Doostmohammadian, Mohammadreza; Aghasi, Alireza; Pirani, Mohammad; Nekouei, Ehsan; Khan, Usman A.; Charalambous, Themistoklis

**Fast-Convergent Anytime-Feasible Dynamics for Distributed Allocation of Resources over Switching Sparse Networks with Quantized Communication Links**

Published in:
2022 European Control Conference, ECC 2022

**DOI:**
10.23919/ECC55457.2022.9838141

Published: 01/01/2022

*Document Version*
Peer reviewed version

*Please cite the original version:*
Fast-Convergent Anytime-Feasible Dynamics for Distributed Allocation of Resources
Over Switching Sparse Networks with Quantized Communication Links

Mohammadreza Doostmohammadian, Alireza Aghasi, Mohammad Pirani, Ehsan Nekouei, Usman A. Khan, Senior Member, IEEE, Themistoklis Charalambous, Senior Member, IEEE

Abstract—This paper proposes anytime feasible networked dynamics to solve resource allocation problems over time-varying multi-agent networks. The state of agents represents the assigned resources while their total (equal to demand) is constant. The idea is to optimally allocate the resources among the group of agents by minimizing the overall cost subject to fixed sum of resources. Each agent’s information is local and restricted to its own state, cost function, and the ones from its immediate neighbors. This work provides a fast convergent solution (compared to linear dynamics) while considering more-relaxed uniform network connectivity and (logarithmic) quantized communications among agents. The proposed dynamics reaches optimal solution over switching (sparsely-connected) undirected networks as far as their union over some bounded non-overlapping time-intervals has a spanning tree. Moreover, we prove anytime-feasibility of the solution, uniqueness, and convergence to the optimal value irrespective of the specific nonlinearity in the proposed dynamics. Such general proof analysis applies to many similar 1st-order allocation dynamics subject to strongly sign-preserving nonlinearities, e.g., actuator saturation in generator coordination. Further, anytime feasibility (despite the nonlinearities) ensures that our solution satisfies the fixed-sum resources constraint at all times.

Index Terms—Distributed optimization, resource allocation, consensus, logarithmic quantization, spanning tree

I. INTRODUCTION

Distributed optimization in machine learning, signal processing, and control literature, solves the following optimization of a global cost/objective as the sum of local functions:

\[
\min_{\mathbf{x}} F(\mathbf{x}) = \sum_{i=1}^{n} f_i(\mathbf{x}) \quad \text{subject to } C(\mathbf{x}) = 0 \quad (1)
\]

The centralized solution of (1) works under the premise that all information is available and processed at a central computing node. However, in large-scale, every node/agent only has access to local information in its neighborhood, and distributed multi-agent algorithms are needed to cooperatively perform local computations via local information. Such solutions find applications in multi-sensor target tracking [1], edge-computing and load balancing [2], power allocation in cellular networks [3], and distributed support-vector-machine [4], [5]. Different constraints and solutions are considered for problem (1): unconstrained [6]–[8] (for strongly-convex objectives), inequality constraint [9], and consensus constraint [4], [5], [10] aiming also to drive the agents to reach consensus along with optimization. In distributed resource allocation [11]–[19] (also known as network resource allocation) the optimization is constrained with constant summation of states, aiming to allocate a fixed amount of overall resources over a large-scale network. This finds applications in economic dispatch over power networks [13], [14], [20]–[24], networked coverage control [25], [26], and edge-computation offloading [27]. Implementing parallel dynamics at agents via local information requires distributed algorithms, some of which include: preliminary linear solution [11], quantized solution via event-triggered communications (fixed network) [28], accelerated linear solution via adding a momentum term (heavy-ball method) [16], low communication rate protocol converging in quadratic time (via long-term connectivity requirement) [12], game-theoretic approach [19], [29], initialization-free [17], and Lagrangian-based solution [13]–[15]. For many existing solutions, as discussed in [22], there is no guarantee for anytime feasibility, i.e., the summation constraint is not necessarily feasible at all times but only at the final equilibrium state. Particularly, the presence of nonlinearities (e.g., saturation, quantization, or sign-based protocols) may affect solution uniqueness, all-time feasibility, and optimality.

Contributions: This work proposes a nonlinear dynamics for network resource allocation. The main purposes for considering the nonlinearity are: fast convergence [30], [31], quantization [32]–[35], and saturation/clipping [33], among others. Knowing that fixed-time dynamics reaches faster convergence than linear solutions [4], [31], we propose a continuous-time state-update for distributed resource allocation with fast convergence (as compared to linear solution) while considering logarithmic-quantized information exchange among agents. We consider uniform connectivity [7], [8] (instead of all-time connectivity), which only requires the union network over some bounded non-overlapping time-intervals to include a spanning-tree. In contrast to unconstrained or consensus-constrained problems [6]–[8], this work extends the solution to distributed resource allocation over sparsely-connected dynamic networks with quantized communications. Borrowing ideas from convex optimization and level-set methods [36],
[37], we prove uniqueness, feasibility, and optimal convergence under the proposed continuous-time dynamics. The convergence to this optimal value is proved via Lyapunov-type stability analysis. Our convergence analysis, although given for a specific dynamics, can be easily extended to strongly sign-preserving nonlinear 1st-order solutions. Our main contributions are: (i) fast-convergence (compared to linear solutions) while considering quantized data transmissions, (ii) proving anytime feasibility (e.g., versus Lagrangian-based solutions), and (iii) proving convergence for general strongly sign-preserving nonlinearities (e.g., quantization and saturation) over general uniformly-connected dynamic networks.

Paper organization: Section II states the problem and preliminaries. Section III describes the proposed dynamics, while its convergence is proved in Section IV. Section V provides simulations and Section VI concludes the paper.

II. PROBLEM FORMULATION

A. Problem Statement

Distributed resource allocation problem is formulated as,

\[
\min_x F(x) = \sum_{i=1}^{n} f_i(x_i) \quad \text{s.t.} \quad \sum_{i=1}^{n} x_i = K \tag{2}
\]

where \( x_i \in \mathbb{R} \) is the amount of resource allocated to agent \( i \), \( f_i : \mathbb{R} \to \mathbb{R} \) is a convex function known by agent \( i \) representing the cost as a function of resources \( x_i \). The network resource allocation problem \( \text{(2)} \) aims to allocate a fixed quantity of total resources, \( \sum_{i=1}^{n} x_i = K \), among a group of agents communicating over an undirected graph \( \mathcal{G} \), such that the total cost \( F(x) \) is minimized. In \( \text{(2)} \) the states should exactly meet the constraint \( K \), e.g., in economic dispatch problem where the produced power is equal to the demand (anytime-feasibility). This differs from the inequality-constrained power bid cost minimization problem in [9] solved by primal-dual methods. There might be box constraints \( l_i \leq x_i \leq u_i \) involved to bound the amount of resources. One can address these by adding exact penalty functions [36] to the objective as \( f_i^\star(\cdot) = f_i(\cdot) + \epsilon [x_i - \mathbb{m}_i]^+ + \epsilon [\mathbb{m}_i - x_i]^+ \), with \( [u]^+ = \max\{u, 0\} \). Recall that the summation of the strictly convex \( f_i(\cdot) \) and convex penalty \( [\cdot]^+ \) is a strictly convex function. Further, the non-smooth \( [u]^+ \) can be replaced by its smooth equivalent \( \frac{\epsilon}{\epsilon^2} \log(1 + \exp(\epsilon u)) \) as in [5] or quadratic penalty \( ([u]^+)^2 \) [38]. Applications include:

(i) Economic dispatch [20]–[22], [24]: to allocate the electricity generation by facilities to minimize the cost while meeting the required load/demand constraints.
(ii) Congestion-control and load-balancing [2], [27]: to modulate traffics and data routing in telecommunication networks to gain fair allocations among the users.
(iii) Coverage control [25], [26]: the objective is to optimally allocate a group of networked robots/agents over a convex area in order to achieve maximum coverage.

Remark 1: Note the difference of (constrained) problem \( \text{(2)} \) with general (unconstrained) distributed optimization (as in \([6]–[8])\). Other than the constraint, for general distributed optimization the cost at all agents is the same function of \( x \), i.e., \( F(x) = \sum_{i=1}^{n} f_i(x_i) \), while in \( \text{(2)} \) the cost at agent \( i \) is only a function of \( x_i \), i.e., \( F(x) = \sum_{i=1}^{n} f_i(x_i) \). This implies that in \( \text{(2)} \) agent \( i \) only needs to know its own state \( x_i \) and not the other agents’ states \( x_j, j \neq i \).

Table I compares solutions and different constraints in the literature. Recall that the 1st-order dynamics refers to the consensus-type protocols in the form, \( x_i = \sum_{j \in \mathcal{N}_i} f(x_j - x_i) \), while 2nd-order dynamics are in the form, \( x_i = g(y_i), y_i = \sum_{j \in \mathcal{N}_i} h(y_j - y_i) \).

Remark 2: For synchronization and consensus, the 1st-order dynamics, compared to its similar 2nd-order counterparts, is known to have faster convergence [40, page 32].

B. Preliminary Definitions and Lemmas

The communication network of agents is modeled as a sequence of (possibly) time-varying undirected graphs, denoted by \( \mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t)) \) with \( \mathcal{V} = \{1, \ldots, n\} \). Two agents \( i \) and \( j \) can communicate/exchange messages if and only if \( (i,j),(j,i) \in \mathcal{E}(t) \). Define \( \mathcal{N}_i(t) = \{j | (j,i) \in \mathcal{E}(t)\} \) as neighbors of agent \( i \) at time \( t \) and \( n \) by \( n \) matrix \( W(t) \) as the adjacency weight matrix of \( \mathcal{G}(t) \), where \( W_{ij} > 0 \) if link \( (i,j) \in \mathcal{E}(t) \) and \( W_{ij} = 0 \) if \( (i,j) \notin \mathcal{E}(t) \).

Definition 1: In the undirected graph \( \mathcal{G} = (\mathcal{V}, \mathcal{E}) \), define a spanning tree as an undirected spanning subgraph in which any two vertices in \( \mathcal{V} \) are connected by exactly one path. Such tree contains the minimum possible links in \( \mathcal{E} \).

Assumption 1: There exists a sequence of non-overlapping bounded time-intervals, \( [t_k, t_k + l_k] \), where the union network across each interval \( \cup_{t_k=0}^{t_k+l_k} \mathcal{G}(t) \) has a spanning tree, while \( \mathcal{G}(t) \) might be sparsely-connected.

Note that the above weak connectivity requirement ensures a path from any node \( i \) to any node \( j \) infinitely often. A situation where Assumption 1 does not hold is when the network is in the form of two separate graph components.

Definition 2: ([36]) Function \( h(x) : \mathbb{R}^n \to \mathbb{R} \) is strictly convex if \( h(kx_1 + (1 - k)x_2) < kh(x_1) + (1 - k)h(x_2) \) for all \( x_1, x_2 \in \mathbb{R}^n, 0 < k < 1 \). It is twice differentiable \( h(x) \) if \( \nabla^2 h(x) \) is positive definite everywhere.

Assumption 2: The functions \( f_i(x_i), i = 1, \ldots, n \) in problem \( \text{(2)} \) are strictly convex and differentiable.

Lemma 1 ([11], [16]): Under the Assumption 2 the resource allocation problem \( \text{(2)} \) has a unique optimal solution \( x^* \) for which \( \nabla F(x^*) = \psi^* \mathbf{1}_n \), where \( \nabla F(x^*) = \left( \frac{\partial F(x^*)}{\partial x_1}, \ldots, \frac{\partial F(x^*)}{\partial x_n} \right) \) denotes the gradient of \( F \) at \( x^* \), and \( \mathbf{1}_n \) is the column vector of \( 1 \)’s.
The optimality condition of Lemma 1 is simply the KKT condition, with $\psi^*$ as the optimal Lagrange multiplier and $\mathbb{1}_n$ as the gradient of the constraint [36].

Definition 3: Define the feasible set of states as the affine space $S_K = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = K\}.$

Definition 4: [36, 37] Given $h(x) : \mathbb{R}^n \to \mathbb{R}$, define its level set $L_\gamma(h) = \{x \in \mathbb{R}^n \mid h(x) \leq \gamma \in \mathbb{R}\}$. For strictly convex $h$, $L_\gamma(h)$ is closed, compact, and strictly convex. The following lemma describes the solution of (2).

Lemma 2: Under Assumption [3], there is a unique point $x^*$ such that $\nabla F(x^*) = \psi^* \mathbb{1}_n$, for every feasible set $S_K$. In other words, given a feasible set $S_K$, there is only one such point $x^* \in S_K$ for which $\frac{\partial f}{\partial x_i}(x^*) = \frac{\partial f}{\partial x_j}(x^*)$, $\forall i, j \in \{1, \ldots, n\}$.

Proof: Following strict convexity of $F(x)$ (Assumption [3]), only one level sets, say $L_{\gamma_1}$, is adjacent to the constraint facet $S_K$, with touching only at one point $x^*$ for which $\nabla F(x^*)$ is orthogonal to $S_K$, i.e., $\frac{\partial f}{\partial x_i}(x^*) = \frac{\partial f}{\partial x_j}(x^*)$, $\forall i, j$. By contradiction, assume $x^1, x^2 \in S_K$ such that $\nabla F(x^1) = \psi^* \mathbb{1}_n$ and $\nabla F(x^2) = \psi^* \mathbb{1}_n$, implying that the strongly convex $L_{\gamma_1}$, $\Gamma = F_{x^1} = F_{x^2}$ is tangent to the affine facade $S_K$ at two points $x^1$ and $x^2$, or two level sets $L_{\gamma_1}$ and $L_{\gamma_2}$ are adjacent to $S_K$. Both cases contradict the strict convexity and closedness of the level sets. This proves the lemma by contradiction.

Lemma 3: Let $g_i : \mathbb{R} \to \mathbb{R}, l \in \{1, 2\}$ be an odd mapping, matrix $W \in \mathbb{R}^{n \times n}$ be symmetric, and $\varphi \in \mathbb{R}^n$. Then,

$$\sum_{i=1}^n \varphi_i \sum_{j=1}^n W_{ij} g_2(g_1(\varphi_j) - g_1(\varphi_i)) = -\sum_{i,j=1}^n W_{ij} (\varphi_j - \varphi_i) g_2(g_1(\varphi_j) - g_1(\varphi_i)).$$

Proof: For every $i, j$, $W_{ij} = W_{ji}$, and $g_i(x) = -g_i(-x)$, $l \in \{1, 2\}$. Thus, we have,

$$\varphi_i W_{ij} g_2(g_1(\varphi_j) - g_1(\varphi_i)) + \varphi_j W_{ij} g_2(g_1(\varphi_i) - g_1(\varphi_j)) = W_{ij} (\varphi_i - \varphi_j) g_2(g_1(\varphi_i) - g_1(\varphi_j))$$

and the proof follows.

We borrow results on nonsmooth analysis and set-value notions from [41], and skip the details due to space limitation. Define the generalized gradient $\partial h : \mathbb{R}^n \to \mathcal{B}(\mathbb{R})$ for a nonsmooth function $h : \mathbb{R}^n \to \mathbb{R}$ as,

$$\partial h(x) = \{c_0 \lim_{i \to -\infty} \nabla h(x_i) : x_i \to x, x_i \notin \Omega \cup S\} \quad (5)$$

with $\partial h(x)$ denoting convex hull, $S \subset \mathbb{R}^n$ as a zero Lebesgue measure set, and $\Omega_i \subset \mathbb{R}^n$ as the set of non-differentiable points in the domain of $f$. If $h$ is locally Lipschitz, then $\partial h(x)$ is nonempty, compact, and convex, and $\partial h : \mathbb{R}^n \to \mathcal{B}(\mathbb{R})$, $x \mapsto \partial h(x)$, is upper semi-continuous and locally bounded. Then, the set-valued derivative $L \partial h(x) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ of function $h$ with respect to the dynamics $x \in \mathcal{O}(x)$ is,

$$L \partial h = \{\eta \in \mathbb{R} \mid \exists \nu \in \mathcal{H}(x) \text{ s.t. } \zeta^T \nu = \eta, \forall \zeta \in \partial f(x)\} \quad (6)$$

We use this for nonsmooth Lyapunov analysis in Section IV.

III. The Proposed Solution

Consider linear dynamics to solve (2) over undirected graphs (e.g., Laplacian-gradient model in [22]),

$$\dot{x}_i = -\eta_1 \sum_{j \in \mathcal{N}_i} W_{ij} (\frac{df_i}{dx_i} - \frac{df_j}{dx_j})$$

where $\eta_1 > 0$ and we assume symmetric weight matrix $W$ with $W_{ij} \geq 0$. Note that for switching networks the RHS of (2) is discontinuous and the dynamics indeed represents a differential inclusion $\dot{x} \in \partial h(x)$. Throughout the paper for notation simplicity we use equality instead of “=”. By ideas from finite-time consensus [42, 43], to reach faster convergence for $|\frac{df_i}{dx_i} - \frac{df_j}{dx_j}| < 1$, (7) is modified as,

$$\dot{x}_i = -\eta_1 \sum_{j \in \mathcal{N}_i} W_{ij} \text{sgn}^+ \left(\frac{df_i}{dx_i} - \frac{df_j}{dx_j}\right)$$

where $0 < \nu < 1$ and $\text{sgn}^+(x) = x|x|^{\nu-1}$ which is non-Lipschitz at 0 (for $0 < \nu < 1$). In general, non-Lipschitz dynamics in the form $\dot{x}_i = -\eta_1 \sum_{j \in \mathcal{N}_i} W_{ij} \text{sgn}^+(x_i - x_j)$ is proved to reach finite-time convergence [44] and faster than linear case close to the equilibrium (because $|\text{sgn}^+(x)| > |x|$ for $|x| < 1$). Further, consensus is in finite-time [43, 44], however, with slow rate in regions far from the equilibrium (because $|\text{sgn}^+(x)| < |x|$ for $|x| > 1$). To overcome this, in fixed-time consensus protocols [44], typically a second term is added as $\text{sgn}^+(x)$ with $\nu_2 > 1$. Having $|\text{sgn}^+(x)| > |x|$ for $|x| > 1$ implies faster convergence rate than the linear case for states far from the consensus equilibrium. Therefore, the combination of the two gives faster convergence tunable by parameters $\nu_1, \nu_2$. Therefore, this work proposes nonlinear 1st-order dynamics to solve problem (2) via,

$$\dot{x}_i = -\sum_{j \in \mathcal{N}_i} W_{ij} \left(\eta_1 \text{sgn}^+ \left(\frac{df_i}{dx_i} - \frac{df_j}{dx_j}\right) + \eta_2 \text{sgn}^+ \left(\frac{df_i}{dx_i} - \frac{df_j}{dx_j}\right)\right)$$

with $0 < \nu_1 < 1 < \nu_2$, $0 < \eta_2, \eta_1$, and $W_{ij} = W_{ji} \geq 0$, where $|\text{sgn}^+(x)| + |\text{sgn}^+(x)| > |x|$; therefore, for any $|\frac{df_i}{dx_i} - \frac{df_j}{dx_j}|$, $|\dot{x}|$ is greater under dynamics (9) (compared to the linear case). Unlike [11, 16] we do not assume $W$ to be (doubly) stochastic. Further, the convergence rate can be tuned by $\nu_1, \nu_2$. Next, consider quantized link $f$.\footnote{The proposed model $f$ represents a system dynamics whose RHS changes in discrete-time, therefore, a hybrid state $\zeta = (x, \tau)$ with $\tau$ as the timer state and $p : t \in [0, \tau] \to P = \{1, 2, \ldots, \pi\}$ as the index of the switching network $g_p$. Then, the switching flow map is $f : \bar{p} = 0, \bar{\tau} \in [0, \tau]$, along with the dynamics of $x$ in (10) and domain set $\zeta = (x, p, \bar{\tau} \in \mathbb{C} \times \mathbb{R} \times \mathbb{R} \times [0, 1])$. Then, the jump map is $f : x^p = x$, $x^\pi \in P$, $\bar{\tau} = 0$ over the jump domain set $\zeta \in \mathcal{D} = \mathbb{R}^n \times P \times [1]$, implying that the hybrid system jumps to a new mode $p \in P$ whenever $\zeta \in \mathcal{D}$ with the interval length depending on the timer rate $\bar{\tau}$ for each mode $p$. For example, for the minimum-length switching time-interval $m_\omega$, the rate is $\tau = \frac{1}{m_\omega}$. This simply means that after $m$ (and in general $m_\omega \geq m$) periods $\omega$, the jump occurs as $\tau = 1$ and $p$ switches to a new mode (a new topology $G_p$, $\tau$ starts over, while the state $x$ is continuous and unchanged). This is known as piece-wise constant jump mapping and satisfies the so-called Basic Assumptions for stability, see [5], [45] and references therein.}
\[ x_i = -\sum_{j \in N_i} W_{ij}(t, p) \left( \eta_1 \text{sgn}^v(q \left( \frac{df_i}{dx_i} \right) - q \left( \frac{df_j}{dx_j} \right)) \right. \\
+ \eta_2 \text{sgn}^v(q \left( \frac{df_i}{dx_i} \right) - q \left( \frac{df_j}{dx_j} \right)) \left. \right) \]

where function \( q_\rho(z) \) represents the data-quantization (of the shared information \( \frac{df_j}{dx_j} \)) as \([33], [34]\),

\[ q_\rho(z) = \text{sgn}(\rho \exp(\log(|z|, \rho))) \]

with \( q_u(z) = \rho \left( \frac{z}{\rho} \right) \) as the uniform quantizer (\([\cdot]\) denotes rounding to the nearest integer). The strongly sign-preserving odd function \( q_\rho(\cdot) \) represents logarithmic data-quantization with level \( \rho \), where \((1-\frac{\rho}{2})\leq q_\rho(z) \leq (1+\frac{\rho}{2})z\). For notation simplicity in the rest of the paper we use \( \psi_i = \frac{df_i}{dx_i} \) and \( W_{ij}(p, t) \).

In \([10]\), every agent knows its own state and objective along with information of its neighbors over \( \mathcal{G} \). We consider periodic communication, say every \( \omega \) sec, with sufficiently small \( \omega \) (similar to \([46, Theorem 10]\)). Every agent \( i \) shares \( v_i = \frac{df_i}{dx_i} \), with \( j \in N_i(t) \) via (logarithmic) quantized channels, where the agent \( j \) receives \( q_\rho(\frac{df_i}{dx_i}) \). Following the strict convexity of the level-sets of \( F \) and \([41, Proposition S2]\), for initialization point \( x_0 \in \mathcal{S}_K \) and its level set \( L_{F(x_0)} \), the solution in \( L_{F(x_0)} \) on \( \mathcal{S}_K \) under the differential inclusion \((10)\) exists and is locally bounded, upper semi-continuous, with non-empty, compact, and convex values. We use this along with \( \mathcal{H} \) referring to Lie-derivative with respect to \( \dot{x} = \partial \mathcal{H}(x) \) given by differential inclusion \([10]\) in the rest of the paper.

**Lemma 4:** Consider a feasible \( x_0 \in \mathcal{S}_K \). For any symmetric weight matrix \( W \), solution \( x(t) \) keeps its feasibility (sum-preserving) under dynamics \([10]\) for \( t > 0 \).

**Proof:** \( x_0 \in \mathcal{S}_K \) implies that \( \sum_{i=1}^{n} x_i(0) = K \). Then,

\[ \sum_{i=1}^{n} \dot{x}_i = -\sum_{i=1}^{n} W_{ij} \left( \eta_1 \text{sgn}^v(q_\rho(\psi_i) - q_\rho(\psi_j)) \right. \\\n+ \eta_2 \text{sgn}^v(q_\rho(\psi_i) - q_\rho(\psi_j)) \left. \right) \]

\[ = \sum_{i=1}^{n} W_{ij} \left( \eta_1 \text{sgn}^v(q_\rho(\psi_i) - q_\rho(\psi_j)) \right. \\\n+ \eta_2 \text{sgn}^v(q_\rho(\psi_i) - q_\rho(\psi_j)) \left. \right) \]

Note that \( \text{sgn}^v(q_\rho(\psi_i) - q_\rho(\psi_j)) = -\text{sgn}^v(q_\rho(\psi_i) - q_\rho(\psi_j)) \) and \( W_{ij} = W_{ji} \), therefore \( \sum_{i=1}^{n} \dot{x}_i = \sum_{i=1}^{n} x_i = 0 \).

**Remark 3:** In the virtue of Lemma 4 the solution \( x(t) \) remains feasible under dynamics \([10]\), i.e., \( x(t) \in \mathcal{S}_K \) for \( t > 0 \). Then, following Lemma 2 for initial value \( x_0 \in \mathcal{S}_K \), there is only one equilibrium \( x^* \) satisfying \( \nabla F(x^*) = \psi^* \mathbb{1}_n \).

**Theorem 2:** Under Assumptions 1 and 2 and starting from a feasible state \( x_0 \in \mathcal{S}_K \) dynamics \([10]\) solves the resource allocation problem \([2]\) (with feasibility at all times).

**Proof:** Following Lemma 1 for \( x^* \) as the optimal solution of problem \([2]\), \( \nabla F(x^*) = \psi^* \mathbb{1}_n \). From Lemma 4 \( x_0 \in \mathcal{S}_K \) implies solution feasibility under \([10]\), and therefore, \( \sum_{i=1}^{n} x_i^* = K \). Let \( F^* = F(x^*) \) and \( \overline{F}(x) = F(x) - F^* \) denoting the residual with respect to optimal value. Consider this positive \( \overline{F}(x) \) as the Lyapunov function with unique equilibrium \( \overline{F}(x^*) = 0 \). From \([41, Proposition 10]\), for this nonsmooth, regular, and locally Lipschitz Lyapunov function \( F(x) \), its derivative satisfies \( \partial \overline{F}(x(t)) = \mathcal{H}(\overline{F}(x(t))) [41, Proposition 10] \) with \( \mathcal{H} \) referring to dynamics \([10]\). We show here that this non-negative Lyapunov function is monotonically non-increasing under dynamics \([10]\). We have,

\[ \partial \overline{F}(x) = \sum_{i=1}^{n} \psi_i \left( -\sum_{j \in N_i} W_{ij} \left( \eta_1 \text{sgn}^v(q_\rho(\psi_i) - q_\rho(\psi_j)) \right. \\\n+ \eta_2 \text{sgn}^v(q_\rho(\psi_i) - q_\rho(\psi_j)) \left. \right) \right) \]
Then, in Lemma 3 set $g_2(\cdot)$ as $\text{sgn}^n(\cdot)$ and $g_1(\cdot)$ as $q_\rho(\cdot)$,

$$
\begin{align*}
\partial \overline{F}(x) &= -\frac{\eta_1}{2} \sum_{i,j=1}^{n} W_{ij} |q_\rho(\psi_i) - q_\rho(\psi_j)|^{\nu_1+1} \\
&\quad - \frac{\eta_2}{2} \sum_{i,j=1}^{n} W_{ij} |q_\rho(\psi_i) - q_\rho(\psi_j)|^{\nu_2+1}.
\end{align*}
$$

(13)

Therefore, $\partial \overline{F}(x) \leq 0$. Following Lemma 2, Remark 3 and Theorem 1 the unique point $x^*$ satisfying $\nabla F \in \text{span}\{1, n\}$, is the unique equilibrium of dynamics (10), and, based on Lemma 1 it is the optimal solution to the problem (2).

For any initial value $x_0 \in S_K$, the compact and closed (affine) solution set $\{L_{F(x_0)} \cap S_K\}$ is anytime feasible and positively invariant under (10). Thus, using LaSalle invariance principle for differential inclusions [22, Theorem 2.1], the solution converges to the largest invariant set $I$ contained in $\{x \in L_{F(x_0)} \cap S_K|0 \in L_{H}(x(t))\}$. Since $I = \{x^*\}$, $F \leq 0$, and $\max L_{H}(x(t)) < 0$ for all $x \in S_K \setminus I$, dynamics (10) globally asymptotically converges to $I = \{x^*\}$ [41, Theorem 1]. This completes the proof.

Recall that in Lemma 3 the oddness ensures anytime-feasibility. The connectivity requirement in Assumption 1 gives the unique optimal state $x^*$ (with $\nabla F(x^*) = \psi^*$) in Theorem 1 while Theorem 2 proves convergence to $x^*$.

Remark 4: The following gives an estimate of the convergence rate of Eq. (13).

$$
|q_\rho(\psi_i) - q_\rho(\psi_j)|^{\nu_1+1} + |q_\rho(\psi_i) - q_\rho(\psi_j)|^{\nu_2+1} \geq |q_\rho(\psi_i) - q_\rho(\psi_j)|^2
$$

(14)

where the RHS of the above represents the convergence rate of the linear (and linear quantized) protocols [11], [28]. Thus, the dynamics (10) is faster than its linear counterparts.

Remark 5: The results of this paper can be extended to consider saturation effects [33]. For example, in case of actuator saturation one may substitute $\text{sgn}(x)$ or $\text{sat}_\kappa(\cdot)$ instead of $\text{sgn}(\cdot)$ in dynamics (10), where $\text{sat}_\kappa(x) = x$ for $-\kappa \leq x \leq \kappa$ and $\kappa \text{sgn}(x)$ otherwise. Recall that the proofs of the given theorems and lemmas use only having non-zero derivative at zero, oddness, and sign-preserving property of $\text{sgn}(\cdot)$, which are also true for $\text{sat}_\kappa(\cdot)$ function. Therefore, the uniqueness, feasibility, and convergence results can be restated for general strongly sign-preserving nonlinearities on the agent’s dynamics and communications.

V. NUMERICAL SIMULATIONS

For the simulations, smooth penalty ($(|u|)^2$ [38] (with $\epsilon = 1$) for the box constraints is used to satisfy Assumption 2.

A. A Comparison Study

Consider the strictly-convex costs as [12],

$$
f_i(x_i) = b_i(x_i - a_i)^4,
$$

(15)

with random coefficients $b_i \in (0, 4]$, $a_i \in [-2, 4]$, and box constraints $0 \leq x_i \leq 5$. The random initial states satisfy $\sum_{i=1}^{n} x_i(0) = K = 20$ (as in Lemma 4). The multi-agent network is a cycle of $n = 10$ nodes with random stochastic link weights (this is required by [11], [16] and only for the sake of comparison). In Fig. 1 the convergence of the dynamics (10) is compared with linear [11], accelerated linear ($\beta = 0.6$) [16], quantized linear (with all-time triggered communications) [28], finite-time [20], and node-based fixed-time [23] (with time-period $\omega = 2 \times 10^{-5}$, $\eta = 1$, and $v_1 = 0.1, v_2 = 1.6$ in dynamics (10)).

B. Simulation over Weakly Connected Sparse Networks

We consider sparse networks of $n = 100$ agents every 0.05 sec switching between Scale-Free networks, none of which includes a spanning tree, while $\int_{t=0}^{t=0.2} G(t)$ is connected, i.e., $t_k = 0.5$ (100 time-periods with $\omega = 5 \times 10^{-3}$) in Assumption 4. The link weights are random and non-stochastic. Similar to [11], consider strictly-convex objectives as,

$$
f_i(x_i) = \frac{1}{2} a_i(x_i - c_i)^2 + \log(1 + \exp(b_i(x_i - d_i))),
$$

(16)

with random coefficients $a_i \in (0, 0.1]$, $b_i \in [-0.01, 0.01]$, $c_i, d_i \in [-0.5, 0.5]$ and box constraint $3 \leq x_i \leq 7$. The time-evolution of the states and absolute residual cost $\overline{F}(x)$ is shown in Fig. 2 with random initialization $\sum_{i=1}^{n} x_i(0) = K = 500, \eta = 1, \rho = 5 \times 10^{-4}$, and $v_1 = 0.3, v_2 = 1.6$.

VI. CONCLUSION

This work provides a distributed nonlinear 1st-order solution for resource allocation over dynamic undirected networks subject to (logarithmic) quantized data transmission, with convergence proved over sparse (uniformly-connected) networks. The explicit discretization of (10) (e.g., via Euler Forward method) can be used in real implementations assuming certain lower-bound on the sampling rate. For uniform quantization
with \[ \frac{dq_i}{dq} \bigg|_{-0.5 < c < 0.5} = 0 \] (sign-preserving but not strongly), one can prove convergence to \( \varepsilon \)-neighborhood of \( x^* \), as a direction of our future research.

**References**


