tag{A.1}$$

(in the case $W(P) > W(\mathcal{E})$, we would reason with the quantities $b_M = \sup_{x \in \mathcal{U}_\epsilon} W(x)$ and $b_m^P = \inf_{x \in \mathcal{U}_\epsilon^P} W(x)$). Since $\Omega(\varphi)$ is the omega-limit set of φ and $\mathcal{B}_\epsilon$ is a compact neighborhood of $\Omega(\varphi)$, there exists $t_1 > 0$ such that $\varphi(t_1) \in \mathcal{U}_\epsilon^P$ and $\varphi(t) \in \mathcal{B}_\epsilon$ for all $t \geq t_1$. Moreover, since $\Omega(\varphi) \cap \mathcal{E}$ is nonempty, there must also exist $t_2 > t_1$ such that $\varphi(t_2) \in \mathcal{U}_\epsilon$. From continuity of the trajectory we deduce that there exist times $t_1^*, t_2^* \in (t_1, t_2)$, $t_1^* < t_2^*$ such that $\varphi(t_1^*)$ and $\varphi(t_2^*)$ lie on the boundary of the compact set $\mathcal{B}_\epsilon \setminus (\mathcal{U}_\epsilon \cup \mathcal{U}_\epsilon^P)$, with $\varphi(t_1^*)$ belonging to the closure of $\mathcal{U}_\epsilon^P$ and $\varphi(t_2^*)$ to the closure of $\mathcal{U}_\epsilon$. However, this is not possible as $W(\varphi(t_2^*)) \geq b_m > b_M^P \geq W(\varphi(t_1^*))$ and, in the interval $[t_1^*, t_2^*]$, the trajectory belongs to $\mathcal{B}_\epsilon \setminus (\mathcal{U}_\epsilon \cup \mathcal{U}_\epsilon^P)$, where the function W can only decrease due to (A.1), which is a contradiction. □

B Continuity properties of set-valued Lie derivatives

Here we present an auxiliary result for the convergence analysis of the algorithms of Sections 4 and 5.

Lemma A.1 *(Continuity property of set-valued Lie derivatives): Let $W : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz and regular function. Let $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous*

function and define the set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by $F(x) = \{g(x,\zeta) \mid \zeta \in \partial W(x)\}$. Assume that

(i) $\mathcal{S}$ is an embedded submanifold of $\mathbb{R}^n$ such that $\zeta^\top g(x,\zeta) \leq 0$ for all $x \in \mathcal{S}$ and all $\zeta \in \partial W(x)$,

(ii) for any $x \in \mathcal{S}$, if $\zeta^\top g(x,\zeta) = 0$ for some $\zeta \in \partial W(x)$, then $x \in \mathcal{E} = \{z \in \mathcal{S} \mid 0 \in \mathcal{L}_F W(z)\}$.

Then, for any compact set $\mathcal{M} \subset \mathcal{S}$ with $\mathcal{M} \cap \mathcal{E} = \emptyset$, there exists a compact neighborhood $\mathcal{M}_c$ of $\mathcal{M}$ in $\mathbb{R}^n$ and $\delta < 0$ such that $\sup_{x \in \mathcal{M}_c} \max \mathcal{L}_F W(x) \leq \delta$.

PROOF. We reason by contradiction, i.e., assume that for all compact neighborhoods $\mathcal{M}_c$ of $\mathcal{M}$ in $\mathbb{R}^n$ and all $\delta < 0$, we have

$$\sup_{x \in \mathcal{M}_c} \max \mathcal{L}_F W(x) > \delta.$$

Note that this implies that $\sup_{x \in \mathcal{M}_c} \max \mathcal{L}_F W(x) \geq 0$. Now, for each $k \in \mathbb{Z}_{\geq 1}$, consider the compact neighborhood $\mathcal{M}_k = \mathcal{M} + \overline{B(0, \frac{1}{k})}$ of $\mathcal{M}$. From the above, we deduce the existence of a sequence $\{x_k\}_{k=1}^{\infty}$ with $x_k \in \mathcal{M}_k$ such that

$$\lim_{k \to \infty} \max \mathcal{L}_F W(x_k) = \ell \geq 0. \tag{A.1}$$

Since the whole sequence belongs to the compact set $\mathcal{M}_1$, there exists a subsequence, which we denote with the same indices for simplicity, such that

$$\lim_{k \to \infty} x_k = \tilde{x} \in \mathcal{M}. \tag{A.2}$$

From (A.1), there exists a sequence $\zeta_k \in \partial W(x_k)$ such that

$$\lim_{k \to \infty} \zeta_k^\top g(x_k, \zeta_k) \geq 0. \tag{A.3}$$

Since ∂W is upper semicontinuous with compact values, the set $\partial W(\mathcal{M}_1)$ is compact, cf. [Aubin and Cellina, 1984, Proposition 3, p. 42]. This implies that the sequence $\{\zeta_k\}$ belongs to the compact set $\partial W(\mathcal{M}_1)$ and hence, there exists a subsequence, denoted again by the same indices for simplicity, such that $\zeta_k \to \tilde{\zeta}$. Now since ∂W is upper semicontinuous and takes closed values, we deduce from [Aubin and Cellina, 1984, Proposition 2, p. 41] that $\tilde{\zeta} \in \partial W(\tilde{x})$. From (A.2) and (A.3), since g is continuous, we obtain $\tilde{\zeta}^\top g(\tilde{x}, \tilde{\zeta}) \geq 0$. By assumption (i), this implies $\tilde{\zeta}^\top g(\tilde{x}, \tilde{\zeta}) = 0$. Assumption (ii) then implies $\tilde{x} \in \mathcal{E}$, which together with (A.2) contradicts $\mathcal{M} \cap \mathcal{E} = \emptyset$. $\square$