Distributed dynamic economic dispatch of power generators with storage

Ashish Cherukuri Jorge Cortés

***Abstract*— This paper considers the dynamic economic dispatch problem for a group of generators with storage that communicate over a weight-balanced strongly connected digraph. The objective of the generators is to collectively meet a certain load profile, specified over a finite time horizon, while minimizing the aggregate cost. At each time slot, each generator decides on the amount of generated power and the amount sent to/drawn from the storage unit. The amount injected into the grid by each generator to satisfy the load is equal to the difference between the generated and stored powers. Additional constraints on the generators 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 dynamical systems, one estimating the mismatch in the injection and the total load at each time slot, and another using this estimate as a feedback 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 brings in 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 the privacy of the entities involved. With this context in mind, our objective here is to provide a distributed algorithmic solution to the dynamic economic dispatch with storage problem. 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 the optimization of the total cost by a group of generators over a single time slot while collectively meeting a specified load and respecting individual constraints. Traditional solutions to the SED problem have been centralized, but the advent of distributed generation has motivated the design of decentralized algorithmic solutions, see e.g., [1], [2], [3] and our own work [4], [5]. As argued in [6], [7], the dynamic version of the problem, termed dynamic economic dispatch (DED), results in better grid control capabilities as it incorporates optimal planning across a time horizon and specifically accounts for ramp limits in generation and variability of power commitment from renewable sources. Conventional solution methods to the DED problem are centralized [6]. Recent works [7], [8] 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 [9] proposes a Lagrangian relaxation method to solve the DED problem, but the implementation requires a master agent to perform dual updates and communicate with each of the generators, which in turn solve their respective subproblems. MPC methods have also been employed by [10] 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 [11], dual decomposition [12], and optimal condition decomposition [13] methods. However, these methods are either centralized or need a central master unit to coordinate the decentralized subproblems of each generator. Finally, [14] considers strategic scenarios where generators seek to optimize their own profit. Under a suitable pricing mechanism, the global optimization of the DEDS problem corresponds to a Nash equilibrium of the game between the generators and the independent system operator, albeit no algorithm is provided to compute it.

Statement of contributions: Our starting point is the formulation of the DEDS problem for a group of power generators communicating over a weight-balanced strongly connected digraph. Since the individual cost functions are convex and all constraints are linear, the problem is convex in its decision variables (the power to be injected and the power to be sent to storage by each generator at each time slot). Using an exact penalty function approach, we reformulate the DEDS problem as an optimization that retains the load constraints but strips off the inequality constraints. The structure of the modified problem guides our design of the provably-correct distributed strategy termed "dynamic average consensus (`dac`) + Laplacian nonsmooth gradient ($\mathrm{L}\partial$) + nonsmooth gradient (∂)" dynamics to solve the DEDS problem starting from any initial condition. Our algorithm design consists of two interconnected systems. A first block allows generators to track, using `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

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

nonsmooth-gradient dynamics on storage variables) and an error-correcting component that uses the mismatch estimated by the dac system to push the total injection towards the load for each time slot. Simulations illustrate our results. Proofs are omitted for space reasons and will appear elsewhere.

Notation: Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, $\mathbb{R}_{<0}$, $\mathbb{Z}_{\geq 1}$ denote the set of real, nonnegative real, positive real, negative real and positive integer numbers, respectively. The 2- and ∞-norm on $\mathbb{R}^n$ are denoted by $\|\cdot\|$ and $\|\cdot\|_\infty$, respectively. Let $B(x,\delta) = \{y \in \mathbb{R}^n \mid \|y-x\| < \delta\}$ denote the open ball 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 eigenvalue of A is $\lambda_{\min}(A)$. 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 notation $\mathbf{0}_n = (0,\ldots,0) \in \mathbb{R}^n$, $\mathbf{1}_n = (1,\ldots,1) \in \mathbb{R}^n$, 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 their 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 $i \in \{1,\ldots,n\}$. For $\mathfrak{h} > 0$, given $y \in \mathbb{R}^{n\mathfrak{h}}$ and $k \in \{1,\ldots,\mathfrak{h}\}$, the vector containing the $nk-n+1$ to nk components of y is $y^{(k)} \in \mathbb{R}^n$, and so, $y = (y^{(1)}; y^{(2)}; \ldots; y^{(\mathfrak{h})})$. We let $[u]^+ = \max\{0,u\}$ for $u \in \mathbb{R}$. Finally, 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 $\mathsf{A} \in \mathbb{R}^{n\times n}_{\geq 0}$ 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, $N^{\text{out}}(v_i) = \{v_j \in \mathcal{V} \mid (v_i, v_j) \in \mathcal{E}\}$ and $N^{\text{in}}(v_i) = \{v_j \in \mathcal{V} \mid (v_j, v_i) \in \mathcal{E}\}$ are the sets of out- and in-neighbors. The weighted out-degree and in-degree of v_i, $i \in \{1,\ldots,n\}$, are respectively, $d^{\text{out}}(v_i) = \sum_{j=1}^n a_{ij}$ and $d^{\text{in}}(v_i) = \sum_{j=1}^n a_{ji}$. The *Laplacian* matrix is $\mathsf{L} = \mathsf{D}_{\text{out}} - \mathsf{A}$, where D_{out} is the diagonal matrix defined by $(\mathsf{D}_{\text{out}})_{ii} = d^{\text{out}}(v_i)$, for all $i \in \{1,\ldots,n\}$. 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* if $d^{\text{out}}(v) = d^{\text{in}}(v)$, for all $v \in \mathcal{V}$. Equivalently, $\mathcal{G}$ is weight-balanced if and only if $\mathbf{1}_n^\top \mathsf{L} = 0$ if and only if $\mathsf{L} + \mathsf{L}^\top$ 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 $x \in \mathbb{R}^n$ one can deduce from [16, Chapter 8] that

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

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 [17]. 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 [17] for these definitions. A convex function is regular. A set-valued map $\mathcal{H} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *upper semicontinuous (u.s.c)* 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, u.s.c, and takes non-empty compact convex values. For $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $(x,y) \mapsto f(x,y)$, 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 [17], [5] regarding the properties of differential inclusions,

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

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 F is locally bounded, u.s.c, and takes non-empty, compact, and convex values, then existence of solutions is guaranteed. The equilibria of (2) is $\mathrm{Eq}(F) = \{x \in \mathbb{R}^n \mid 0 \in F(x)\}$.

Constrained optimization and exact penalty functions: Here, we introduce some notions on constrained convex optimization problems and exact penalty functions [18], [19]. Consider the optimization problem,

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

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 $\lambda \in \mathbb{R}^m_{\geq 0}$ and $\nu \in \mathbb{R}^p$ such that

$$g(x) \leq \mathbf{0}_m, \quad h(x) = \mathbf{0}_p, \quad \lambda^\top g(x) = 0,$$
$$\nabla f(x) + \textstyle\sum_{i=1}^m \lambda_i \nabla g_i(x) + \sum_{i=1}^p \nu_i \nabla h_i(x) = 0.$$

If strong duality holds then, a point is a solution of (3) if and only if 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 2.1: (Bound on Lagrange multiplier [20, Remark 2.3.3]): Assume that (3) satisfies 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 use an exact penalty function to eliminate the inequality constraints in (3) while keeping the equality constraints intact. To this end, we follow [19] to construct a nonsmooth exact penalty function $f^\epsilon : \mathbb{R}^n \to \mathbb{R}$, $f^\epsilon(x) = f(x) + \frac{1}{\epsilon}\sum_{i=1}^m [g_i(x)]^+$, with $\epsilon > 0$, and define

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

Note that f^ϵ is convex as f and $t \mapsto \frac{1}{\epsilon}[t]^+$ are convex. Hence, the problem (4) is convex. The next result, see e.g. [19, Proposition 1], identifies conditions under which the solutions of the problems (3) and (4) coincide.

Proposition 2.2: (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 $n \in \mathbb{Z}_{\geq 1}$ power generators with communication topology described by a strongly connected and weight-balanced digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathsf{A})$. Each generator corresponds to a vertex. An edge (i, j) represents the capability of j to transmit information to i. Each generator i is equipped with a storage unit with minimum and maximum capacities $C_i^m \in \mathbb{R}_{\geq 0}$ and $C_i^M \in \mathbb{R}_{>0}$, resp. The group collectively aims to meet a power demand profile during a finite-time horizon $\mathcal{K} = \{1, \dots, \mathfrak{h}\}$ specified by $l \in \mathbb{R}^{\mathfrak{h}}_{>0}$, i.e., $l^{(k)}$ is the demand at time slot $k \in \mathcal{K}$. We assume the load profile is known to an arbitrarily selected generator r. Along with load satisfaction, the group aims to minimize the total generation cost and satisfy individual constraints.

Each generator decides at every time slot in $\mathcal{K}$, the amount of power it generates, the portion of it that it injects into the grid to meet the load, and the remaining part that it sends to the storage unit. Specifically, we let $I_i^{(k)} \in \mathbb{R}$ and $S_i^{(k)} \in \mathbb{R}$ denote the power injected into the grid and the power sent to the storage, resp., by the generator i at time slot k. The power generated by i at k is then $I_i^{(k)} + S_i^{(k)}$. For convenience, 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, resp. The load satisfaction at each time slot reads as $\mathbf{1}_n^\top I^{(k)} = l^{(k)}$, for all $k \in \mathcal{K}$. The cost of generation for i at time k is given by $f_i^{(k)} : \mathbb{R} \to \mathbb{R}_{\geq 0}$, which is assumed to be convex and continuously differentiable. Thus, the cost incurred by i at time k to generate power $I_i^{(k)} + S_i^{(k)}$ is $f_i^{(k)}(I_i^{(k)} + S_i^{(k)})$. We denote the network cost function at time k by $f^{(k)} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, i.e., given $(I^{(k)}, S^{(k)})$ the network cost at time slot k is

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

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

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

Note that the functions $\{f^{(k)}\}_{k\in\mathcal{K}}$ and f are also convex and continuously differentiable. Next, we describe the physical constraints on the generators. Each generator'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 generator 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 $R_i^u \in \mathbb{R}_{>0}$ and $-R_i^l \in \mathbb{R}_{<0}$, resp. At each time slot, the power injected into the grid by each generator must be nonnegative, i.e., $I_i^{(k)} \geq 0$. Also, 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 before 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,

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

We refer to (5b)–(5f) as the *load*, *box*, *storage limits*, *injection*, and *ramp* constraints, resp. We denote by $\mathcal{F}_{\text{DEDS}}$ and $\mathcal{F}^*_{\text{DEDS}}$ the feasibility set and the solution set of the DEDS problem (5), resp., 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 constraints (5b)–(5f) are affine. Additionally, we assume that the DEDS problem satisfies the strong Slater condition with parameter $\rho \in \mathbb{R}_{>0}$ and feasible point $(I^\rho, S^\rho) \in \mathbb{R}^{2n\mathfrak{h}}$. Our goal is to design a distributed algorithmic solution that allows the network of generators interacting over $\mathcal{G}$ to solve the DEDS problem.

IV. Distributed algorithm for the DEDS problem

Our design strategy builds on an alternative formulation of the optimization problem using penalty functions (cf. Section IV-A). This step allows us to get rid of the inequality constraints, resulting into an optimization whose particular 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

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

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 u.s.c. 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 the penalty parameter ϵ such that the modified DEDS problem and the DEDS problem have the exact same solutions. The proof is a direct application of Lemma 2.1 and Proposition 2.2 using the fact 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

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

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

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

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

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

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 $p \in [0, \bar{p})$. •

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. The design of this routine combines ideas of Laplacian-gradient dynamics [4] and dynamic average consensus algorithms [21]. Consider the set-valued dynamics,

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

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 can be understood as 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 $l^{(k)}$ and the current injection level $\mathbf{1}_n^\top I^{(k)}$ for generator 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 the set-valued map $X_{\texttt{dac+}(\mathsf{L}\partial,\partial)} : \mathbb{R}^{4n\mathfrak{h}} \rightrightarrows \mathbb{R}^{4n\mathfrak{h}}$. Note that $\text{Eq}(X_{\texttt{dac+}(\mathsf{L}\partial,\partial)}) = \mathcal{F}^*_{\text{DEDS}}$ and since $\partial_I f^\epsilon$ and $\partial_S f^\epsilon$ are locally bounded, u.s.c and take nonempty convex compact values, the solutions of $X_{\texttt{dac+}(\mathsf{L}\partial,\partial)}$ exist (cf. Section II).

Remark 4.3: (Distributed implementation of the $\texttt{dac+}(\mathsf{L}\partial, \partial)$ *dynamics):* Writing the (z,v) dynamics componentwise, note that for each i and each k, $(\dot{z}_i^{(k)}, \dot{v}_i^{(k)})$ can be computed using $(z_i^{(k)}, \{z_j^{(k)}\}_{j \in N^{\text{out}}}, v_i^{(k)}, I_i^{(k)})$ only. Hence, (10c) and (10d) can be implemented in a distributed manner where each generator only requires information from its out-neighbors. Further, f^ϵ can be written as

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

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 depends on each i's own state and for the I-dynamics, i needs its out-neighbors' information.•

Next, we establish the convergence of the $\texttt{dac+}(\mathsf{L}\partial, \partial)$ dynamics. The proof involves a refinement of the LaSalle Invariance Principle for differential inclusions from [5].

Theorem 4.4: (Convergence of the $\texttt{dac+}(\mathsf{L}\partial, \partial)$ *dynamics to the solutions of the* DEDS *problem):* Let $\mathcal{F}^*_{\text{DEDS}}$ 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 $\mathbb{R}^{n\mathfrak{h}} \times \mathbb{R}^{n\mathfrak{h}} \times \mathbb{R}^{n\mathfrak{h}} \times (\mathcal{H}_0)^{\mathfrak{h}}$ converges to $\mathcal{F}^*_{\text{aug}} = \{(I,S,z,v) \in \mathcal{F}^*_{\text{DEDS}} \times \{0\} \times \mathbb{R}^{n\mathfrak{h}} \mid v = \nu_2(l \otimes e_r - I)\}$.

V. Simulations

Here, we show the application of the $\texttt{dac+}(\mathsf{L}\partial, \partial)$ dynamics to solve the DEDS problem for 6 generators communicating over a digraph with adjacency matrix in Table II(b).

1) Case 1: Constant cost across time slots: The planning horizon is $\mathfrak{h} = 5$ with load profile l is given in Figure 1(a). Generators have storage capacities $C^M = 100\mathbf{1}_n$ and $C^m =$

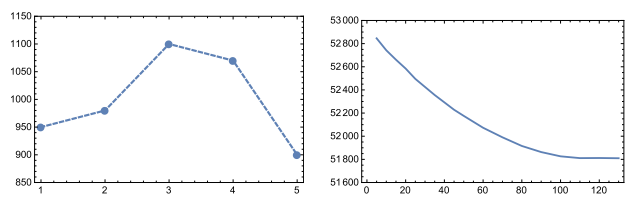
$S^{(0)} = 5\mathbf{1}_n$. For each generator, the cost is quadratic and it remains constant across time. Table I details the cost function coefficients, box and ramp constraints taken from [8].

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	100	500	120	80
2	200	10.0	0.0095	50	200	90	50
3	220	8.5	0.0090	80	300	100	65
4	200	11.0	0.0090	50	150	90	50
5	220	10.5	0.0080	50	200	90	50
6	190	12.0	0.0075	50	120	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 $f_i(P_i) = a_i + b_i P_i + c_i P_i^2$.

Figures 2(a) and 2(b) illustrate the evolution of the total power injected at each time slot and the total cost incurred by the network, resp., along a trajectory of the dynamics. As established in Theorem 4.4 and shown in Figure 2 (c)-(h), the total injection asymptotically converges to the load profile l, the total aggregate cost converges to the minimum 59459 and the converged solution satisfies the constraints (5c)-(5f).



(a) Load profile (b) Cost vs Storage capacity

Fig. 1. (a) Load profile $l = (950, 980, 1100, 1070, 900)$ over a 5 time-slot horizon and (b) dependence of the aggregate cost as a function of the storage capacity available at each generator. In (b), cost functions at time slots $k = 1, 2$ are given in Table II(a) and at time slots $k = 3, 4, 5$ are given in Table I. Box and ramp constraints given in Table I.

2) Case 2: Changing cost across time slots: Here, we retain the setup of Case 1 except for the cost functions: for $k = 1, 2$, the cost is defined by $(\bar{a}, \bar{b}, \bar{c})$ given in Table II(a) and, for $k = 3, 4, 5$, the cost is given by (a, b, c), cf. Table I. With these costs, producing energy is cheaper earlier than later which can be attributed to evolving sunlight or changing wind patterns. Figure 3 (g)-(i) show the converged values that satisfies all the constraints with the total cost 51833. As expected, the network generates more energy than required initially to store it and use it later when the cost is higher.

Unit	$\bar{a}_i$	$\bar{b}_i$	$\bar{c}_i$
1	200	5.0	0.0060
2	100	3.0	0.0090
3	200	7.5	0.0090
4	180	6.0	0.0090
5	200	9.0	0.0080
6	180	12.0	0.006

(a)

$$\begin{bmatrix} 0 & 2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(b)

TABLE II

COST COEFFICIENTS $(\bar{a}_i, \bar{b}_i, \bar{c}_i)$ AND THE ADJACENCY MATRIX.

In Case 2, we have also studied the effect of storage capacity on the optimal total cost incurred by the network. As expected, cf. Figure 1(b), the optimal cost decreases as the storage capacity increases because higher capacity enables the network to produce more at lower cost.

VI. CONCLUSIONS

We have formulated the DEDS problem for a group of generators with storage capabilities that communicate over a strongly connected, weight-balanced digraph. Using exact penalty functions we have provided an alternative formulation of the problem that lead to the design of the distributed $\texttt{dac+}(\mathsf{L}\partial, \partial)$ dynamics. We have established that this dynamics converges to the set of solutions of the problem from any initial condition. For future work, we plan to extend our formulation to include power flow equations, constraints on the power lines, and various losses. Further, we wish to design robust distributed algorithms for stochastic versions of the problem that include uncertainties in load profiles, cost functions, and availability of generators across time slots.

ACKNOWLEDGMENTS

Research supported by NSF award ECCS-1307176.

REFERENCES

[1] 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*, (Hawaii, USA), pp. 3688–3693, Dec. 2012.

[2] 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.

[3] 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.

[4] 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, 2015.

[5] A. Cherukuri and J. Cortés, "Initialization-free distributed coordination for economic dispatch under varying loads and generator commitment," *Automatica*, 2014. Submitted.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

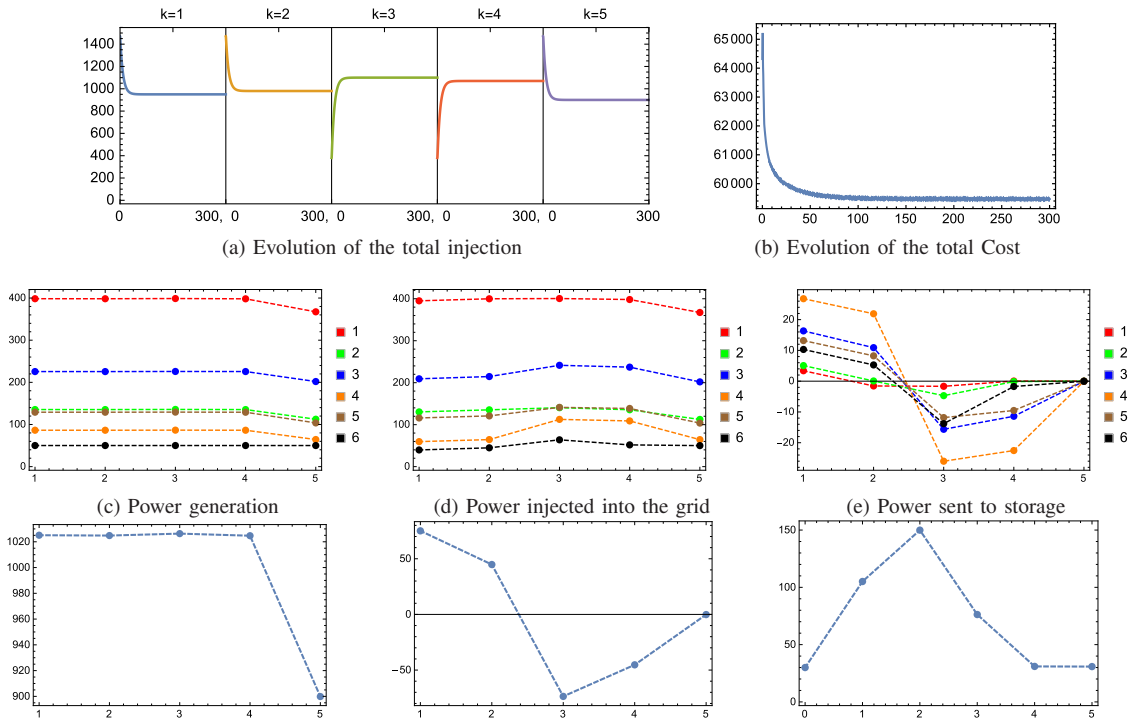
[12] Y. Zhang, N. Gatsis, and G. B. Giannakis, "Robust energy management for microgrids with high-penetration renewables," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 4, pp. 944–953, 2013.

[13] 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.

[14] W. Tang and R. Jain, "Dynamic economic dispatch among strategic generators with storage systems," in *Allerton Conf. on Communications, Control and Computing*, (Monticello, IL), Oct. 2014.

[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] D. S. Bernstein, *Matrix Mathematics*. Princeton University Press, 2 ed., 2009.

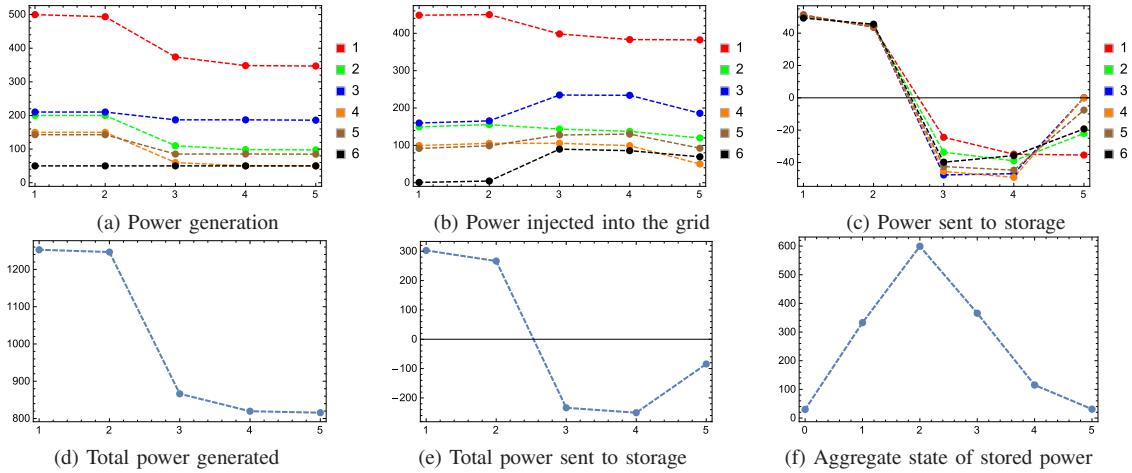


(a) Evolution of the total injection (b) Evolution of the total Cost

(c) Power generation (d) Power injected into the grid (e) Power sent to storage

(f) Total power generated (g) Total power sent to storage (h) Aggregate state of stored power

Fig. 2. Illustration of the execution of $\texttt{dac+}(\mathsf{L}\partial, \partial)$ dynamics for a network of 6 generators with communication topology given by the adjacency matrix in Table II(b). Table I gives the box constraints, the ramp constraints, and the cost functions. The load profile is given in Figure 1(a) and $C^M = 1001\mathbf{1}_n$, $C^m = S^{(0)} = 51_n$. 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 $\texttt{dac+}(\mathsf{L}\partial, \partial)$ dynamics starting at $I(0) = (P^M, P^M, P^m, P^m, P^M)$, $S(0) = z(0) = v(0) = \mathbf{0}_{n\mathfrak{h}}$. The parameters are $\epsilon = 0.007$, $\alpha = 4$, $\beta = 10$, and $\nu_1 = \nu_2 = 0.7$ (which satisfy (8) and (11)). Plots (c) to (h) illustrate the obtained solution, that is exactly same as that obtained from centralized solvers. Plots (d) and (e) 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 generation in (c). The collective behavior is given in (f)-(h), where we plot the total power generated, the total power sent to storage, and the aggregate of the power stored in the storage units, resp. Since the time-independent cost is quadratic with positive coefficients and the storage capacity is large enough, one can show that the optimal policy is to produce the same power, i.e., $\frac{1}{4}\sum_{k=1}^{4} l^{(k)}$, at each slot $k = 1, \ldots, 4$, as seen in plot (c) and (f). The initial excess generation (due to the lower load) at slots $k = 1, 2$ is stored and used in slots $k = 3, 4, 5$, as indicated in plots (g) and (h).



(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. 3. Solution, as determined by the $\texttt{dac+}(\mathsf{L}\partial, \partial)$ dynamics, of the DEDS_problem for the setup described in Figure 2 but with time-dependent quadratic cost functions. The cost functions are determined by the coefficients $(\bar{a}, \bar{b}, \bar{c})$ (cf. Table II(a)) for time slots $k = 1, 2$ and the coefficients (a, b, c) (cf. Table I) for $k = 3, 4, 5$. This example illustrates further the importance of storage. With lower cost functions in slots 1 and 2, the generators produce even more power (as compared to the solution in Figure 2) in the first two slots, using the excess stored power in later time slots when the cost is high.

[17] 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.

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

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

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

[21] 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.