Distributed coordination of DERs with storage for dynamic economic dispatch

Ashish Cherukuri Jorge Cortés

Abstract—**This paper considers the dynamic economic dispatch problem for a group of distributed energy resources (DERs) with storage that communicate over a weight-balanced strongly connected digraph. The objective is to collectively meet a certain load profile over a finite time horizon while minimizing the aggregate cost. At each time slot, each DER decides on the amount of generated power, the amount sent to/drawn from the storage unit, and the amount injected into the grid to satisfy the load. Additional constraints include bounds on the amount of generated power, ramp constraints on the difference in generation across successive time slots, and bounds on the amount of power in storage. We synthesize a provably-correct distributed algorithm that solves the resulting finite-horizon optimization problem starting from any initial condition. Our design consists of two interconnected systems, one estimating the mismatch between the injection and the total load at each time slot, and another using this estimate to reduce the mismatch and optimize the total cost of generation while meeting the constraints.**

I. Introduction

The current electricity grid is up for a major transformation to enable the widespread integration of distributed energy resources and flexible loads to improve efficiency and reduce emissions without affecting reliability and performance. This presents the need for novel coordinated control and optimization strategies which, along with suitable architectures, can handle uncertainties and variability, are fault-tolerant and robust, and preserve privacy. With this context in mind, our objective here is to provide a distributed algorithmic solution to the dynamic economic dispatch problem with storage. We see the availability of such strategies as a necessary building block in realizing the vision of the future grid.

Literature review: Static economic dispatch (SED) involves a group of generators collectively meeting a specified load for a single time slot while minimizing the total cost and respecting individual constraints. In recent years, distributed generation has motivated the shift from traditional solutions of the SED problem to decentralized ones, see e.g., [2], [3], [4] and our own work [5], [6]. As argued in [7], [8], the dynamic version of the problem, termed dynamic economic dispatch (DED), results in better grid control as it optimally plans generation across a time horizon, specifically taking into account ramp limits and variability of power commitment from renewable sources. Conventional solution methods to the DED problem are centralized [7]. Recent works [8], [9] have employed model predictive control (MPC)-based algorithms to deal more effectively with complex constraints and uncertainty, but the resulting methods are still centralized and do not provide theoretical guarantees on the optimality of the solution. The work [10] proposes a Lagrangian relaxation method to solve the DED problem, but the implementation requires a master agent that communicates with and coordinates the generators. MPC methods have also been employed by [11] in the dynamic economic dispatch with storage (DEDS) problem, which adds storage units to the DED problem to lower the total cost, meet uncertain demand under uncertain generation, and smooth out the generation profile across time. The stochastic version of the DEDS problem adds uncertainty in demand and generation by renewables. Algorithmic solutions for this problem put the emphasis on breaking down the complexity to speed up convergence for large-scale problems and include stochastic MPC [12], dual decomposition [13], and optimal condition decomposition [14] methods. However, these methods are either centralized or need a coordinating central master.

Statement of contributions: Our starting point is the formulation of the DEDS problem for a group of power DERs communicating over a weight-balanced strongly connected digraph. Since the cost functions are convex and all constraints are linear, the problem is convex in its decision variables, which are the power to be injected and the power to be sent to storage by each DER at each time slot. Using exact penalty functions, we reformulate the DEDS problem as an equivalent optimization that retains equality constraints but removes inequality ones. The structure of the modified problem guides our design of the provably-correct distributed strategy termed "dynamic average consensus ($\mathtt{dac}$) + Laplacian nonsmooth gradient ($\mathsf{L}\partial$) + nonsmooth gradient (∂)" dynamics to solve the DEDS problem starting from any initial condition. This algorithm consists of two interconnected systems. A first block allows DERs to track, using $\mathtt{dac}$, the mismatch between the current total power injected and the load for each time slot of the planning horizon. A second block has two components, one that minimizes the total cost while keeping the total injection constant (employing Laplacian-nonsmooth-gradient dynamics on injection variables and nonsmooth-gradient dynamics on storage variables) and an error-correcting component that uses the mismatch signal estimated by the first block to adjust, exponentially fast, the total injection towards the load for each time slot.

Notation: Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, $\mathbb{Z}_{\geq 1}$ denote the set of real, nonnegative real, positive real, and positive integer numbers, respectively. The 2- and ∞-norm on $\mathbb{R}^n$ are denoted by $\|\cdot\|$ and $\|\cdot\|_\infty$, respectively. We let $B(x,\delta)$ denote the open ball

A preliminary version appeared as [1] at the IEEE Conference on Decision and Control.

Ashish Cherukuri and Jorge Cortés are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, {acheruku,cortes}@ucsd.edu.

centered at $x \in \mathbb{R}^n$ with radius $\delta > 0$. Given $r \in \mathbb{R}$, we denote $\mathcal{H}_r = \{x \in \mathbb{R}^n \mid \mathbf{1}_n^\top x = r\}$. For a symmetric matrix $A \in \mathbb{R}^{n \times n}$, the minimum and maximum eigenvalues of A are $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$. The Kronecker product of $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{p \times q}$ is $A \otimes B \in \mathbb{R}^{np \times mq}$. We use $\mathbf{0}_n = (0, \dots, 0) \in \mathbb{R}^n$, $\mathbf{1}_n = (1, \dots, 1) \in \mathbb{R}^n$, and $\mathrm{I}_n \in \mathbb{R}^{n \times n}$ for the identity matrix. For $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, the vector $(x; y) \in \mathbb{R}^{n+m}$ denotes the concatenation. Given $x, y \in \mathbb{R}^n$, x_i denotes the i-th component of x, and $x \leq y$ denotes $x_i \leq y_i$ for $i \in \{1, \dots, n\}$. For $\mathfrak{h} > 0$, given $y \in \mathbb{R}^{n\mathfrak{h}}$ and $k \in \{1, \dots, \mathfrak{h}\}$, the vector containing the $nk - n + 1$ to nk components of y is $y^{(k)} \in \mathbb{R}^n$, and so, $y = (y^{(1)}; y^{(2)}; \dots; y^{(\mathfrak{h})})$. We let $[u]^+ = \max\{0, u\}$ for $u \in \mathbb{R}$. A set-valued map $f : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ associates to each point in $\mathbb{R}^n$ a set in $\mathbb{R}^m$.

II. Preliminaries

This section introduces concepts from graph theory, nonsmooth analysis, differential inclusions, and optimization.

Graph theory: Following [15], a *weighted directed graph*, is a triplet $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathsf{A})$, where $\mathcal{V}$ is the vertex set, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set, and $\mathsf{A} \in \mathbb{R}_{\geq 0}^{n \times n}$ is the *adjacency matrix* with the property that $a_{ij} > 0$ if $(v_i, v_j) \in \mathcal{E}$ and $a_{ij} = 0$, otherwise. A path is an ordered sequence of vertices such that any consecutive pair of vertices is an edge. A digraph is *strongly connected* if there is a path between any pair of distinct vertices. For a vertex v_i, $N^{\text{out}}(v_i) = \{v_j \in \mathcal{V} \mid (v_i, v_j) \in \mathcal{E}\}$ is the set of its out-neighbors. The *Laplacian* matrix is $\mathsf{L} = \mathsf{D}_{\text{out}} - \mathsf{A}$, where D_{out} is the diagonal matrix defined by $(\mathsf{D}_{\text{out}})_{ii} = \sum_{j=1}^n a_{ij}$, for all $i \in \{1, \dots, n\}$. Note that $\mathsf{L}\mathbf{1}_n = 0$. If $\mathcal{G}$ is strongly connected, then zero is a simple eigenvalue of L. $\mathcal{G}$ is *weight-balanced* iff $\mathbf{1}_n^\top \mathsf{L} = 0$ iff $\mathsf{L} + \mathsf{L}^\top$ is positive semidefinite. If $\mathcal{G}$ is weight-balanced and strongly connected, then zero is a simple eigenvalue of $\mathsf{L} + \mathsf{L}^\top$ and, for $x \in \mathbb{R}^n$,

$$\lambda_2(\mathsf{L} + \mathsf{L}^\top) \big\| x - \frac{1}{n}(\mathbf{1}_n^\top x)\mathbf{1}_n \big\|^2 \leq x^\top (\mathsf{L} + \mathsf{L}^\top) x, \qquad (1)$$

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

Nonsmooth analysis: Here, we introduce some notions on nonsmooth analysis from [16]. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *locally Lipschitz* at $x \in \mathbb{R}^n$ if there exist $L, \epsilon \in \mathbb{R}_{>0}$ such that $\|f(y) - f(y')\| \leq L\|y - y'\|$, 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 directional derivative and the generalized directional derivative of f at x along the direction v coincide, see [16] for these definitions. 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 $\epsilon \in \mathbb{R}_{>0}$, there exists $\delta \in \mathbb{R}_{>0}$ 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 $\epsilon, \delta \in \mathbb{R}_{>0}$ such that $\|z\| \leq \epsilon$ for all $z \in \mathcal{H}(y)$, and all $y \in B(x, \delta)$. 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* $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ of f is

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

where co is the convex hull and $S \subset \mathbb{R}^n$ is any set of measure zero. The set-valued map ∂f is locally bounded, upper semicontinuous, and takes non-empty, compact, and convex values. For a function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$, the partial generalized gradient with respect to x and y are denoted by $\partial_x f$ and $\partial_y f$, respectively.

Differential inclusions: We gather here tools from [16], [6] to analyze the stability properties of differential inclusions,

$$\dot{x} \in F(x), \qquad (2)$$

where $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued map. A solution of (2) on $[0, T] \subset \mathbb{R}$ is an absolutely continuous map $x : [0, T] \to \mathbb{R}^n$ that satisfies (2) for almost all $t \in [0, T]$. If the set-valued map F is locally bounded, upper semicontinuous, and takes non-empty, compact, and convex values, then the existence of solutions is guaranteed. The set of equilibria of (2) is $\mathrm{Eq}(F) = \{x \in \mathbb{R}^n \mid 0 \in F(x)\}$. Given a locally Lipschitz function $W : \mathbb{R}^n \to \mathbb{R}$, the *set-valued Lie derivative* $\mathcal{L}_F W : \mathbb{R}^n \rightrightarrows \mathbb{R}$ of W with respect to (2) at $x \in \mathbb{R}^n$ is

$$\mathcal{L}_F W = \{a \in \mathbb{R} \mid \exists v \in F(x) \text{ s.t. } \zeta^\top v = a, \forall \zeta \in \partial W(x)\}.$$

The *ω-limit set* of a trajectory $t \mapsto \varphi(t)$, $\varphi(0) \in \mathbb{R}^n$ of (2), denoted $\Omega(\varphi)$, is the set of all points $y \in \mathbb{R}^n$ for which there exists a sequence of times $\{t_k\}_{k=1}^\infty$ with $t_k \to \infty$ such that $\lim_{k \to \infty} \varphi(t_k) = y$. If the trajectory is bounded, then the ω-limit set is nonempty, compact, connected. The next result from [6] is a refinement of the LaSalle Invariance Principle for differential inclusions that establishes convergence of (2).

Proposition 2.1: (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 everywhere in $\mathbb{R}^n$. Let $t \mapsto \varphi(t)$ be a bounded solution of (2) whose ω-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. Then, $\Omega(\varphi) \subset \mathcal{E}$ if the following holds,

(i) $\mathcal{E} = \{x \in \mathcal{S} \mid 0 \in \mathcal{L}_F W(x)\}$ 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 $\sup_{x \in \mathcal{M}_c} \max \mathcal{L}_F W(x) \leq \delta$.

Constrained optimization and exact penalty functions: Here, we introduce some notions on constrained convex optimization following [17], [18]. Consider the optimization problem,

$$\text{minimize} \quad f(x), \qquad (3a)$$

$$\text{subject to} \quad g(x) \leq \mathbf{0}_m, \quad h(x) = \mathbf{0}_p, \qquad (3b)$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, are continuously differentiable and convex, and $h : \mathbb{R}^n \to \mathbb{R}^p$ with $p \leq n$ is affine. The *refined Slater condition* is satisfied by (3) if there exists $x \in \mathbb{R}^n$ such that $h(x) = \mathbf{0}_p$, $g(x) \leq \mathbf{0}_m$, and $g_i(x) < 0$ for all nonaffine functions g_i. The refined Slater condition implies that strong duality holds. A point $x \in \mathbb{R}^n$ is a Karush-Kuhn-Tucker (KKT) point of (3) if there exist Lagrange multipliers $\lambda \in \mathbb{R}_{\geq 0}^m$ and $\nu \in \mathbb{R}^p$ such that

$$g(x) \leq \mathbf{0}_m, \quad h(x) = \mathbf{0}_p, \quad \lambda^\top g(x) = 0,$$

$$\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) + \sum_{i=1}^p \nu_i \nabla h_i(x) = 0.$$

If strong duality holds then, a point is a solution of (3) iff it is a KKT point. The optimization (3) satisfies the *strong Slater condition* with parameter $\rho \in \mathbb{R}_{>0}$ and feasible point $x^\rho \in \mathbb{R}^n$ if $g(x^\rho) < -\rho \mathbf{1}_m$ and $h(x^\rho) = \mathbf{0}_p$.

Lemma 2.2: (Bound on Lagrange multiplier [19, Remark 2.3.3]): If (3) satisfies the strong Slater condition with parameter $\rho \in \mathbb{R}_{>0}$ and feasible point $x^\rho \in \mathbb{R}^n$, then any primal-dual optimizer (x, λ, ν) of (3) satisfies

$$\|\lambda\|_\infty \leq \frac{f(x^\rho) - f(x)}{\rho}.$$

We are interested in eliminating the inequality constraints in (3) while keeping the equality constraints intact. To this end, we use [18] to construct a nonsmooth exact penalty function $f^\epsilon : \mathbb{R}^n \to \mathbb{R}$, given as $f^\epsilon(x) = f(x) + \frac{1}{\epsilon}\sum_{i=1}^m [g_i(x)]^+$, with $\epsilon > 0$, and define the minimization problem

$$\begin{aligned} &\text{minimize} \quad f^\epsilon(x), && \text{(4a)}\\ &\text{subject to} \quad h(x) = \mathbf{0}_p. && \text{(4b)} \end{aligned}$$

Note that f^ϵ is convex as f and $t \mapsto \frac{1}{\epsilon}[t]^+$ are convex. Hence, the problem (4) is convex. The following result, see e.g. [18, Proposition 1], identifies conditions under which the solutions of the problems (3) and (4) coincide.

Proposition 2.3: (Equivalence of (3) *and* (4)*):* Assume (3) has nonempty, compact solution set, and satisfies the refined Slater condition. Then, (3) and (4) have the same solutions if $\frac{1}{\epsilon} > \|\lambda\|_\infty$, for some Lagrange multiplier $\lambda \in \mathbb{R}^m_{\geq 0}$ of (3).

III. Problem statement

Consider a network of $n \in \mathbb{Z}_{\geq 1}$ distributed energy resources (DERs) whose communication topology is a strongly connected and weight-balanced digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathsf{A})$. For simplicity, we assume DERs to be generator units. In our discussion, DERs can also be flexible loads (where the cost function corresponds to the negative of the load utility function). An edge (i, j) represents the capability of unit j to transmit information to unit i. Each unit i is equipped with storage capabilities with minimum $C_i^m \in \mathbb{R}_{\geq 0}$ and maximum $C_i^M \in \mathbb{R}_{>0}$ capacities. The network collectively aims to meet a power demand profile during a finite-time horizon $\mathcal{K} = \{1, \dots, \mathfrak{h}\}$ specified by $l \in \mathbb{R}^{\mathfrak{h}}_{>0}$, that is, $l^{(k)}$ is the demand at time slot $k \in \mathcal{K}$. This demand can either correspond to a load requested from an outside entity, denoted $L^{(k)} \geq 0$ for slot k, or each DER i might have to satisfy a load at the bus it is connected to, denoted $\tilde{l}_i^{(k)} \geq 0$ for slot k. Thus, for each $k \in \mathcal{K}$, $l^{(k)} = L^{(k)} + \sum_{i=1}^n \tilde{l}_i^{(k)}$. We assume that the external demand $L = (L^{(1)}, \dots, L^{(\mathfrak{h})}) \in \mathbb{R}^{\mathfrak{h}}_{\geq 0}$ is known to an arbitrarily selected unit $r \in \{1, \dots, n\}$, whereas the demand at bus i, $\tilde{l}_i = (\tilde{l}_i^{(1)}, \dots, \tilde{l}_i^{(\mathfrak{h})}) \in \mathbb{R}^{\mathfrak{h}}_{\geq 0}$, is known to unit i. For convenience, $\tilde{l} = (\tilde{l}^{(1)}, \dots, \tilde{l}^{\mathfrak{h}})$, where $\tilde{l}^{(k)} = (\tilde{l}_1^{(k)}, \dots, \tilde{l}_n^{(k)})$ collects the load known to each unit at slot $k \in \mathcal{K}$. Along with load satisfaction, the group also aims to minimize the total cost of generation and to satisfy the individual physical constraints for each DER. We make these elements precise next.

Each unit i decides at every time slot k in $\mathcal{K}$ the amount of power it generates, the portion $I_i^{(k)} \in \mathbb{R}$ of it that it injects into the grid to meet the load, and the remaining part $S_i^{(k)} \in \mathbb{R}$ that it sends to the storage unit. The power generated by i at k is then $I_i^{(k)} + S_i^{(k)}$. We denote by $I^{(k)} = (I_1^{(k)}, \dots, I_n^{(k)}) \in \mathbb{R}^n$ and $S^{(k)} = (S_1^{(k)}, \dots, S_n^{(k)}) \in \mathbb{R}^n$ the collective injected and stored power at time k, respectively. The load satisfaction is then expressed as $\mathbf{1}_n^\top I^{(k)} = l^{(k)} = L^{(k)} + \mathbf{1}_n^\top \tilde{l}^{(k)}$, for all $k \in \mathcal{K}$. The cost $f_i^{(k)}(I_i^{(k)} + S_i^{(k)})$ of power generation $I_i^{(k)} + S_i^{(k)}$ by unit i at time k is specified by the function $f_i^{(k)} : \mathbb{R} \to \mathbb{R}_{\geq 0}$, which we assume convex and continuously differentiable. Given $(I^{(k)}, S^{(k)})$, the cost incurred by the network at time slot k is

$$f^{(k)}(I^{(k)} + S^{(k)}) = \sum_{i=1}^n f_i^{(k)}(I_i^{(k)} + S_i^{(k)}).$$

The cumulative cost of generation for the network across the time horizon is $f : \mathbb{R}^{n\mathfrak{h}} \to \mathbb{R}_{\geq 0}$, $f(x) = \sum_{k=1}^{\mathfrak{h}} f^{(k)}(x^{(k)})$. Given injection $I = (I^{(1)}, \dots, I^{(\mathfrak{h})}) \in \mathbb{R}^{n\mathfrak{h}}$ and storage $S = (S^{(1)}, \dots, S^{(\mathfrak{h})}) \in \mathbb{R}^{n\mathfrak{h}}$ values, the total network cost is

$$f(I + S) = \sum_{k=1}^{\mathfrak{h}} f^{(k)}(I^{(k)} + S^{(k)}).$$

The functions $\{f^{(k)}\}_{k \in \mathcal{K}}$ and f are also convex and continuously differentiable. Next, we describe the physical constraints on the DERs. Each unit's power must belong to the range $[P_i^m, P_i^M] \subset \mathbb{R}_{>0}$, representing lower and upper bounds on the amount of power it can generate at each time slot. Each unit i also respects upper and lower ramp constraints: the change in the generation level from any time slot k to $k+1$ is upper and lower bounded by R_i^u and $-R_i^l$, respectively, with $R_i^u, R_i^l \in \mathbb{R}_{>0}$. At each time slot, the power injected into the grid by each unit must be nonnegative, i.e., $I_i^{(k)} \geq 0$. Furthermore, the amount of power stored in any storage unit i at any time slot $k \in \mathcal{K}$ must belong to the range $[C_i^m, C_i^M]$. Finally, we assume that at the beginning of the time slot $k = 1$, each storage unit i starts with some stored power $S_i^{(0)} \in [C_i^m, C_i^M]$. With the above model, the *dynamic economic dispatch with storage* (DEDS) problem is formally defined by the following convex optimization problem,

$$\begin{aligned} &\underset{(I,S) \in \mathbb{R}^{2n\mathfrak{h}}}{\text{minimize}} \quad f(I + S), && \text{(5a)}\\ &\text{subject to for } k \in \mathcal{K}, \\ &\qquad \mathbf{1}_n^\top I^{(k)} = l^{(k)}, && \text{(5b)}\\ &\qquad P^m \leq I^{(k)} + S^{(k)} \leq P^M, && \text{(5c)}\\ &\qquad C^m \leq S^{(0)} + \textstyle\sum_{k'=1}^k S^{(k')} \leq C^M, && \text{(5d)}\\ &\qquad \mathbf{0}_n \leq I^{(k)}, && \text{(5e)}\\ &\qquad \text{for } k \in \mathcal{K} \setminus \{\mathfrak{h}\}, \\ &\qquad -R^l \leq I^{(k+1)} + S^{(k+1)} - I^{(k)} - S^{(k)} \leq R^u. && \text{(5f)} \end{aligned}$$

We refer to (5b)–(5f) as the *load conditions*, *box constraints*, *storage limits*, *injection constraints*, and *ramp constraints*, respectively. We denote by $\mathcal{F}_{\text{DEDS}}$ and $\mathcal{F}^*_{\text{DEDS}}$ the feasibility set and the solution set of the DEDS problem (5), respectively, and assume them to be nonempty. Since $\mathcal{F}_{\text{DEDS}}$ is compact, so is $\mathcal{F}^*_{\text{DEDS}}$. Moreover, the refined Slater condition is satisfied for DEDS as all the constraints (5b)–(5f) are affine in the decision variables. Additionally, we assume that the DEDS problem satisfies the strong Slater condition with parameter $\rho \in \mathbb{R}_{>0}$ and feasible point $(I^\rho, S^\rho) \in \mathbb{R}^{2n\mathfrak{h}}$.

Remark 3.1: (General setup for storage): The DEDS formulation above can be modified to consider scenarios where only some DERs $\mathcal{V}_{gs}$ are equipped with storage and others $\mathcal{V}_g$ are not, with $\{1, \dots, n\} = \mathcal{V}_{gs} \cup \mathcal{V}_g$. The formulation can also be extended to consider the cost of storage, inefficiencies, and constraints on (dis)charging of the storage units, as in [11], [13]. These factors either affect the constraint (5d), add additional conditions on the storage variables, or modify the objective function. As long as the resulting cost and constraints are convex in S, all these can be treated within (5) without affecting the design methodology. •

Our aim is to design a distributed algorithm that allows the network interacting over $\mathcal{G}$ to solve the DEDS problem.

IV. Distributed algorithmic solution

We describe here the distributed algorithm that asymptotically finds the optimizers of the DEDS problem. Our design strategy builds on an alternative formulation of the optimization problem using penalty functions (cf. Section IV-A). This allows us to get rid of the inequality constraints, resulting into an optimization whose structure guides our algorithmic design (cf. Section IV-B).

A. Alternative formulation of the DEDS *problem:* The procedure here follows closely the theory of exact penalty functions outlined in Section II. For an $\epsilon \in \mathbb{R}_{>0}$, consider the modified cost function $f^\epsilon : \mathbb{R}^{n\mathfrak{h}} \times \mathbb{R}^{n\mathfrak{h}} \to \mathbb{R}_{\geq 0}$,

$$f^\epsilon(I,S) = f(I+S) + \frac{1}{\epsilon}\Big(\sum_{k=1}^{\mathfrak{h}} \mathbf{1}_n^\top \big([T_1^{(k)}]^+ + [T_2^{(k)}]^+ + [T_3^{(k)}]^+ + [T_4^{(k)}]^+ + [T_5^{(k)}]^+\big) + \sum_{k=1}^{\mathfrak{h}-1} \mathbf{1}_n^\top \big([T_6^{(k)}]^+ + [T_7^{(k)}]^+\big)\Big),$$

where

$$\begin{aligned}
&T_1^{(k)} = P^m - I^{(k)} - S^{(k)},\ T_2^{(k)} = I^{(k)} + S^{(k)} - P^M, \\
&T_3^{(k)} = C^m - S^{(0)} - \textstyle\sum_{k'=1}^{k} S^{(k')}, \\
&T_4^{(k)} = S^{(0)} + \textstyle\sum_{k'=1}^{k} S^{(k')} - C^M,\ T_5^{(k)} = -I^{(k)}, \\
&T_6^{(k)} = -R^l - I^{(k+1)} - S^{(k+1)} + I^{(k)} + S^{(k)}, \\
&T_7^{(k)} = I^{(k+1)} + S^{(k+1)} - I^{(k)} - S^{(k)} - R^u.
\end{aligned} \tag{6}$$

This cost contains the penalty terms for all the inequality constraints of the DEDS problem. Note that f^ϵ is locally Lipschitz, jointly convex in I and S, and regular. Thus, the partial generalized gradients $\partial_I f^\epsilon$ and $\partial_S f^\epsilon$ take nonempty, convex, compact values and are locally bounded and upper semicontinuous. Consider the modified DEDS problem

$$\begin{aligned}
&\text{minimize} \quad f^\epsilon(I,S), && (7a)\\
&\text{subject to} \quad \mathbf{1}_n^\top I^{(k)} = l^{(k)},\ \forall k \in \mathcal{K}. && (7b)
\end{aligned}$$

The next result provides a criteria for selecting ϵ such that the modified DEDS and the DEDS problems have the exact same solutions. The proof is a direct application of Lemmas 2.2 and 2.3 using that the DEDS problem satisfies the strong Slater condition with parameter ρ and feasible point (I^ρ, S^ρ).

Lemma 4.1: (Equivalence of DEDS *and modified* DEDS *problems):* Let $(I^*, S^*) \in \mathcal{F}^*_{\text{DEDS}}$. Then, the optimizers of the problems (5) and (7) are the same for $\epsilon \in \mathbb{R}_{>0}$ satisfying

$$\epsilon < \frac{\rho}{f(I^\rho + S^\rho) - f(I^* + S^*)}. \tag{8}$$

As a consequence, if ϵ satisfies (8) then, writing the Lagrangian and the KKT conditions for (7) gives the following characterization of the solution set of the DEDS problem

$$\begin{aligned}
\mathcal{F}^*_{\text{DEDS}} = \{&(I,S) \in \mathbb{R}^{2n\mathfrak{h}} \mid \mathbf{1}_n^\top I^{(k)} = l^{(k)} \text{ for all } k \in \mathcal{K}, \\
&0 \in \partial_S f^\epsilon(I,S), \text{ and } \exists \nu \in \mathbb{R}^{\mathfrak{h}} \text{ such that} \\
&(\nu^{(1)}\mathbf{1}_n; \dots; \nu^{(\mathfrak{h})}\mathbf{1}_n) \in \partial_I f^\epsilon(I,S)\}.
\end{aligned} \tag{9}$$

Recall that $\mathcal{F}^*_{\text{DEDS}}$ is bounded. Next, we stipulate a mild regularity assumption on this set which implies that perturbing it by a small parameter does not result into an unbounded set. This property is of use in our convergence analysis later.

Assumption 4.2: (Regularity of $\mathcal{F}^*_{\text{DEDS}}$): For $p \in \mathbb{R}_{\geq 0}$, define the map $p \mapsto \mathcal{F}(p) \subset \mathbb{R}^{2n\mathfrak{h}}$ as

$$\begin{aligned}
\mathcal{F}(p) = \{&(I,S) \in \mathbb{R}^{2n\mathfrak{h}} \mid \left|\mathbf{1}_n^\top I^{(k)} - l^{(k)}\right| \leq p \text{ for all } k \in \mathcal{K}, \\
&0 \in \partial_S f^\epsilon(I,S) + pB(0,1), \text{ and } \exists \nu \in \mathbb{R}^{\mathfrak{h}} \text{ such that} \\
&(\nu^{(1)}\mathbf{1}_n; \dots; \nu^{(\mathfrak{h})}\mathbf{1}_n) \in \partial_I f^\epsilon(I,S) + pB(0,1)\}.
\end{aligned}$$

Note that $\mathcal{F}(0) = \mathcal{F}^*_{\text{DEDS}}$. Then, there exists a $\bar{p} > 0$ such that $\mathcal{F}(p)$ is bounded for all $p \in [0, \bar{p})$. •

We end this section by stating a property of the generalized gradient of f^ϵ that will be employed later in the analysis.

Lemma 4.3: (Uniform bound on the difference between $\partial_I f^\epsilon$ and $\partial_S f^\epsilon$): For $(I,S) \in \mathbb{R}^{2n\mathfrak{h}}$, any two elements $\zeta_1 \in \partial_I f^\epsilon(I,S)$ and $\zeta_2 \in \partial_S f^\epsilon(I,S)$ satisfy

$$\|\zeta_1 - \zeta_2\|_\infty \leq \frac{\mathfrak{h}+4}{\epsilon}.$$

Proof: Write $f^\epsilon(I,S) = f_a(I+S) + f_b(I) + f_c(S)$ where the functions $f_a, f_b, f_c : \mathbb{R}^{n\mathfrak{h}} \to \mathbb{R}_{\geq 0}$ are

$$\begin{aligned}
f_a(I+S) &= f(I+S) + \frac{1}{\epsilon}\Big(\sum_{k=1}^{\mathfrak{h}} \mathbf{1}_n^\top([T_1^{(k)}]^+ + [T_2^{(k)}]^+) + \sum_{k=1}^{\mathfrak{h}-1} \mathbf{1}_n^\top([T_6^{(k)}]^+ + [T_7^{(k)}]^+)\Big), \\
f_b(I) &= \frac{1}{\epsilon}\sum_{k=1}^{\mathfrak{h}} \mathbf{1}_n^\top [T_5^{(k)}]^+, \\
f_c(S) &= \frac{1}{\epsilon}\sum_{k=1}^{\mathfrak{h}} \mathbf{1}_n^\top([T_3^{(k)}]^+ + [T_4^{(k)}]^+).
\end{aligned}$$

From the sum rule of generalized gradients [16], any element $\zeta_1 \in \partial_I f^\epsilon(I,S)$ can be expressed as a sum of the vectors $\zeta_{1,a}$ and $\zeta_{1,b} \in \mathbb{R}^{n\mathfrak{h}}$ such that $\zeta_{1,a} \in \partial f_a(I+S)$ and $\zeta_{1,b} \in \partial f_b(I)$. Similarly, $\zeta_2 = \zeta_{2,a} + \zeta_{2,c}$ where $\zeta_{2,a} \in \partial f_a(I+S)$ and $\zeta_{2,c} \in \partial f_c(S)$. By the definition of f_b, we get $\|\zeta_{1,b}\|_\infty \leq \frac{1}{\epsilon}$. For the function f_c, note that for any $i \in \{1, \dots, n\}$ and any $k \in \mathcal{K}$, either $([T_3^{(k)}]^+)_i$ is zero or $([T_4^{(k)}]^+)_i$ is zero. Considering extreme case, if for a particular i, either $([T_3^{(k)}]^+)_i > 0$ or $([T_4^{(k)}]^+)_i > 0$ for all $k \in \mathcal{K}$ then, we obtain $\left|(\zeta_{2,c})_i^{(1)}\right| = \frac{\mathfrak{h}}{\epsilon}$. This implies that $\|\zeta_{2,c}\|_\infty \leq \frac{\mathfrak{h}}{\epsilon}$. Now consider any two elements $\zeta_{1,a}, \zeta_{2,a} \in \partial f_a(I+S)$. Note that for any $i \in \{1, \ldots, n\}$, either $([T_1^{(k)}]^+)_i$ is zero or $([T_2^{(k)}]^+)_i$ is zero. Similarly, either $([T_6^{(k)}]^+)_i$ or $([T_7^{(k)}]^+)_i$ is zero. Further, note that $I_i^{(k)} + S_i^{(k)}$ appears in $([T_6^{(k)}]^+)_i$ and $([T_7^{(k)}]^+)_i$ as well as in $([T_6^{(k-1)}]^+)_i$ and $([T_7^{(k-1)}]^+)_i$. At the same time, only two of these four terms are nonzero for any $k \in \mathcal{K} \setminus \mathfrak{h}$ and any $i \in \{1, \ldots, n\}$. Using these facts one can obtain the bound $\|\zeta_{1,a} - \zeta_{2,a}\|_\infty \leq \frac{3}{\epsilon}$. Finally, the proof concludes noting

$$\begin{aligned}\|\zeta_1 - \zeta_2\|_\infty &= \|\zeta_{1,a} + \zeta_{1,b} - \zeta_{2,a} - \zeta_{2,c}\|_\infty \\ &\leq \|\zeta_{1,a} - \zeta_{2,a}\|_\infty + \|\zeta_{1,b}\|_\infty + \|\zeta_{2,c}\|_\infty = \frac{\mathfrak{h}+4}{\epsilon}. \quad \blacksquare\end{aligned}$$

B. The $\texttt{dac+}(\mathsf{L}\partial, \partial)$ *coordination algorithm:* Here, we present our distributed algorithm and establish its asymptotic convergence to the set of solutions of the DEDS problem starting from any initial condition. Our design combines ideas of Laplacian-gradient dynamics [5] and dynamic average consensus [20]. Consider the set-valued dynamics,

$$\begin{aligned}\dot{I} &\in -(\mathrm{I}_{\mathfrak{h}} \otimes \mathsf{L})\partial_I f^\epsilon(I,S) + \nu_1 z, && \text{(10a)}\\ \dot{S} &\in -\partial_S f^\epsilon(I,S), && \text{(10b)}\\ \dot{z} &= -\alpha z - \beta(\mathrm{I}_{\mathfrak{h}} \otimes \mathsf{L})z - v + \nu_2(L \otimes e_r + \tilde{l} - I), && \text{(10c)}\\ \dot{v} &= \alpha\beta(\mathrm{I}_{\mathfrak{h}} \otimes \mathsf{L})z, && \text{(10d)}\end{aligned}$$

where $\alpha, \beta, \nu_2, \nu_2 \in \mathbb{R}_{>0}$ are design parameters and $e_r \in \mathbb{R}^n$ is the unit vector along the r-th coordinate. This dynamics is an interconnected system with two parts: the (I,S)-component seeks to adjust the injection levels to satisfy the load profile and search for the optimizers of the DEDS problem while the (z,v)-component corresponds to the dynamic average consensus part, with $z_i^{(k)}$ aiming to track the difference between the load $l^{(k)} = L^{(k)} + \mathbf{1}_n^\top \tilde{l}_i^{(k)}$ and the current injection level $\mathbf{1}_n^\top I^{(k)}$ for unit i. Our terminology $\texttt{dac+}(\mathsf{L}\partial, \partial)$ dynamics to refer to (10) is motivated by this "dynamic average consensus in (z,v)+ Laplacian gradient in I + gradient in S" structure. For convenience, we denote (10) by $X_{\texttt{dac+}(\mathsf{L}\partial,\partial)} : \mathbb{R}^{4n\mathfrak{h}} \rightrightarrows \mathbb{R}^{4n\mathfrak{h}}$. Note $\mathrm{Eq}(X_{\texttt{dac+}(\mathsf{L}\partial,\partial)}) = \mathcal{F}^*_{\mathrm{DEDS}}$ and since $\partial_I f^\epsilon$ and $\partial_S f^\epsilon$ are locally bounded, upper semicontinuous and take nonempty convex compact values, the solutions of $X_{\texttt{dac+}(\mathsf{L}\partial,\partial)}$ exist starting from any initial condition (cf. Section II).

Remark 4.4: (Distributed implementation of the $\texttt{dac+}(\mathsf{L}\partial, \partial)$ *dynamics):* Writing the (z,v) dynamics componentwise, one can see that for each i and each k, the values $(\dot{z}_i^{(k)}, \dot{v}_i^{(k)})$ can be computed using the state variables $(z_i^{(k)}, \{z_j^{(k)}\}_{j \in N^{\mathrm{out}}(i)}, v_i^{(k)}, I_i^{(k)})$ only. Hence, (10c) and (10d) can be implemented in a distributed manner where each unit only requires information from its out-neighbors. Subsequently, f^ϵ can be written in the separable form

$$f^\epsilon(I,S) = \sum_{i=1}^n f_i^\epsilon(I_i^{(1)}, \ldots, I_i^{(\mathfrak{h})}, S_i^{(1)}, \ldots, S_i^{(\mathfrak{h})}).$$

Thus, if $\zeta_1 \in \partial_I f^\epsilon(I,S)$ and $\zeta_2 \in \partial_S f^\epsilon(I,S)$ then, for all $k \in \mathcal{K}$, $(\zeta_1)_i^{(k)}, (\zeta_2)_i^{(k)} \in \mathbb{R}$ only depend on the state of unit i, i.e., $(I_i^{(1)}, \ldots, I_i^{(\mathfrak{h})}, S_i^{(1)}, \ldots, S_i^{(\mathfrak{h})})$ and are computable by i. Hence, the S-dynamics can implemented by the DERs using their own state and to execute the I-dynamics, each i needs information from its out-neighbors. $\bullet$

We next address the convergence analysis of (10). For convenience, let $\mathfrak{M}_g = \mathbb{R}^{n\mathfrak{h}} \times \mathbb{R}^{n\mathfrak{h}} \times \mathbb{R}^{n\mathfrak{h}} \times (\mathcal{H}_0)^{\mathfrak{h}}$ and $\mathfrak{M}_o = \prod_{k=1}^{\mathfrak{h}} \mathcal{H}_{l^{(k)}} \times \mathbb{R}^{n\mathfrak{h}} \times (\mathcal{H}_0)^{\mathfrak{h}} \times (\mathcal{H}_0)^{\mathfrak{h}}$.

Theorem 4.5: (Convergence of the $\texttt{dac+}(\mathsf{L}\partial, \partial)$ *dynamics to the solutions of the* DEDS *problem):* Let $\mathcal{F}^*_{\mathrm{DEDS}}$ satisfy Assumption 4.2, ϵ satisfy (8), and $\alpha, \beta, \nu_1, \nu_2 > 0$ satisfy

$$\frac{\nu_1}{\beta\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)} + \frac{\nu_2^2\lambda_{\max}(\mathsf{L}^\top\mathsf{L})}{2\alpha} < \lambda_2(\mathsf{L}+\mathsf{L}^\top). \qquad (11)$$

Then, any trajectory of (10) starting in $\mathfrak{M}_g$ converges to $\mathcal{F}^*_{\mathrm{aug}} = \{(I,S,z,v) \in \mathcal{F}^*_{\mathrm{DEDS}} \times \{0\} \times \mathbb{R}^{n\mathfrak{h}} \mid v = \nu_2(l \otimes e_r - I)\}$.

Proof: Our first step is to show that the ω-limit set of any trajectory of (10) with initial condition $(I_0, S_0, z_0, v_0) \in \mathfrak{M}_g$ is contained in $\mathfrak{M}_o$. To this end, write (10d) as

$$\dot{v}^{(k)} = \alpha\beta\mathsf{L}z^{(k)} \quad \text{for all } k \in \mathcal{K}.$$

Note that $\mathbf{1}_n^\top \dot{v}^{(k)} = \alpha\beta\mathbf{1}_n^\top\mathsf{L}z^{(k)} = 0$ for all $k \in \mathcal{K}$ because $\mathcal{G}$ is weight-balanced. Therefore, the initial condition $v_0 \in (\mathcal{H}_0)^{\mathfrak{h}}$ implies that $v(t) \in (\mathcal{H}_0)^{\mathfrak{h}}$ for all $t \geq 0$ along any trajectory of (10) starting at (I_0, S_0, z_0, v_0). Now, if $\zeta \in \partial_I f^\epsilon(I,S)$ then, from (10a) and (10c), we get for any $k \in \mathcal{K}$

$$\begin{aligned}\dot{I}^{(k)} &= -\mathsf{L}\zeta^{(k)} + \nu_1 z^{(k)},\\ \dot{z}^{(k)} &= -\alpha z^{(k)} - \beta\mathsf{L}z^{(k)} - v^{(k)} + \nu_2(l^{(k)}e_r - I^{(k)}).\end{aligned}$$

Let $\xi_k = \mathbf{1}_n^\top I^{(k)} - l^{(k)}$. Then, from the above equations we get $\dot{\xi}_k = \mathbf{1}_n^\top \dot{I}^{(k)} = \nu_1\mathbf{1}_n^\top z^{(k)}$. Further, we have

$$\begin{aligned}\ddot{\xi}_k &= \nu_1\mathbf{1}_n^\top\dot{z}^{(k)} = -\alpha\nu_1\mathbf{1}_n^\top z^{(k)} + \nu_1\nu_2(l^{(k)} - \mathbf{1}^\top I^{(k)})\\ &= -\alpha\dot{\xi}_k - \nu_1\nu_2\xi_k,\end{aligned}$$

forming a second-order linear system for ξ_k. The LaSalle Invariance Principle [21] with the function $\nu_1\nu_2\|\xi_k\|^2 + \|\dot{\xi}_k\|^2$ implies that as $t \to \infty$ we have $(\xi_k(t); \dot{\xi}_k(t)) \to 0$ and so $\mathbf{1}_n^\top I^{(k)}(t) \to l^{(k)}$ and $\mathbf{1}_n^\top z^{(k)}(t) \to 0$ as $t \to \infty$.

Next, proceeding to the convergence analysis, consider the change of coordinates$D : \mathbb{R}^{4n\mathfrak{h}} \to \mathbb{R}^{4n\mathfrak{h}}$ defined by

$$\begin{aligned}(I,S,\omega_1,\omega_2) &= D(I,S,z,v)\\ &= (I,S,z,v + \alpha z - \nu_2(l \otimes e_r - I)).\end{aligned}$$

In these coordinates, the set-valued map (10) takes the form

$$\begin{aligned}X_{\texttt{dac+}(\mathsf{L}\partial,\partial)}(I,S,\omega_1,\omega_2) = \{(&-(\mathrm{I}_{\mathfrak{h}} \otimes \mathsf{L})\zeta_1 + \nu_1\omega_1, -\zeta_2,\\ &-\beta(\mathrm{I}_{\mathfrak{h}} \otimes \mathsf{L})\omega_1 - \omega_2,\\ &\nu_1\nu_2\omega_1 - \alpha\omega_2 - \nu_2(\mathrm{I}_{\mathfrak{h}} \otimes \mathsf{L})\zeta_1) \in \mathbb{R}^{4n\mathfrak{h}} \mid\\ &\zeta_1 \in \partial_I f^\epsilon(I,S), \zeta_2 \in \partial_S f^\epsilon(I,S)\}.\end{aligned} \qquad (12)$$

This transformation helps in identifying the LaSalle-type function for the dynamics. We now focus on proving that, in the new coordinates, the trajectories of (10) converge to

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

Note that $D(\mathfrak{M}_o) = \mathfrak{M}_o$ and so, from the property of the ω-limit set of trajectories above, we get that $t \mapsto (I(t), S(t), \omega_1(t), \omega_2(t))$ starting in $D(\mathfrak{M}_g)$ belongs to $\mathfrak{M}_o$. Next, we show the hypotheses of Proposition 2.1 are satisfied, where $\mathfrak{M}_o$ plays the role of $\mathcal{S} \subset \mathbb{R}^{4n\mathfrak{h}}$ and $V : \mathbb{R}^{4n\mathfrak{h}} \to \mathbb{R}_{\geq 0}$,

$$V(I, S, \omega_1, \omega_2) = f^{\epsilon}(I, S) + \tfrac{1}{2}(\nu_1\nu_2\|\omega_1\|^2 + \|\omega_2\|^2).$$

plays the role of W, resp. Let $(I, S, \omega_1, \omega_2) \in \mathfrak{M}_o$ then any element of $\mathcal{L}_{X_{\mathrm{dac+(L\partial,\partial)}}} V(I, S, \omega_1, \omega_2)$ can be written as

$$\begin{aligned}&- \zeta_1^\top(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L})\zeta_1 + \nu_1\zeta_1^\top\omega_1 - \|\zeta_2\|^2 - \beta\nu_1\nu_2\omega_1^\top(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L})\omega_1 \\ &- \alpha\|\omega_2\|^2 - \nu_2\omega_2^\top(\mathrm{I}_\mathfrak{h} \otimes L)\zeta_1, \end{aligned} \tag{13}$$

where $\zeta_1 \in \partial_I f^\epsilon(I, S)$ and $\zeta_2 \in \partial_S f^\epsilon(I, S)$. Since the digraph $\mathcal{G}$ is strongly connected and weight-balanced, we use (1) and $\mathbf{1}_{n\mathfrak{h}}^\top\omega_1 = 0$ to bound the above expression as

$$\begin{aligned}&- \tfrac{1}{2}\lambda_2(\mathsf{L} + \mathsf{L}^\top)\|\eta\|^2 + \nu_1\eta^\top\omega_1 - \|\zeta_2\|^2 \\ &- \tfrac{1}{2}\beta\nu_1\nu_2\lambda_2(\mathsf{L} + \mathsf{L}^\top)\|\omega_1\|^2 - \alpha\|\omega_2\|^2 - \nu_2\omega_2^\top(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L})\eta \\ &= \gamma^\top M\gamma - \|\zeta_2\|^2,\end{aligned}$$

where $\eta = (\eta^{(1)}; \ldots; \eta^{(\mathfrak{h})})$ with $\eta^{(k)} = \zeta^{(k)} - \frac{1}{n}(\mathbf{1}_n^\top\zeta^{(k)})\mathbf{1}_n$, the vector $\gamma = (\eta; \omega_1; \omega_2)$, and the matrix

$$M = \begin{bmatrix} -\frac{1}{2}\lambda_2(\mathsf{L} + \mathsf{L}^\top)\mathrm{I}_{n\mathfrak{h}} & B^\top \\ B & C \end{bmatrix},$$

with $B^\top = \begin{bmatrix} \frac{1}{2}\nu_1\mathrm{I}_{n\mathfrak{h}} & -\frac{1}{2}\nu_2(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L})^\top \end{bmatrix}$, and

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

Resorting to the Schur complement [17], $M \in \mathbb{R}^{3n\mathfrak{h}\times 3n\mathfrak{h}}$ is neg. definite if $-\frac{1}{2}\lambda_2(\mathsf{L} + \mathsf{L}^\top)\mathrm{I}_{n\mathfrak{h}} - B^\top C^{-1}B$, that equals

$$-\tfrac{1}{2}\lambda_2(\mathsf{L} + \mathsf{L}^\top)\mathrm{I}_{n\mathfrak{h}} + \tfrac{\nu_1}{2\beta\nu_2\lambda_2(\mathsf{L}+\mathsf{L}^\top)}\mathrm{I}_{n\mathfrak{h}} + \tfrac{\nu_2^2}{4\alpha}(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L})^\top(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L}),$$

is negative definite, which follows from (11). Hence, for any $(I, S, \omega_1, \omega_2) \in \mathfrak{M}_o$, we have $\max \mathcal{L}_{X_{\mathrm{dac+(L\partial,\partial)}}} V(I, S, \omega_1, \omega_2) \leq 0$ and also $0 \in \mathcal{L}_{X_{\mathrm{dac+(L\partial,\partial)}}} V(I, S, \omega_1, \omega_2)$ iff $\eta = \zeta_2 = \omega_1 = \omega_2 = 0$, which means $\zeta^{(k)} \in \mathrm{span}\{\mathbf{1}_n\}$ for each $k \in \mathcal{K}$. Consequently, using the characterization of optimizers in (9), we deduce that (I, S) is a solution of (7) and so, $(I, S, \omega_1, \omega_2) \in \overline{\mathcal{F}}_{\mathrm{aug}}$. Since, $\overline{\mathcal{F}}_{\mathrm{aug}}$ belongs to a level set of V, we conclude that Proposition 2.1(i) holds. Further, using [6, Lemma A.1] one can show that Proposition 2.1(ii) holds too (we omit the details due to space constraints).

To apply Proposition 2.1, it remains to show that the trajectories starting from $D(\mathfrak{M}_g)$ are bounded. We reason by contradiction. Assume there exists $t \mapsto (I(t), S(t), \omega_1(t), \omega_2(t))$, with $(I(0), S(0), \omega_1(0), \omega_2(0)) \in D(\mathfrak{M}_g)$, of $X_{\mathrm{dac+(L\partial,\partial)}}$ such that $\|(I(t), S(t), \omega_1(t), \omega_2(t)\| \to \infty$. Since V is radially unbounded, this implies $V(I(t), S(t), \omega_1(t), \omega_2(t)) \to \infty$. Also, as established above, we know $\mathbf{1}_n^\top I^{(k)}(t) \to l^{(k)}$ and $\mathbf{1}_n^\top\omega_1^{(k)}(t) \to 0$ for each $k \in \mathcal{K}$. Thus, there exist times $\{t_m\}_{m=1}^\infty$ with $t_m \to \infty$ such that for all $m \in \mathbb{Z}_{\geq 1}$,

$$\left|\mathbf{1}_n^\top\omega_1^{(k)}(t_m)\right| < 1/m \text{ for all } k \in \mathcal{K}, \tag{14}$$

$$\max \mathcal{L}_{X_{\mathrm{dac+(L\partial,\partial)}}} V(I(t_m), S(t_m), \omega_1(t_m), \omega_2(t_m)) > 0.$$

The second inequality implies the existence of $\{\zeta_{1,m}\}_{m=1}^\infty$ and $\{\zeta_{2,m}\}_{m=1}^\infty$ with $(\zeta_{1,m}, \zeta_{2,m}) \in (\partial_I f^\epsilon(I(t_m), S(t_m)), \partial_S f^\epsilon(I(t_m), S(t_m)))$, such that

$$\begin{aligned}&-\zeta_{1,m}^\top(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L})\zeta_{1,m} + \nu_1\zeta_{1,m}^\top\omega_1(t_m) - \|\zeta_{2,m}\|^2 \\ &\quad - \beta\nu_1\nu_2\omega_1(t_m)^\top(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L})\omega_1(t_m) - \alpha\|\omega_2(t_m)\|^2 \\ &\quad - \nu_2\omega_2(t_m)^\top(\mathrm{I}_\mathfrak{h} \otimes \mathsf{L})\zeta_{1,m} > 0,\end{aligned}$$

for all $m \in \mathbb{Z}_{\geq 1}$, where we have used (13) to write an element of $\mathcal{L}_{X_{\mathrm{dac+(L\partial,\partial)}}} V(I, S, \omega_1, \omega_2)$. Letting $\eta_m^{(k)} = \zeta_{1,m}^{(k)} - \frac{1}{n}(\mathbf{1}_n^\top\zeta_{1,m}^{(k)})\mathbf{1}_n$, using (1), and using the relation $\|\omega_1^{(k)}(t_m) - \frac{1}{n}(\mathbf{1}_n^\top\omega_1^{(k)}(t_m))\mathbf{1}_n\|^2 = \|\omega_1^{(k)}(t_m)\|^2 - \frac{1}{n}(\mathbf{1}_n^\top\omega_1^{(k)}(t_m))^2$, the above inequality can be rewritten as

$$\begin{aligned}&\gamma_m^\top M\gamma_m + \tfrac{1}{n}\nu_1\sum_{k\in\mathcal{K}}(\mathbf{1}_n^\top\zeta_{1,m}^{(k)})(\mathbf{1}_n^\top\omega_1^{(k)}(t_m)) - \|\zeta_{2,m}\|^2 \\ &\quad + \tfrac{\beta\nu_1\nu_2}{2n}\lambda_2(\mathsf{L} + \mathsf{L}^\top)\sum_{k\in\mathcal{K}}(\mathbf{1}_n^\top\omega_1^{(k)}(t_m))^2 > 0,\end{aligned} \tag{15}$$

with $\gamma_m = (\eta_m; \omega_1(t_m); \omega_2(t_m))$. Using (14) on (15),

$$\begin{aligned}&\gamma_m^\top M\gamma_m - \|\zeta_{2,m}\|^2 + \tfrac{\nu_1}{nm}\sum_{k\in\mathcal{K}}\left|\mathbf{1}_n^\top\zeta_{1,m}^{(k)}\right| \\ &\quad + \tfrac{\beta\nu_1\nu_2\mathfrak{h}}{2nm^2}\lambda_2(\mathsf{L} + \mathsf{L}^\top) > 0\end{aligned} \tag{16}$$

for all $m \in \mathbb{Z}_{\geq 1}$. Next, we consider two cases, depending on whether the sequence $\{(I(t_m), S(t_m))\}_{m=1}^\infty$ is (a) bounded or (b) unbounded. In case (a), $\{(\omega_1(t_m), \omega_2(t_m))\}_{m=1}^\infty$ must be unbounded. Since M is negative definite, we have $\gamma_m^\top M\gamma_m \leq \lambda_{\max}(M)\|(\omega_1(t_m), \omega_2(t_m))\|^2$. Thus, by (16)

$$\begin{aligned}&\lambda_{\max}(M)\|(\omega_1(t_m),\omega_2(t_m))\|^2 + \tfrac{\nu_1}{nm}\sum_{k\in\mathcal{K}}\left|\mathbf{1}_n^\top\zeta_{1,m}^{(k)}\right| \\ &\quad + \tfrac{\beta\nu_1\nu_2\mathfrak{h}}{2nm^2}\lambda_2(\mathsf{L} + \mathsf{L}^\top) > 0.\end{aligned} \tag{17}$$

Since $\partial_I f^\epsilon$ is locally bounded and $\{(I(t_m), S(t_m))\}_{m=1}^\infty$ is bounded, we deduce $\{\zeta_{1,m}\}$ is bounded [19, Proposition 6.2.2]. Combining these facts with $\lambda_{\max}(M) < 0$ and $\|(\omega_1(t_m), \omega_2(t_m))\| \to \infty$, one can find $\bar{m} \in \mathbb{Z}_{\geq 1}$ such that (17) is violated for all $m \geq \bar{m}$, a contradiction. Now consider case (b) where $\{(I(t_m), S(t_m))\}_{m=1}^\infty$ is unbounded. We divide this case further into two, based on the sequence $\{\sum_{k=1}^{\mathfrak{h}}\left|\mathbf{1}_n^\top\zeta_{1,m}^{(k)}\right|\}_{m=1}^\infty$ being bounded or not. Using $\gamma_m^\top M\gamma_m \leq \lambda_{\max}(M)\|\eta_m\|^2$, the inequality (16) implies

$$\begin{aligned}&\lambda_{\max}(M)\|\eta_m\|^2 - \|\zeta_{2,m}\|^2 + \frac{\nu_1}{nm}\sum_{k=1}^{\mathfrak{h}}\left|\mathbf{1}_n^\top\zeta_{1,m}^{(k)}\right| \\ &\quad + \frac{\beta\nu_1\nu_2\mathfrak{h}}{2nm^2}\lambda_2(\mathsf{L} + \mathsf{L}^\top) > 0.\end{aligned} \tag{18}$$

Consider the case when $\{\sum_{k=1}^{\mathfrak{h}}\left|\mathbf{1}_n^\top\zeta_{1,m}^{(k)}\right|\}_{m=1}^\infty$ is unbounded. Partition $\mathcal{K}$ into disjoint sets $\mathcal{K}_u$ and $\mathcal{K}_b$ such that $\left|\mathbf{1}_n^\top\zeta_{1,m}^{(k)}\right| \to \infty$ for all $k \in \mathcal{K}_u$ and $\{\left|\mathbf{1}_n^\top\zeta_{1,m}^{(k)}\right|\}_{m=1}^\infty$ is uniformly bounded for all $k \in \mathcal{K}_b$. For convenience, rewrite (18) as $\sum_{k=1}^{\mathfrak{h}} U_{k,m} + \frac{Z_1}{m} > 0$, where $Z_1 = \frac{\beta\nu_1\nu_2\mathfrak{h}}{2nm}\lambda_2(\mathsf{L} + \mathsf{L}^\top)$ and, for each $k \in \mathcal{K}$,

$$U_{k,m} = \lambda_{\max}(M)\|\eta_m^{(k)}\|^2 - \|\zeta_{2,m}^{(k)}\|^2 + \frac{\nu_1}{nm}\left|\mathbf{1}_n^\top\zeta_{1,m}^{(k)}\right|.$$

By definition of $\mathcal{K}_b$, there exists $Z_2 > 0$ with $\sum_{k\in\mathcal{K}_b} U_{k,m} \le \frac{Z_2}{m}$. Hence, if (18) holds for all $m \in \mathbb{Z}_{\geq 1}$, then so is

$$\sum_{k\in\mathcal{K}_u} U_{k,m} + \frac{Z_1+Z_2}{m} > 0.$$

Next we show that for each $k \in \mathcal{K}_u$ there exists $m_k \in \mathbb{Z}_{\geq 1}$ such that $U_{k,m} + \frac{Z_1+Z_2}{m} < 0$ for all $m \ge m_k$. This will lead to the desired contradiction. Assume without loss of generality that $\mathbf{1}_n^\top \zeta_{1,m}^{(k)} \to \infty$ (reasoning for the case when the sequence approaches negative infinity follows analogously). Then, for

$$\lambda_{\max}(M)\|\eta_m^{(k)}\|^2 - \|\zeta_{2,m}\|^2 + \frac{\nu_1}{nm}\left|\mathbf{1}_n^\top \zeta_{1,m}^{(k)}\right| + \frac{Z_1+Z_2}{m} > 0,$$

for all $m \in \mathbb{Z}_{\geq 1}$, we require $(\zeta_{1,m}^{(k)})_i \to \infty$ for all $i \in \{1,\dots,n\}$. Indeed, otherwise, recalling that $\eta_m^{(k)} = \zeta_{1,m}^{(k)} - \frac{1}{n}(\mathbf{1}_n^\top \zeta_{1,m}^{(k)})\mathbf{1}_n$, it can be shown that there exist an $\bar{m}$ such that

$$\lambda_{\max}\|\eta_m^{(k)}\|^2 < \frac{\nu_1}{nm}\left|\mathbf{1}_n^\top \zeta_{1,m}^{(k)}\right| + \frac{Z_1+Z_2}{m} \quad \text{for all } m \ge \bar{m}.$$

Note that from Lemma 4.3 we have $\|\zeta_{1,m}^{(k)} - \zeta_{2,m}^{(k)}\|_\infty \le \frac{\mathfrak{h}+4}{\epsilon}$ which further implies that $(\zeta_{2,m}^{(k)})_i \to \infty$ for all $i \in \{1,\dots,n\}$. With these facts in place, we write

$$U_{k,m} + \frac{Z_1+Z_2}{m} < -\sum_{i=1}^{n}(\zeta_{2,m}^{(k)})_i^2 + \frac{\nu_1}{m}\left|\sum_{i=1}^{n}(\zeta_{1,m}^{(k)})_i\right| + \frac{Z_1+Z_2}{m}$$

and deduce that there exists an $m_k \in \mathbb{Z}_{\geq 1}$ such that the right-hand side of the above expression is negative for all $m \ge m_k$, which is what we wanted to show.

Finally, consider the case when the sequence $\{\sum_{k=1}^{\mathfrak{h}} \left|\mathbf{1}_n^\top \zeta_{1,m}^{(k)}\right|\}_{m=1}^{\infty}$ is bounded. For (18) to be true for all $m \in \mathbb{Z}_{\geq 1}$, we require $\|\gamma_m\| \to 0$ and $\|\zeta_{2,m}\| \to 0$ as $m \to \infty$. This further implies that $\eta_m \to 0$ and, from Assumption 4.2, this is only possible if $\{(I(t_m), S(t_m))\}_{m=1}^{\infty}$ is bounded, which is a contradiction. ■

Remark 4.6: (General setup for storage: revisited): The $\mathtt{dac}$+$(\mathsf{L}\partial,\partial)$ dynamics (10) can be modified to scenarios that include more general descriptions of storage capabilities, as in Remark 3.1. For instance, if only a subset of units have storage capabilities, the only modification is to set the variables $\{S_i^{(k)}\}_{i\in\mathcal{V}_g, k\in\mathcal{K}}$ to zero and execute (10b) only for the variables $\{S_i^{(k)}\}_{i\in\mathcal{V}_{gs}, k\in\mathcal{K}}$. The resulting strategy converges to the solution set of the corresponding DEDS problem. •

Remark 4.7: (Distributed selection of design parameters): The implementation of the $\mathtt{dac}$+$(\mathsf{L}\partial,\partial)$ dynamics requires the selection of parameters $\alpha, \beta, \nu_1, \nu_2, \epsilon$ satisfying (8) and (11). Condition (11) involves knowledge of network-wide quantities, but the units can resort to various distributed procedures to collectively select appropriate values. Regarding (8), an upper bound on the denominator of the right-hand side can be computed aggregating, using consensus, the difference between the max and the min values that each DER's aggregate cost function takes in its respective feasibility set (neglecting load conditions). The challenge for the units, however, is to estimate the parameter ρ if it is not known a priori. •

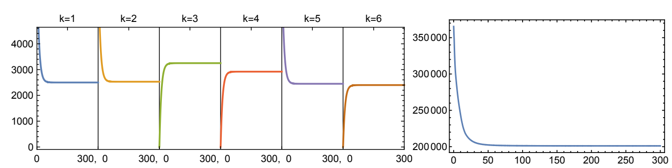
V. Simulations

We illustrate the application of the $\mathtt{dac}$+$(\mathsf{L}\partial,\partial)$ dynamics to solve the DEDS problem for a group of $n = 10$ generators with communication defined by a directed ring with bi-directional edges $\{(1,5),(2,6),(3,7),(4,8)\}$ (all edge weights are 1). The planning horizon is $\mathfrak{h} = 6$ and the load profile consists of the external load $L = (1950, 1980, 2700, 2370, 1900, 1850)$ and the load at each generator i for each slot k given by $\tilde{l}_i^{(k)} = 10i$. Thus, for each slot k, $\tilde{l}^{(k)} = \sum_{i=1}^{10}\tilde{l}_i^{(k)} = 550$ and so, $l = (2500, 2530, 3250, 2920, 2450, 2400)$. Generators have storage capacities determined by $C^M = 100\mathbf{1}_n$ and $C^m = S^{(0)} = 5\mathbf{1}_n$. The cost function of each unit is quadratic and constant across time. Table I details the cost function coefficients, generation limits, and ramp constraints, which are modified from the data for 39-bus New England system [22].

Unit	a_i	b_i	c_i	P_i^m	P_i^M	R_i^l	R_i^u
1	240	7.0	0.0070	0	1040	120	80
2	200	10.0	0.0095	0	646	90	50
3	220	8.5	0.0090	0	725	100	65
4	200	11.0	0.0090	0	652	90	50
5	220	10.5	0.0080	0	508	90	50
6	190	12.0	0.0075	0	687	90	50
7	200	10.0	0.0100	0	580	120	80
8	170	9.0	0.0090	0	564	90	50
9	190	11.0	0.0072	0	865	100	65
10	220	8.8	0.0080	0	1100	90	50

TABLE I
Cost coefficients (a_i, b_i, c_i) and bounds P_i^M, P_i^m, R_i^l, R_i^u. The cost function of i is $f_i(P_i) = a_i + b_iP_i + c_iP_i^2$.

Figure 1 illustrates the evolution of the total power injected at each time slot and the total cost incurred by the network, respectively. As established in Theorem 4.5 and shown in Figure 2, the total injection asymptotically converges to the load profile l, the total aggregate cost converges to the minimum 201092 and the converged solution satisfies (5c)-(5f).

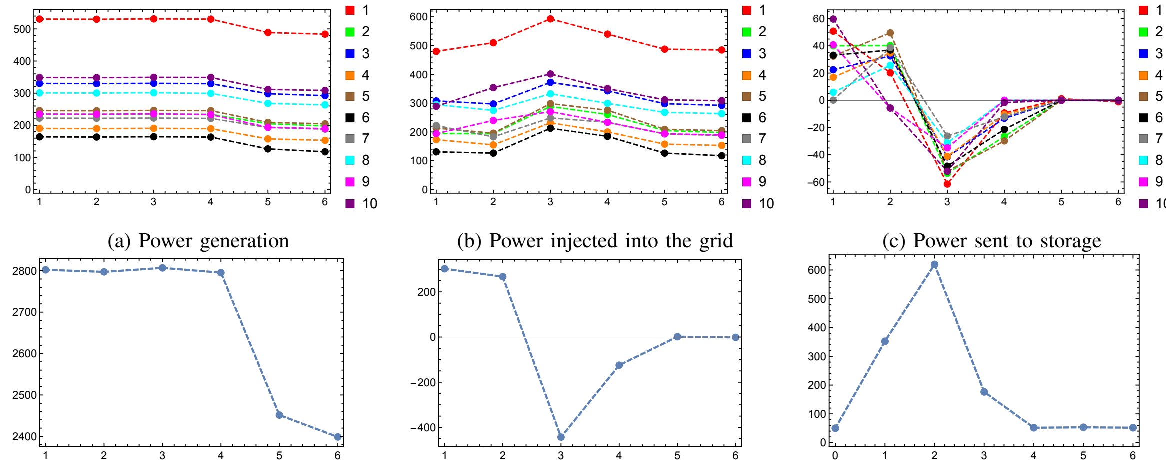


(a) Total injection (b) Total cost

Fig. 1. Illustration of the execution of $\mathtt{dac}$+$(\mathsf{L}\partial,\partial)$ dynamics for a network of 10 generators with communication topology given by a directed ring among the generators with bi-directional edges $\{(1,5),(2,6),(3,7),(4,8)\}$ where all edge weights are 1. Table I gives the box constraints, the ramp constraints, and the cost functions. The load profile is $l = (2500, 2530, 3250, 2920, 2450, 2400)$ and $C^M = 100\mathbf{1}_n$, $C^m = S^{(0)} = 5\mathbf{1}_n$. Plots (a) and (b) show the time evolution of the total injection at each time slot and the aggregate cost along a trajectory of the $\mathtt{dac}$+$(\mathsf{L}\partial,\partial)$ dynamics starting at $I(0) = (P^M, P^M, P^m, P^m, P^M, P^m)$, $S(0) = z(0) = v(0) = \mathbf{0}_{n\mathfrak{h}}$. The parameters are $\epsilon = 0.007$, $\alpha = 4$, $\beta = 10$, and $\nu_1 = \nu_2 = 0.65$ (which satisfy conditions (8) and (11)).

VI. Conclusions

We have studied the DEDS problem for a group of generators with storage capabilities that communicate over a strongly connected, weight-balanced digraph. Using exact



(a) Power generation (b) Power injected into the grid (c) Power sent to storage

(d) Total power generated (e) Total power sent to storage (f) Aggregate state of stored power

Fig. 2. Plots (a) to (f) illustrate the solution obtained in Figure 1. Plots (b) and (c) show the power injected and power sent to storage across the time horizon, with unique colors for each generator. These values add up to the total generation in (a). The collective behavior is represented in (d)-(f), where we plot the total power generated, the total power sent to storage, and the aggregate of the power stored in the storage units, respectively. The profile of total injection is the same as that of load profile. Since the time-independent cost is quadratic with positive coefficients and the storage capacity is large enough, one can show that the optimal strategy is to produce the same power, i.e., $\frac{1}{5}\sum_{k=1}^{5} l^{(k)}$, at each time slot $k = 1, \dots, 5$, as seen in (a) and (d). The initial excess generation (due to the lower required load) at slots $k = 1, 2$ is stored and used in slots $k = 3, 4, 5, 6$, as indicated in (e) and (f).

penalty functions, we have provided an alternative problem formulation, upon which we have built to design the distributed $\mathtt{dac+}(\mathsf{L}\partial, \partial)$ dynamics. This dynamics provably converges to the set of solutions of the problem from any initial condition. For future work, we plan to extend the scope of our formulation to include power flow equations, constraints on the power lines, various losses, and stochasticity of the available data (loads, costs, and generator availability). We also intend to explore the use of our dynamics as a building block in solving grid control problems across different time scales (e.g., implementations at long time scales on high-inertia generators and at short time scales on low-inertia generators in the face of highly-varying demand) and hierarchical levels (e.g., in multi-layer architectures where aggregators at one layer coordinate their response to a request for power production, and feed their decisions as load requirements to the devices in lower layers).

References

[1] A. Cherukuri and J. Cortés, "Distributed dynamic economic dispatch of power generators with storage," in *IEEE Conf. on Decision and Control*, (Osaka, Japan), pp. 2365–2370, 2015.

[2] A. D. Domínguez-García, S. T. Cady, and C. N. Hadjicostis, "Decentralized optimal dispatch of distributed energy resources," in *IEEE Conf. on Decision and Control*, (Hawaii, USA), pp. 3688–3693, 2012.

[3] 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.

[4] 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*, vol. 30, no. 1, pp. 35–45, 2014.

[5] A. Cherukuri and J. Cortés, "Distributed generator coordination for initialization and anytime optimization in economic dispatch," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 226–237, 2015.

[6] A. Cherukuri and J. Cortés, "Initialization-free distributed coordination for economic dispatch under varying loads and generator commitment," *Automatica*, 2014. Available at http://carmenere.ucsd.edu/jorge.

[7] X. Xia and A. M. Elaiw, "Optimal dynamic economic dispatch of generation: A review," *Electric Power Systems Research*, vol. 80, no. 8, pp. 975–986, 2010.

[8] M. D. Ilić, L. Xie, and J. Joo, "Efficient coordination of wind power and price-responsive demand – part 1: theoretical foundations," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1875–1884, 2011.

[9] X. Xia, J. Zhang, and A. Elaiw, "An application of model predictive control to the dynamic economic dispatch of power generation," *Control Engineering Practice*, vol. 19, no. 6, pp. 638–648, 2011.

[10] Z. Li, W. Wu, B. Zhang, H. Sun, and Q. Guo, "Dynamic economic dispatch using Lagrangian relaxation with multiplier updates based on Quasi-Newton method," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4516–4527, 2013.

[11] A. Hooshmand, J. Mohammadpour, H. Malki, and H. Daneshi, "Power system dynamic scheduling scheduling with high penetration of renewable sources," in *American Control Conference*, (Washington, D.C.), pp. 5847–5852, June 2013.

[12] D. Zhu and G. Hug, "Decomposed stochastic model predictive control for optimal dispatch of storage and generation," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2044–2053, 2014.

[13] Y. Zhang, N. Gatsis, and G. B. Giannakis, "Robust energy management for microgrids with high-penetration renewables," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 4, pp. 944–953, 2013.

[14] A. Soroudi, A. Rabiee, and A. Keane, "Stochastic real-time scheduling of wind-thermal generation units in an electric utility," *IEEE Systems Journal*, 2015. To appear.

[15] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*. Applied Mathematics Series, Princeton University Press, 2009. Electronically available at http://coordinationbook.info.

[16] J. Cortés, "Discontinuous dynamical systems - a tutorial on solutions, nonsmooth analysis, and stability," *IEEE Control Systems Magazine*, vol. 28, no. 3, pp. 36–73, 2008.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[18] D. P. Bertsekas, "Necessary and sufficient conditions for a penalty method to be exact," *Mathematical Programming*, vol. 9, no. 1, pp. 87–99, 1975.

[19] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I*. Grundlehren Text Editions, New York: Springer, 1993.

[20] 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*, vol. 25, no. 13, pp. 1941–1966, 2015.

[21] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, 3 ed., 2002.

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