1st-Order Dynamics on Nonlinear Agents for Resource Allocation over Uniformly-Connected Networks

Mohammadreza Doostmohammadian, Alireza Aghasi, Maria Vrakopoulou, Themistoklis Charalambous

Abstract—A general nonlinear 1st-order consensus-based solution for distributed constrained convex optimization is proposed with network resource allocation applications. The solution is used to optimize continuously-differentiable strictly convex cost functions over weakly-connected undirected networks, while it is anytime feasible and models various nonlinearities to account for imperfections and constraints on the (physical model of) agents in terms of limited actuation capabilities, e.g., quantization and saturation. Due to such inherent nonlinearities, the existing linear solutions considering ideal agent models may not necessarily converge with guaranteed optimality and anytime feasibility. Some applications also impose specific nonlinearities, e.g., convergence in fixed/finite-time or sign-based robust disturbance-tolerant dynamics. Our proposed distributed protocol generalizes such nonlinear models. Putting convex set analysis together with nonsmooth Lyapunov analysis, we prove convergence, (i) regardless of the particular type of nonlinearity, and (ii) with weak network-connectivity requirements (uniform-connectivity).

Index Terms—Network resource allocation, graph theory, spanning tree, convex optimization.

I. INTRODUCTION

CONSSENSUS has been infiltrated into control and machine learning, e.g., in distributed optimization [1], estimation [2, 3], and resource allocation [4]. Network resource allocation is the problem of allocating constant amount of resources among agents to minimize the cost, with application to several fields, such as, the distributed Economic Dispatch Problem (EDP) [5–12], distributed coverage control [13], congestion control [14], and distributed load balancing [15]. Such problems are subject to inherent physical constraints on the agents, leading to nonlinear dynamics with respect to actuation and affecting the stability. This work formulates a general solution considering such nonlinear agents to solve distributed allocation. Another example is the Automatic Generation Control (AGC) in electric power systems [17, 18], which regulates the generators’ output power compensating for any generation-load mismatch in the system. The AGC generators’ deviations are subject to limits based on the available power reserves and also on Ramp Rate Limits (RRLs) (or rate saturation), i.e., the speed their produced power can increase or decrease is constrained. Under such nonlinear constraints, a linear (ideal) model for generators as given by [4, 5] may not remain feasible or result in a sub-optimal solution.

Related literature: The literature spans from preliminary linear [4, 5, 19] and accelerated linear [20] solutions to more recent sign-based consensus [21], Newton-based [22], derivative-free swarm-based [23], predictive online saddle-point methods [24], 2nd-order autonomous dynamics [6], [25–27], distributed mechanism over local message-passing networks [28], multi-objective [29], primal-dual [30–32], Lagrangian-based [11, 33–35], and projected proximal sub-gradient algorithms [36], among others. These works cannot address different inherent physical nonlinearities on the agents’ model, such as the RRL for distributed AGC, or some other designed nonlinear models intended for improving computation load and convergence rate, e.g., reaching fast convergence. In general, model nonlinearities such as limited computational capacities, constrained actuation, and model imperfections may significantly affect the convergence or degrade the resource allocation performance. For example, none of the mentioned references can address quantization, saturation, and sign-based actuation altogether, while ensuring feasibility at all times. In reality, under such model nonlinearities there is no guarantee that the existing solutions accurately follow the ideally-designed dynamics and preserve feasibility, optimality, or specified convergence rate. Some existing Lagrangian-based methods [34, 35] are not anytime feasible, but reach feasibility upon the convergence [5]. In a different line of research, inequality-constrained problems are solved via primal-dual methods and Lagrangian relaxation [30–32]. This differs from equality-constrained problems which are typically solved via Laplacian gradient methods. The latter is is used for the optimal resource allocation in EDP [5, 7], but without addressing the RRL nonlinearity on the power rate.

Main contributions: We propose a general 1st-order Laplacian-gradient dynamics for distributed resource allocation. The proposed localized solution generalizes many nonlinear constraints on the agents including, but not limited to, (i) saturation and (ii) quantization. Further, some specific constraints (e.g., on the convergence or robustness), impose nonlinearities on the agents’ dynamics. For example, it is practical in applications to design (iii) fixed-time and finite-time convergent solutions, and/or (iv) robust protocols to impulsive noise and uncertainties. Our proposed dynamics generalizes many symmetric sign-preserving model nonlinearities. We prove uniqueness, anytime feasibility, and convergence over generally sparse, time-varying, undirected (and not necessarily connected) networks, referred to as uniform-connectivity. The proofs are based on nonsmooth Lyapunov theory [37], graph theory, and convex analysis, irrespective of the type of nonlinearity. This generalized 1st-order solution is more practical as it considers all possible sign-preserving physical constraints on...
the agents dynamics, and further, can be extended to consider nonlinearities on the agents’ communications [38, 39].

II. PROBLEM STATEMENT

The network resource allocation problem is in the form

$$\min_{\mathbf{X}} F(\mathbf{X}, t) = \sum_{i=1}^{n} f_i(\mathbf{x}_i, t), \quad \text{s.t.} \quad \mathbf{Xa} = \mathbf{b} \quad (1)$$

with $\mathbf{x}_i \in \mathbb{R}^d$, $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$, vectors $\mathbf{a} = [a_1; \ldots; a_n] \in \mathbb{R}^n$, and $\mathbf{b} = [b_1; \ldots; b_d] \in \mathbb{R}^d$. The entries of $\mathbf{a}$ are assumed to not be very close to zero to avoid unbounded solutions. If $a_j = 0$ for agent $j$, its state $\mathbf{x}_j$ is decoupled from the other agents, and problem (1) can be restated for $n-1$ coupled agents plus an unconstrained optimization on $f_j(\mathbf{x}_j, t)$. $f_i(\mathbf{x}_i, t) : \mathbb{R}^{d+1} \to \mathbb{R}$ in (1) denotes the local time-varying cost at agent $i$ as $f_i(\mathbf{x}_i, t) = f_i(x_i) + f_i(t)$, with $f_i(t) \neq 0$ representing the time-varying cost. In some applications, the states are subject to the box constraints, $\mathbf{m} \leq \mathbf{x} \leq \mathbf{M}$, denoting element-wise comparison. Using exact penalty functions, these constraints are added into the local objectives as $f_i^p(\mathbf{x}_i, t) = f_i(\mathbf{x}_i, t) + \epsilon^h(\mathbf{x}_i - \mathbf{m}) + \epsilon^h(\mathbf{m} - \mathbf{x}_i)$ with $h^*(u) = \max\{u, 0\}$. The smooth equivalent substitutes are $\frac{1}{2} \log(1 + e^{-\epsilon^h(u)})$ quadratic penalty $(\max\{u, 0\})^2$ (or $\theta$-logarithmic barrier [11]) with the gap inversely scaling with $\epsilon$.

Assumption 1: The (time-independent part of) local functions, $f_i(\mathbf{x}_i) : \mathbb{R}^d \to \mathbb{R}$, are strictly convex and differentiable. This assumption ensures unique optimizer (see Lemma 3) and existence of function gradient. This paper aims to design a localized general nonlinear dynamic to solve (1) based on partial information at agents over a network.

III. DEFINITIONS AND AUXILIARY RESULTS

A. Graph Theory and Nonsmooth Analysis

The multi-agent network is modeled as a time-varying undirected graph $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$ with links $\mathcal{E}(t)$ and nodes $\mathcal{V} = \{1, \ldots, n\}$. $(i, j) \in \mathcal{E}(t)$ denotes a link from agent $i$ to $j$, and the set $\mathcal{N}_i(t) = \{j \mid (i, j) \in \mathcal{E}(t)\}$ represents the direct neighbors of agent $i$ over $\mathcal{G}(t)$. Every link $(i, j) \in \mathcal{E}(t)$ is assigned with a positive weight $W_{ij} > 0$, in the associated weight matrix $W(t) = [W_{ij}(t)] \in \mathbb{R}_{\geq 0}^{n \times n}$ of $\mathcal{G}(t)$. In $\mathcal{G}(t)$ define a spanning tree as a subset of links in which there exists only one path between every two nodes (for all $n$ nodes).

Assumption 2: The following assumptions hold on $\mathcal{G}(t)$:
- The network $\mathcal{G}(t)$ is undirected. This implies a symmetric associated weight matrix $W(t)$, i.e., $W_{ij}(t) = W_{ji}(t) \geq 0$ for $i, j \in \{1, \ldots, n\}$ at all time $t \geq 0$, which is not necessarily row, column, or doubly stochastic.
- There exist a sequence of non-overlapping finite time-intervals $[t_k, t_k + k_b]$ in which $\bigcup_{i=1}^{t_k + k_b} \mathcal{G}(t)$ includes an undirected spanning tree (uniform-connectivity).

Next, we restate some nonsmooth set-valuation analysis from [37]. For a nonsmooth function $h : \mathbb{R}^m \to \mathbb{R}$, define its generalized gradient as

$$\partial h(\mathbf{x}) = \text{co}\{\lim \nabla h(\mathbf{x}_i) : \mathbf{x}_i \to \mathbf{x}, \mathbf{x}_i \notin \Omega_h \cup S\} \quad (2)$$

1Note the subtle abuse of notation where the overall state $\mathbf{X}$ is represented in matrix form to simplify the notation in proof analysis throughout the paper.
Proof: From strict convexity of $\tilde{F}(X)$ (Assumption 1), only one of its strict convex level sets, say $L_1(\tilde{F})$, touches the constraint facet $S_b$ only at a single point, say $X^*$. Clearly, the gradient $\nabla \tilde{F}(X^*)$ is orthogonal to $S_b$, and $\nabla \tilde{F}(x_i) = \frac{\nabla f_i(x_i)}{a_i} = \Lambda$ for all $i$. By contradiction, consider two points $X_1, X_2 \in S_b$ for which $\nabla \tilde{F}(X_1^*) = \Lambda_1 \otimes a_i^T$, $\nabla \tilde{F}(X_2^*) = \Lambda_2 \otimes a_i^T$ (two possible optimum), implying that either (i) one level set $L_1(\tilde{F})$, $\gamma = \tilde{F}(X^*) = \tilde{F}(X_1^*)$ is adjacent to the affine constraint $S_a$ at both $X_1^*$, $X_2^*$, or (ii) there are two level sets $L_1(\tilde{F}(X), L_2(\tilde{F}(X))$, touching the affine set $S_a$ at $X_1^*$ and $X_2^*$ respectively, and thus, at both points $\nabla \tilde{F}(X_1^*)$ and $\nabla \tilde{F}(X_2^*)$ need to be orthogonal to $(X_1^* - X_2^*)$ in $S_b$. Since $S_b$ forms a linear facet, the former case contradicts the strict convexity of the level sets. In the latter case,

$$e_p^T(\nabla \tilde{F}(X_1^*) - \nabla \tilde{F}(X_2^*)) (X_1^* - X_2^*) p = 0, \forall p$$

which contradicts (4). This proves the lemma.

This proof analysis is further recalled in the next sections.

IV. THE PROPOSED 1ST-ORDER NONLINEAR DYNAMICS

We propose a 1st-order protocol $F: \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ coupling the agents’ dynamics to solve problem (1), while addressing model nonlinearities and satisfying feasibility at all times,

$$\dot{x}_i = -\frac{1}{a_i} \sum_{j \in N_i} W_{ij} g \left( \frac{\nabla \tilde{f}_j(x_i)}{a_i} - \frac{\nabla \tilde{f}_j(x_j)}{a_j} \right) : \mathcal{F}_i(x_j),$$

(6)

with $W_{ij}$ as the weight of the link between agents $i$ and $j$ and $\nabla \tilde{f}_j(x_j)$ as the gradient of (time-invariant part of) the local objective $f_i$ with respect to $x_i$ and $g$ defines the nonlinearity to be explained later. Following Assumption 1, given a state point $X_0$, the level set $L_1(\tilde{F}(X_0))$ is closed, convex, and compact. Then, the solution set $L_1(\tilde{F}(x_i)) \cap S_a$ under (4) is closed and bounded. Indeed (6) represents a differential inclusion due to discontinuity of RHS of (4) [37], where for the sake of notation simplicity “=” is used instead of “∈”. From [37], it is straightforward to see that the trajectory $\mathcal{F}$ is locally bounded, upper semi-continuous, with non-empty, compact, and convex values, and thus, from [37] Proposition S2 and similar to [41], [42], the solution under (6) for initial condition $X_0 \in S_b$ exists and is unique. Recall that the time-varying and time-invariant parts of the local objectives are decoupled. Dynamics (6) represents a 1st-order weighted gradient tracking, with no use of the Hessian matrix. Thus, function $\tilde{f}_i(\cdot)$ is not needed to be twice-differentiable (in contrast to 2nd-order dynamics, e.g., in [25]). This allows to incorporate smooth penalty functions to address the box constraints. In case of communication network among agents, periodic communication with sufficiently small period $T$ is considered, see [43] for details. The state of every agent $i$ evolves under influence of its direct neighbors $j \in N_i$, weighted by $W_{ij}$, e.g., via information sharing networks [43] where every agent $i$ shares its local gradients $\nabla \tilde{f}_i(x_i)$ along with the weight $W_{ij}$. Therefore, the proposed resource allocation dynamics (6) is only based on local information-update, and is distributed over the multi-agent network.

Assumption 3: (Strongly sign-preserving nonlinearity) In dynamics (6), $g: \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a nonlinear odd mapping such that $g(x) = -g(-x)$, $g(x) \geq 0$ for $x \geq 0$, $g(0) = 0$, and $g(x) \geq 0$ for $x < 0$. Further, $\nabla g(0) \neq 0$.

Some causes of such practical nonlinearities as function $g(\cdot)$ in (6), e.g., physics-based nonlinearities, are given next.

**Application 1:** Function $g(\cdot)$ can be adopted from finite-time and fixed-time literature [38], [39], [42] as $\text{sgn}^u(x) = \frac{x}{|x|^\mu - 1}$, where $|| \cdot ||$ denotes the Euclidean norm and $\mu \geq 1$.

In general, system dynamics as $\dot{x}_i = -\sum_{j = 1}^n W_{ij} (\text{sgn}^u(x_i - x_j) + \text{sgn}^{u^2}(x_i - x_j))$ converge in finite/fix-time [42], motivating fast-convergent allocation dynamics [39] as,

$$\dot{x}_i = -\sum_{j \in N_i} W_{ij} (\text{sgn}^u(z) + \text{sgn}^{u^2}(z)),$$

(7)

with $z = \frac{\nabla \tilde{f}_i(x_i) - \nabla \tilde{f}_j(x_j)}{a_i - a_j}$, $0 < \mu_1 < 1$, and $0 < \mu < 1$ (finite-time case) or $1 < \mu_2$ (finite-time case).

**Application 2:** Quantized allocation by choosing $g(\cdot)$ as,

$$g(z) = \text{sgn}(z) \exp(g_a(\log(|z|))),$$

(8)

where $g_a(z) = \delta \left( \frac{z}{\bar{z}} \right)$ represents the uniform quantizer with $\cdot$ as rounding operation to the nearest integer [44]–[46], $\text{sgn}(\cdot)$ follows $\text{sgn}^u(\cdot)$ with $\mu = 0$, $\delta$ is the quantization level, and function $g_a$ denotes logarithmic quantizer.

**Application 3:** Sign-preserving nonlinear dynamics [47], [48] robust to impulsive noise can be achieved via $g_p(z) = \frac{1}{g_a(\log(z))}$ with $p$ as the noise density. For example, for $p$ following approximately uniform $P_1$, or Laplace class $P_2$ [48],

$$p \in P_1: g_p(z) = \begin{cases} \frac{1}{\epsilon p} & \text{sgn}(z) \ |z| > d \\ 0 & \text{otherwise} \end{cases}$$

(9)

$$p \in P_2: g_p(z) = 2c \text{sgn}(z),$$

(10)

with $0 < \epsilon < 1$, $d > 0$.

**Application 4:** Saturation nonlinearities [49], [50] (or clipping) are due to limited actuation range for which the saturation level may affect the stability, convergence, and general behavior of the system. For a given saturation level $\kappa > 0$,

$$g_\kappa(z) = \begin{cases} \kappa \text{sgn}(z) & |z| \geq \kappa \\ z & |z| \leq \kappa \end{cases}$$

(11)

**Remark 1:** Recall that Eq. (6) represents Laplacian-gradient-type dynamics (see [4] for details) which can ensure feasibility at all times under various nonlinearities of $g(\cdot)$ in contrast to the Lagrangian-type methods [11], [30]–[32], [34], [35]. If the actuator is not subject to nonlinearities, one may select a linear function for $g(\cdot)$, i.e., $g(z) = z$ and utilize linear methods [1], [3], [20]. However, our focus is to provide a more general solution method that is applicable also to agents with nonlinearities (inherent or by design). For example, the generators are known to be physically constrained with RRLs which is a determining factor on the stability of the grid [18]. Linear methods cannot consider RRLs and may result in solutions with a high rate of change in power generation $\xi_i$, which cannot be followed in reality and may result in infeasibility or sub-optimality. However, such limits can be satisfied considering $g(\cdot)$ as in (11) where the limits can be tuned by $\kappa$. 


V. Analysis of Convergence

In this section, combining convex analysis from Lemma 12 with Lyapunov theory, we prove the convergence of the general protocol (6) to the optimal value of problem (1) subject to the constraint on the weighted-sum of resources. The proof is, in general, irrespective of the nonlinearity types, i.e., holds for any nonlinearity satisfying Assumption 3 including [7], [11].

Lemma 3: (Anytime Feasibility) Suppose Assumption 3 holds. The states of the agents under dynamics (6) remain feasible, i.e., if \( x_0 \in S_0 \), then \( x(t) \in S_0 \) for \( t > 0 \).

Proof: Having \( x_0 \in S_0 \) implies that \( x_0 \in b = b \). For the general state dynamics (6),

\[
\frac{d}{dt} (\mathbf{x}_a) = \sum_{i=1}^{n} \mathbf{x}_a i a_i = -\sum_{i=1}^{n} \sum_{j \in N_i} W_{ij} g(\frac{\nabla f_i(x_i)}{a_i} - \frac{\nabla f_j(x_j)}{a_j}).
\]

(12)

From Assumptions 2 and 3, \( W_{ij} = W_{ji} = g(-x) = -g(x) \). Therefore, the summation in (12) is equal to zero, \( \frac{d}{dt} (\mathbf{x}_a) = 0 \), and \( \mathbf{x}_a \) is time-invariant under dynamics (6). Thus, having feasible initial state \( x_0 \in b = b \), then \( x(t) / a = b \) remains feasible over time, i.e., \( x(t) \in S_0 \) for all \( t > 0 \).

The above proves anytime feasibility, i.e., nonlinear dynamics (6) remains feasible at all times, which is privileged over consensus-based solutions [11], [34], [35]. For AGC subject to RRL, \( x \), and thus \( g(\cdot) \) needs to be further of limited range. Further, Lemma 1 shows that \( S_0 \) is positively invariant under the nonlinear dynamics (6).

Theorem 1: (Equilibrium-Uniqueness) Under Assumptions 2 and 3, the equilibrium point \( X^* \) of the solution dynamics (6) is only in the form \( \nabla \mathcal{F}(X^*) = \Lambda \otimes a^T \) with \( \Lambda \in \mathbb{R}^{d} \), and coincides with the unique optimal point of (1).

Proof: From dynamics (6), \( x^*_a = 0, \forall i \) for \( X^* \) satisfying \( \nabla \mathcal{F}(X^*) = \Lambda \otimes a^T \), and such point \( X^* \) is clearly an equilibrium of (6). We prove that there is no other equilibrium with \( \nabla \mathcal{F}(X^*) \neq \Lambda \otimes a^T \) by contradiction. Assume \( X \) as the equilibrium of (6) such that \( \frac{\nabla f_i(x_i)}{a_i} \neq \frac{\nabla f_j(x_j)}{a_j} \) for at least two agents \( i, j \). Let \( \nabla \mathcal{F}(\hat{X}) = (\hat{a}_1, \ldots, \hat{a}_d) \). Consider two agents \( \alpha = \arg \max_{x_i \in \{1, \ldots, n\}} \hat{a}_i \) and \( \beta = \arg \max_{x_j \in \{1, \ldots, n\}} \hat{a}_j \) for any entry \( x \in \{1, \ldots, d\} \). Following the Assumption 2 the existence of an (unordered) spanning tree in the union network \( \bigcup_{t \leq t_k} G(t) \) implies that there is a mutual path between nodes (agents) \( \alpha \) and \( \beta \). In this path, there exists at least two agents \( \bar{a} \) and \( \bar{b} \) for which \( \Lambda_{a \bar{a}} \geq \Lambda_{a \bar{b}} \), \( \Lambda_{a \bar{b}} = \Lambda_{b \bar{a}} \) with \( \Lambda_{a \bar{a}} \) and \( \Lambda_{b \bar{a}} \) as the neighbors of \( \bar{a} \) and \( \bar{b} \), respectively. The strict inequality holds for at least one neighboring node in \( \Lambda_{a \bar{a}} \) and \( \Lambda_{b \bar{a}} \). From Assumption 2 and 3 in a sub-domain of \( \{l_k, t_k + l_k\} \), we have \( \hat{\mathbf{x}}_a < 0 \) and \( \hat{\mathbf{x}}_b < 0 \). Therefore, \( \hat{X} \neq 0 \) which contradicts the assumption that \( \hat{X} \) is the equilibrium of (6).

Recall that, from Lemma 2, this point coincides with the optimal solution of (1). As for every feasible initialization in \( S_0 \) there is only one such point \( X^* \) satisfying \( \nabla \mathcal{F}(X^*) = \Lambda \otimes a^T \). This completes the proof.

The above lemma paves the way for convergence analysis via the Lyapunov stability theorem, as it shows that the dynamics (6) has a unique equilibrium for any feasible initial condition.

Lemma 4: [5] Let nonlinearity \( g(\cdot) \) and matrix \( W \) satisfy Assumptions 2 and 3. Then, for \( \psi \in \mathbb{R}^d \) we have,

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} g(\psi_j - \psi_i) = \frac{1}{2} \sum_{i=1}^{n} W_{ij} (\psi_j - \psi_i)^T g(\psi_j - \psi_i).
\]

Following the convex analysis in Lemmas 14 and Theorem 1 along with Lemma 4 we provide our main theorem next.

Theorem 2: (Convergence) Suppose Assumptions 1, 3 hold. Then, initializing by \( x_0 \in S_0 \), the proposed dynamics (6) solves the network resource allocation problem (1).

Proof: Following Lemmas 2 and 3 and Theorem 1 and initializing from \( x_0 \in S_0 \), for any \( b \in \mathbb{R}^d \), there is a unique feasible equilibrium \( X^* \) for solution dynamics (6). Let \( F(X) = F(X, t) - F(X^*, t) \). Clearly, \( F(X) = \sum_{i=1}^{n} (\hat{f}_i(x_i) - \hat{f}_i(x_i^*)) > 0 \) is purely a function of \( X \), with \( X^* \) as its unique equilibrium. For this continuous (but nonsmooth) regular and locally Lipschitz Lyapunov function \( F(X) \), its generalized derivative \( t \to F(x(t)) \), for \( x \) as the solution to (4), satisfies \( \partial_t F(x(t)) \in L^2 F(X(t)), \) see [5] Proposition 10. Then (dropping \( t \) for notation simplicity),

\[
\partial_t F = \nabla F^T X = \sum_{i=1}^{n} \mathbf{X}_a i = \frac{\nabla f_i(x_i)}{a_i} \sum_{j \in N_i} W_{ij} g(\frac{\nabla f_j(x_j)}{a_i} - \frac{\nabla f_j(x_j)}{a_j}).
\]

(13)

Following Lemma 4

\[
\partial_t F = -\sum_{i=1}^{n} W_{ij} \left( \frac{\nabla f_i(x_i)}{a_i} - \frac{\nabla f_j(x_j)}{a_j} \right)^T g(\frac{\nabla f_i(x_i)}{a_i} - \frac{\nabla f_j(x_j)}{a_j}).
\]

From Assumption 3 \( g(x) \) is odd and strongly sign-preserving, i.e., \( x^* g(x) \geq 0 \). Therefore, \( \partial_t F \leq 0 \) with the largest invariant set \( I \) contained in \( \{ X \in L \mathcal{F}(X) \cap \mathbb{S} \} \), \( I \) includes the unique point \( X^* \in S_0 \) for which \( \nabla \mathcal{F} \in \) span \( \{ a \} \) (or \( \frac{\nabla f_i(x_i)}{a_i} \otimes \frac{\nabla f_j(x_j)}{a_j} = \varphi^*, \forall i, j \) from Lemmas 1 and 2). Using LaSalle invariance principle for differential inclusions [5] Theorem 2.1, initializing by \( x_0 \in S_0 \), the trajectory set \( \{ L \mathcal{F}(X) \cap \mathbb{S} \} \) remains feasible and positively invariant under (6) (Lemma 3), and converges to the largest invariant set \( I \) \{ \( X^* \) \} including the unique equilibrium of (6) (as shown in Theorem 1). \( F \) is monotonically non-decreasing and radially unbounded, \( \max L \mathcal{F}(X(t)) < \infty \) for all \( X \in S_0 \), and thus, from [5] Theorem 1, \( X^* \) is globally strongly asymptotically stable. This proves that agents’ states under dynamics (6) converge to \( X^* \).

The above proof holds for any \( b \) value and any initialization state \( x_0 \in S_0 \), and the solution converges to \( X^* \) in Lemma 1.

Remark 2: Following similar analysis as in [5], assuming \( \exists u_{\text{min}} K_{\text{min}} \) such that \( u_{\text{min}} \leq \nabla f_i(x_i) \) (strongly convex cost with smooth gradient) and \( K_{\text{min}} \leq \frac{g(\mathbf{x})}{\mathbf{x}} \). Eq. (13) over a connected network \( G \) with \( \lambda_2 \) as its algebraic connectivity (Fiedler-value) and \( a = 1_n \) gives the decay rate of \( F \) as,

\[
\partial_t F \leq -2u_{\text{min}} K_{\text{min}} \lambda_2 F
\]

(14)

For a disconnected network with at least one link \( (i, j) \), the summation in (13) is positive and \( \partial_t F \) is negative if \( \frac{\nabla f_i(x_i)}{a_i} \neq \frac{\nabla f_j(x_j)}{a_j} \). From Assumption 2 \( \partial_t F \) is negative.
over sub-intervals of every time-interval \([t_k, t_k + l_k]\) (infinitely often) having \(\nabla \bar{f}_i(x_i) \neq \nabla \bar{f}_j(x_j)\) for (at least) 2 neighbors \(i, j\) till reaching the optimizer \(X^*\) (for which \(\nabla \bar{f}_i(x_i^*) = \nabla \bar{f}_j(x_j^*) \forall i, j\)). One may also consider discrete Lyapunov analysis and simply prove that \(\mathcal{F}(X(t_k + l_k)) < \mathcal{F}(X(t_k))\) for all \(X(t_k) \in \mathcal{S}_b \setminus \mathcal{I}\).

VI. SIMULATION OVER SPARSE NETWORKS

We simulate protocol (6) for (i) quantized and (ii) saturated resource allocation over 4 weakly-connected Erdos-Rényi networks of \(n = 100\) agents changing every 0.1 second with switching command \(s : [10t - 4, 2.5t]\) satisfying Assumption 2. Consider strictly convex cost as [19],

\[
\begin{align*}
\bar{f}_i(x_i) &= \sum_{j=1}^{4} \tilde{a}_{i,j}(x_{i,j} - \tilde{c}_{i,j})^2 + \log(1 + \exp(b_{i,j}(x_{i,j} - \tilde{d}_{i,j}))) \\
\bar{f}_i(t) &= \sum_{j=1}^{4} \tilde{a}_{i,j} \sin(\alpha_{i,j}t + \phi_{i,j})
\end{align*}
\]

with random parameters. Assume \(b = 101.4\) and \(\alpha_i\) in \([0.1, 1]\). To solve (14), we accommodate (6) for two cases: (i) quantized actuation via the logarithmic quantizer (8) with \(\delta = 1\), and (ii) saturated actuation (11) with \(\kappa = 1\). The time-evolution of the cost (15) and the Lyapunov \(\mathcal{F}(X) = F(X, t) - F^*(t)\) are shown in Fig. 1. As it is clear, the cost functions converge to the optimal (time-varying) values, with Lyapunov functions (residuals) decreasing in time.

VII. APPLICATION: AUTOMATIC GENERATION CONTROL

The AGC adjusts the power generation based on predetermined reserve limits to compensate for any generation-load mismatch in a time scale of minutes. We assume that the generation-load mismatch is known (e.g. generator outage) and we aim to allocate that mismatch to the generators by minimizing their power deviation cost. Let \(x_i\) represent the power deviation for generator \(i\). The optimization problem finds the optimal mismatch allocation to \(n\) generators while satisfying the reserve limits and is given by:

\[
\min_{X} \sum_{i=1}^{n} \gamma_i x_i^2 + \beta_i x_i + \alpha_i,
\]

\[
s.t. \sum_{i=1}^{n} x_i = P_{\text{mis}}, \quad -R_i \leq x_i \leq R_i, \quad i = 1, ..., n.
\]

The generation-load mismatch is \(P_{\text{mis}}\) and the reserve limits for decreasing and increasing the power generation are \(R_i\). Mapping the problem to formulation (1)

\[
\begin{align*}
\bar{f}_i(x_i) &= \sum_{j=1}^{4} \tilde{a}_{i,j}(x_{i,j} - \tilde{c}_{i,j})^2 + \log(1 + \exp(b_{i,j}(x_{i,j} - \tilde{d}_{i,j}))) \\
\bar{f}_i(t) &= \sum_{j=1}^{4} \tilde{a}_{i,j} \sin(\alpha_{i,j}t + \phi_{i,j})
\end{align*}
\]

with random parameters. Assume \(b = 101.4\) and \(\alpha_i\) in \([0.1, 1]\). To solve (14), we accommodate (6) for two cases: (i) quantized actuation via the logarithmic quantizer (8) with \(\delta = 1\), and (ii) saturated actuation (11) with \(\kappa = 1\). The time-evolution of the cost (15) and the Lyapunov \(\mathcal{F}(X) = F(X, t) - F^*(t)\) are shown in Fig. 1. As it is clear, the cost functions converge to the optimal (time-varying) values, with Lyapunov functions (residuals) decreasing in time.

VII. APPLICATION: AUTOMATIC GENERATION CONTROL

The AGC adjusts the power generation based on predetermined reserve limits to compensate for any generation-load mismatch in a time scale of minutes. We assume that the generation-load mismatch is known (e.g. generator outage) and we aim to allocate that mismatch to the generators by minimizing their power deviation cost. Let \(x_i\) represent the power deviation for generator \(i\). The optimization problem finds the optimal mismatch allocation to \(n\) generators while satisfying the reserve limits and is given by:

\[
\min_{X} \sum_{i=1}^{n} \gamma_i x_i^2 + \beta_i x_i + \alpha_i,
\]

\[
s.t. \sum_{i=1}^{n} x_i = P_{\text{mis}}, \quad -R_i \leq x_i \leq R_i, \quad i = 1, ..., n.
\]
