This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2023.3272220

A Finite-Time Protocol for Distributed Time-Varying Optimization over a Graph

Matteo Santilli, Antonio Furchì, Gabriele Oliva, *Senior Member, IEEE* and Andrea Gasparri, *Senior Member, IEEE*

***Abstract*—In this paper we address a time-varying quadratic optimization problem over a graph under the assumption that the problem shares the same sparsity pattern as the static graph encoding the undirected network topology over which the multi-agent systems interacts. Notably, this framework allows to effectively model scenarios in which the optimization problem is inherently embedded within the network topology, e.g., flow balancing, electrical power system managements or packet routing problems. In this regard, we propose a finite-time distributed algorithm which allows the multi-agent system to track the time-varying optimal solution over time. Specifically, we first solve the frozen-time optimization problem, providing a necessary and sufficient condition for a solution to be globally optimal. Then, based on such condition, a continuous-time distributed nonsmooth algorithm is developed. Numerical simulations are provided to corroborate the theoretical findings.**

***Index Terms*—Distributed Time-Varying Optimization, Graph Sparsity, Finite-Time Protocol, Nonsmooth Analysis.**

I. Introduction

IN recent years, a lot of attention has been devoted to multi-agent systems in several fields, including Internet of Things, social dynamics, or precision farming. In particular, multi-agent systems have been successfully adopted for solving distributed optimization problems [1], [2]–[14]. Generally speaking, in these settings the network mainly acts as a distributed computational resource over which the given optimization problem is solved. Differently, in our context we focus on optimization problems which are inherently embedded within the network topology, e.g., flow balancing [15], electrical power system management [16] or packet routing [17] problems. In this view, it is reasonable to assume that the matrices involved within the optimization problem are indeed sparse and their sparsity shares the underlying sparsity pattern of the network topology, while the time-varying terms in both the constraints and objective function account for the variability due to external factors. Notably, this setting is fundamentally different from classical distributed optimization approaches (e.g., see [18] and references therein). Our strategy to achieve finite-time tracking consist in deriving a nonsmooth algebraic optimality condition for the "frozen-time" problem, i.e., for the problem obtained by considering the actual values of the formulation at a specific time instant, and then developing a distributed algorithm able to track the time-varying solution of the nonsmooth algebraic optimality condition in finite-time. We resort to nonsmooth stability theory to prove convergence and finite-time tracking. To the best of our knowledge, this is the first work where finite-time tracking of a quadratic optimization problem with time-varying and coupling objective function and time-varying inequality constraints is obtained.

This research was partially supported by the Regione Lazio within the Program POR FESR Lazio 2014-2020, Avviso Pubblico "Progetti di Gruppi di Ricerca 2020," AGR-o-RAMA, under Agreement B85F21001360006.

M. Santilli, A.Furchì and A. Gasparri are with the Department of Civil, Computer Science and Aeronautical Technologies Engineering, Roma Tre University, 00146 Rome, Italy, (e-mail: matteo.santilli@uniroma3.it, antonio.furchi@uniroma3.it, gasparri@inf.uniroma3.it).

G. Oliva is with the Department of Engineering, Università Campus Bio-Medico di Roma, 00128, Rome, Italy (e-mail: g.oliva@unicampus.it, corresponding author).

In the literature there has been a large deal of work related to time-varying optimization for the centralized (e.g, see [20]–[22]) and distributed settings [1], [2]–[14], even though very few works were able to provide finite-time convergence guarantees [8], [13]. In particular, some approaches address time-invariant problem over a time-varying graph [1], [2]–[4], [23], while others focus on problems with time-varying objective functions and/or constraints [5]–[14]. Let us now briefly discuss the state of the art for the case of asymptotic convergence; existing works with asymptotic convergence are compared in Table I. Notably, discrete-time approaches [1], [7], [10], besides considering a distributed interaction protocol, rely on the iterative solution of local optimization problems and on additional local computations to be executed at each step, and only guarantee boundedness of the tracking error. Concerning continuous-time, except for [8], [13] and for the proposed algorithm, approaches in the literature either guarantee boundedness of the tracking error [9], [11] or asymptotic convergence [1], [5], [6], [12], [20]. It is noteworthy that [12] is the only work to guarantee asymptotic tracking in the general case of convex programming, under suitable assumptions; moreover, only [9] developed an algorithm that deals with directed communication between agents. Let us now discuss the algorithms with finite-time tracking capabilities, which are compared in Table I. In [19] an unconstrained optimization problem with time-varying objective functions is considered, where the agents are able to reach and maintain a consensus on their decision variables in finite time, but convergence to the time-varying optimal solution is asymptotic. Notice that [8] considers a particular resource allocation

		General Convex Programming	Quadratic Programming	Linear Programming	Resource Allocation Problem	Constrained	Coupled objective function	Coupled Constraints	Time-varying objective function	Time-varying constraints	1-hop information	2-hop information (via 1-hop interaction)
Asymptotic Convergence	Rahili and Ren, 2016 [5]	✓							✓		✓	
	Sun et al., 2017 [6]	✓				✓			✓	✓	✓	
	Vaquero and Cortés, 2018 [9]	✓				✓	✓	✓	✓	✓	✓	
	Hosseinzadeh et al., 2018 [11]			✓		✓	✓	✓		✓	✓	
	Sun and Ren, 2020 [12]	✓				✓			✓	✓	✓	
	Sun et. al, 2022 [14]	✓				✓			✓	✓	✓	
Finite-time Consensus Only	B. Ning, et al., 2017 [19]	✓							✓		✓	
Finite-time Convergence	Bai et al., 2018 [8]				✓	✓		✓		✓	✓	
	Santilli et al., 2020 [13]		✓			✓		✓		✓		✓
	Proposed Algorithm		✓			✓	✓	✓	✓	✓		✓

TABLE I**:** Comparison of the proposed algorithm against the state of the art.

problem with quadratic, invariant and decoupled objective functions, while time variance only occurs in an equality constraint that couples the variables. Moreover, [13] focuses on a particular quadratic problem where the aim is to minimize the square norm of the agents' variables, while the agents are coupled by a linear inequality constraint with time varying known term. However, [13] requires 2-hop information, while [8] relies on 1-hop information; such a 2-hop information can be retrieved by resorting to a state of the art finite-time k-hop distributed observer which can be implemented using only 1-hop information [24]. To summarize, most of previous literature is not able to provide finite-time convergence guarantees, while the few works that have this property are tailored to a particular class of problems [8] or consider simple and invariant objective functions [13]. In this paper, we aim to fill this gap. Specifically, this paper represents an extension of [13]; in fact, we introduce a number of improvements: a) we extend the setting to quadratic programming problems having objective function that is also time-varying; b) the objective function considered in this paper couples the agents, while in [13] the agents only aim to minimize the square norm of the decision variables; c) we allow the time-varying constraint vector to possess non derivable points; d) we provide a finite upper bound on the convergence time; e) we provide a computationally efficient bound on the gain required to guarantee finite-time tracking.

II. Notation and Preliminaries

We denote vectors by boldface lowercase letters and matrices with uppercase letters. We refer to the (i, j)-th entry of a matrix A by A_{ij}. We represent by $\mathbf{0}_n$ and $\mathbf{1}_n$ vectors with n entries, all equal to zero and to one, respectively. We use $2^{\mathbb{R}^n}$ to denote the power set, i.e., the set of all subsets of $\mathbb{R}^n$. We denote with I_n the identity matrix of size n and with $O_{n\times m}$ the zeros matrix with dimension $n\times m$. Given two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, we use $\max\{\boldsymbol{x}, \boldsymbol{y}\} \in \mathbb{R}^n$ and $\min\{\boldsymbol{x}, \boldsymbol{y}\} \in \mathbb{R}^n$ to denote the component-wise maximum and minimum, respectively. We denote with $\lambda_i(A)$ ($\sigma_i(A)$) the i-th largest eigenvalue (singular value) of the matrix $A \in \mathbb{R}^{n\times n}$, respectively. Moreover, we use $\lambda_{\max}(A)$ ($\sigma_{\max}(A)$) and $\lambda_{\min}(A)$ ($\sigma_{\min}(A)$) to denote the maximum and minimum eigenvalue (singular value) of A, respectively. We use $\|A\|$ and $\|A\|_F$ to denote the 2-norm and the Frobenius norm of a matrix A, respectively, while we use $\|\boldsymbol{x}\|$ and $\|\boldsymbol{x}\|_\infty$ to denote the Euclidean and the infinity norm of a vector $\boldsymbol{x}$, respectively. In addition, we introduce the discontinuous sign function $\text{sign}(\cdot) \in \mathbb{R}$:

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0; \end{cases}$$

and the discontinuous set-valued sign function $\text{SIGN}(x) \subset \mathbb{R}$:

$$\text{SIGN}(x) \in \begin{cases} \{1\} & \text{if } x > 0, \\ [-1, 1] & \text{if } x = 0, \\ \{-1\} & \text{if } x < 0, \end{cases}$$

and also define their respective vector forms:

$$\mathbf{sign}(\boldsymbol{x}) = [\text{sign}(x_1), ..., \text{sign}(x_n)]^T \in \mathbb{R}^n,$$
$$\mathbf{SIGN}(\boldsymbol{x}) = [\text{SIGN}(x_1), ..., \text{SIGN}(x_n)]^T \subset \mathbb{R}^n.$$

Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be an *undirected graph* with node set $\mathcal{V} = \{1, \dots, n\}$ with $|\mathcal{V}| = n$ and edges $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$, where $(i, j) \in \mathcal{E}$ captures the existence of a link from node i to node j. Note that, since the graph is undirected, the existence of an edge $(i, j) \in \mathcal{E}$ implies the existence of the edge $(j, i) \in \mathcal{E}$. Let us define a path between agents i and j as the set of edges through which an agent j can be reached by an agent i; in the following we will denote a path which involves k edges from agent i to reach agent j as a k-hop path between agents i and j. Let $\mathcal{N}_i^k$ denote the k-hop neighborhood of an agent i, that is the set of agents j for which there exists a p-hop path from agent j to i with $p \le k$. In addition, given a graph $\mathcal{G}$, let $\mathbb{A}_{\mathcal{G}}$ be the set of matrices compatible with it defined as

$$\mathbb{A}_{\mathcal{G}} = \left\{\Gamma \in \mathbb{R}^{n\times n} : \Gamma_{ij} = 0, \forall (i, j) \notin \mathcal{E}\cup\mathcal{C}\right\},$$

with $\mathcal{C} = \{(i, i)\}$, $i = 1, \dots, n$. Note that by definition the matrix $\Gamma \in \mathbb{A}_{\mathcal{G}}$ is not required to be symmetric and can have nonzero diagonal entries. The *degree* d_i of a node v_i

is the number of its incoming edges, i.e., $d_i = |\mathcal{N}_i|$. Given an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ with n nodes, we define the Laplacian matrix L as the $n \times n$ matrix such that

$$L_{ij} = \begin{cases} d_i, & \text{if } i = j \\ -1, & \text{if } (v_j, v_i) \in \mathcal{E}, \\ 0, & \text{otherwise.} \end{cases}$$

It is well known [25] that when $\mathcal{G}$ is connected L has a unique eigenvalue equal to zero and that the corresponding left eigenvector is $\mathbf{1}_n^T$.

Let us now review the Filippov solution concept for differential equations with discontinuous right-hand side, the nonsmooth analysis of Clarke's Generalized Gradient, and the chain-rule for differentiating regular functions along Filippov solution trajectories. The reader is referred to [26]–[28] and references therein for a comprehensive overview of the topic. Let us consider the differential equation

$$\dot{\boldsymbol{z}}(t) = \boldsymbol{f}(\boldsymbol{z}(t), t), \tag{1}$$

with $\boldsymbol{f} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ a measurable and essentially locally bounded function. In the following, where understood, we omit the time-dependency. First, we need to clarify what it means to be a solution of this equation.

Definition 1 (Filippov Solution): A vector function $\boldsymbol{z}(\cdot)$ is called solution of Eq. (1) on a time interval $[t_0, t_1]$ if $\boldsymbol{z}(\cdot)$ is absolutely continuous on $[t_0, t_1]$ and for almost all $t \in [t_0, t_1]$ it holds $\dot{\boldsymbol{z}} \in K[\boldsymbol{f}](\boldsymbol{z}, t)$, with

$$K[\boldsymbol{f}](\boldsymbol{z}, t) = \bigcap_{\delta > 0} \bigcap_{\mu\{H\}=0} \overline{\text{co}}\{\boldsymbol{f}(B(\boldsymbol{z}, \delta) \setminus H, t)\},$$

where $\bigcap_{\mu\{H\}=0}$ denotes the intersection over all sets H of Lebesgue measure zero, $B(\boldsymbol{z}, \delta)$ denotes the ball of radius δ centered at $\boldsymbol{z}$ and $\overline{\text{co}}$ the convex closure.
The ability to disregard sets of measure zero represents an interesting feature of the above definition that makes it possible to identify solutions even at locations where the vector field is not defined.

We now recall from [28] the conditions for the existence and the uniqueness of the Filippov solutions.

Proposition 1 (Existence and Uniqueness [28]): Let $\boldsymbol{f}(\boldsymbol{z}, t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ be measurable and locally essentially bounded. Assume that $\forall\, \boldsymbol{z} \in \mathbb{R}^n$ there exists $\epsilon > 0$ such that $\boldsymbol{f}(\cdot)$ is essentially one-sided Lipschitz on $B(\boldsymbol{z}, \epsilon)$. Then, for all $\boldsymbol{z}_0 \in \mathbb{R}^n$, there exists a unique Filippov solution of Eq. (1) with initial condition $\boldsymbol{z}(0) = \boldsymbol{z}_0$.

We now review the concept of Clarke's Generalized Gradient, an essential tool in the machinery of nonsmooth analysis.

Definition 2 (Clarke's Generalized Gradient): Consider a locally Lipschitz function $V : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$. Then, the generalized gradient at $(\boldsymbol{z}, t)$ is defined as

$$\partial V(\boldsymbol{z}, t) = \overline{\text{co}} \left\{ \lim_{k \to \infty} \nabla V(\boldsymbol{z}_k, t_k) : \Omega_V \not\ni (\boldsymbol{z}_k, t_k) \to (\boldsymbol{z}, t) \right\},$$

where Ω_V is the set of measure zeros where the gradient of V is not defined. Note that the gradient ∇ includes the derivative with respect to time $(\partial / \partial t)$.

We now review the chain rule which allows to differentiate Lipschitz regular functions along the Filippov's solution trajectories.

Theorem 1 (Chain Rule [27]): Let $\boldsymbol{z}(\cdot)$ be a Filippov solution to Eq. (1) on an interval containing t and $V : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ be a Lipschitz and, in addition, regular function. Then $V(\boldsymbol{z}(t), t)$ is absolutely continuous, $\frac{d}{dt}(V(\boldsymbol{z}(t), t))$ exists almost everywhere (i.e., save for a set of measure zero) and

$$\frac{d}{dt} V(\boldsymbol{z}(t), t) \in^{\text{a.e.}} \dot{\tilde{V}}(\boldsymbol{z}(t), t),$$

where a.e. is a shorthand for "almost everywhere" (The reader is referred [29] for a more comprehensive overview on nonsmooth analysis) and $\dot{\tilde{V}}(\boldsymbol{z}, t)$ is defined as

$$\dot{\tilde{V}}(\boldsymbol{z}, t) = \bigcap_{\boldsymbol{\xi} \in \partial V(\boldsymbol{z}(t), t)} \boldsymbol{\xi}^T \begin{pmatrix} K[\boldsymbol{f}](\boldsymbol{z}, t) \\ 1 \end{pmatrix}.$$

Let us now recall a revised version of the Generalized Lyapunov theorem given in [26] based on the results given in [27]. This will prove useful to establish finite-time stability results for dynamical systems described by differential equations with discontinuous right-hand side.

Theorem 2 (Finite-Time Stability Theorem): Consider a Filippov solution $\boldsymbol{z}(t) : \mathbb{R} \to \mathbb{R}^n$ to Eq. (1) and let $V(\boldsymbol{z}, t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, be a time dependent regular function such that $V(\boldsymbol{z}, t) = 0 \;\forall \boldsymbol{z} \in \mathcal{C}(t)$ and $V(\boldsymbol{z}, t) > 0 \;\forall \boldsymbol{z} \notin \mathcal{C}(t)$, with $\mathcal{C}(t) \subset \mathbb{R}^n$ a compact set. Furthermore, let $\boldsymbol{z}(t)$ and $V(\boldsymbol{z}, t)$ be absolutely continuous on $[t_0, \infty)$ with

$$\tfrac{d}{dt}(V(\boldsymbol{z}, t)) \leq -\epsilon < 0$$

almost everywhere on $\{t : \boldsymbol{z}(t) \notin \mathcal{C}(t)\}$. Then, $V(\boldsymbol{z}(t), t)$ converges to 0 in finite-time and $\boldsymbol{z}(t)$ reaches the compact set $\mathcal{C}(t)$ in finite-time as well.

Let us now also introduce the concept of generalized Jacobian. Consider a Lipschitz vector-valued function $F : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$, with $F = [F_1, \ldots, F_m]$. It follows from [29] that the generalized Jacobian $\partial F(\boldsymbol{z}, t)$ is

$$\partial F(\boldsymbol{z}, t) = \overline{\text{co}} \left\{ \lim_{i \to \infty} JF(\boldsymbol{z}_k, t_k) \,:\, \Omega_F \not\ni (\boldsymbol{z}_k, t_k) \to (\boldsymbol{z}, t) \right\},$$

with $JF(\boldsymbol{z}, t) \in \mathbb{R}^{m \times n}$ the classical Jacobian whenever it exists and Ω_F the set of measure zeros where $JF(\boldsymbol{z}, t)$ is not defined.

III. Problem Statement

Let us consider the following quadratic optimization problem with time-varying linear objective term and time-varying linear constraints with the same sparsity pattern as the static graph encoding the undirected network topology over which the multi-agent systems interacts.

Problem 1: Consider the following optimization problem over a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ with $|\mathcal{V}| = n$.

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \frac{1}{2} \boldsymbol{x}^T Q \boldsymbol{x} + \boldsymbol{\varphi}^T(t)\, \boldsymbol{x} \quad \text{s.t.} \quad A\, \boldsymbol{x} \geq \boldsymbol{b}(t)$$

with $Q \in \mathbb{A}_{\mathcal{G}}$, $\boldsymbol{\varphi}(t) \in \mathbb{R}^n$, $\boldsymbol{b}(t) \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$ with $m \leq n$. In particular, $A = DP$, where $P \in \mathbb{A}_{\mathcal{G}}$, while $D \in \mathbb{R}^{m \times n}$ is in the form $D = [\boldsymbol{e}_{i(1)}, \ldots, \boldsymbol{e}_{i(m)}]^T$, with $\boldsymbol{e}_j$ the j-th vector in the canonical basis in $\mathbb{R}^n$ and

$$\mathcal{H} = \{i(1), \ldots, i(m)\} \subseteq \{1, \ldots, n\} = \mathcal{V}.$$

In other words, the matrix A essentially amounts to a subset of the rows of a matrix P that is compatible with the graph underlying the agents' interaction.

In this paper we are interested in solving Problem 1 in a distributed manner as detailed in the following.

Problem 2: Let us consider a multi-agent system composed of n agents interconnected by a communication network $\mathcal{G}$. Our problem consist in designing a distributed control protocol that drives each agent i to a component of the optimal solution $\boldsymbol{x}^*(t)$ and the optimal Lagrange multiplier $\boldsymbol{\zeta}^*(t)$ of Problem 1 in finite-time T, i.e.,

$$\|x_i(t) - x_i^*(t)\| = 0, \quad \forall\, i \in \mathcal{V}, \quad \forall\, t \geq T,$$

$$\|\zeta_i(t) - \zeta_i^*(t)\| = 0, \quad \forall\, i \in \mathcal{H}, \quad \forall\, t \geq T.$$

Before moving forward with the technical derivations of the paper, we now discuss an interesting subclass of problems that represents a motivational example for the proposed framework.

Example 1 (Motivational Example): Let us consider a problem where each agent has two clashing objectives: from one side, the agents want their variables $x_i(t)$ to track a time-varying reference signal $\phi_i(t)$, while from another side the agents want to have values as similar as possible with each other (e.g., [30], where multi-objective approaches are used to drive the exploration task of mobile robots). This kind of setting has a number of application such as in exploration problems in the context of mobile robotics, where agents may want to explore different zones but also to stick with each other. Another interesting case is in the context of networks of distributed electrical prosumers, able to consume, provide or exchange energy with their neighbors; in this context, an interesting feature is the ability to mediate between the local utility of the agent, which can be considered to be time-varying based, for instance, on the energy prices (in this context the possibility to handle nonsmooth variation could be useful to model abrupt price changes), and requirements that the energy provided by the agents is is similar, in order to reduce the risk of instability. An example in this direction is given in [31], where *prosumers* interact by exchanging energy with their neighbors, and the amount of energy produced, exchanged or consumed is decided by solving a multi-objective optimization problem. From a practical standpoint, the objective function is in the form

$$f(\boldsymbol{x}(t), t) = \gamma \frac{1}{2} \sum_{i=1}^{n} (x_i(t) - \psi_i(t))^2 + (1-\gamma)\frac{1}{2} \sum_{i=1}^{n} \sum_{j \in \mathcal{N}_i} (x_i(t) - x_j(t))^2,$$

where the parameter γ is used to mediate between the two objectives. By some algebra, the above objective function can be equivalently expressed as

$$f(\boldsymbol{x}(t), t) = \frac{1}{2}\boldsymbol{x}^T \left(\gamma I_n + (1-\gamma)L\right)\boldsymbol{x} - \gamma \boldsymbol{\psi}^T(t)\boldsymbol{x} + \gamma \frac{1}{2}\boldsymbol{\psi}^T(t)\boldsymbol{\psi}(t)$$

notice that the term $\gamma\boldsymbol{\psi}^T(t)\boldsymbol{\psi}(t)/2$, being independent on $\boldsymbol{x}(t)$, can be neglected, in that the optimal solution does not change when a constant is added to the objetive function. As a consequence, the objective function becomes

$$\overline{f}(\boldsymbol{x}(t), t) = \frac{1}{2}\boldsymbol{x}^T \Bigg(\underbrace{\gamma I_n + (1-\gamma)L}_{Q}\Bigg)\boldsymbol{x} \underbrace{-\gamma\boldsymbol{\psi}^T(t)}_{\boldsymbol{\phi}^T(t)}\boldsymbol{x}$$

and it can be noted that, within the objective function $\overline{f}(\boldsymbol{x}(t), t)$, time variability only occurs in the linear term. As for the constraints, for simplicity, one may consider time-varying lower limits on $x_i(t) \geq b_i(t)$, i.e., constraints in the form $I_n\boldsymbol{x}(t) \geq \boldsymbol{b}(t)$. In the case of smart grids, such constraints can model minimum requirements for the power generation at each agent, which are in general time-varying to account for the demand at different time instants.

Let us now collect a set of technical assumptions that will be required to prove convergence of the proposed algorithm.

Assumption 1: The graph $\mathcal{G}$ is static, connected and undirected.

Assumption 2: Matrix Q is symmetric and positive definite.

Assumption 3: Matrix A is full row rank.

Assumption 4: The time-varying signals $\boldsymbol{\varphi}(t), \boldsymbol{b}(t)$ are absolutely continuous, bounded, and such that $\dot{\boldsymbol{\varphi}}(t) = \omega(t)$ and $\dot{\boldsymbol{b}}(t) = \chi(t)$ for all t with $\omega(t)$ and $\chi(t)$ locally essentially bounded functions. Moreover, it holds $\|\boldsymbol{\psi}\| < \kappa_\varphi$ for all $\boldsymbol{\psi} \in K[\omega](t)$ and $\|\boldsymbol{\beta}\| < \kappa_b$ for all $\boldsymbol{\beta} \in K[\chi](t)$.

We point out that, while we require $\boldsymbol{\varphi}(t)$ and $\boldsymbol{b}(t)$ to be bounded, we do not need to know their actual bounds; instead, as discussed later, the agents will need to know a bound on their derivative. In addition, such assumption is intrinsically satisfied for well-posed problems (i.e., when $\boldsymbol{b}(t)$ is not bounded, the problem can be easily show to be either unfeasible or unconstrained, while when $\boldsymbol{\varphi}(t)$ is not bounded the problem reduces to finding a feasible solution).

Let us now define the set

$$\mathbb{X}(t) = \{\boldsymbol{y} \in \mathbb{R}^n \mid A\boldsymbol{y} + \boldsymbol{b}_i(t) \leq \boldsymbol{0}_m\}$$

The next assumption is required to guarantee that the problem at hand is feasible at all time instants.

Assumption 5: For all time instants $t \in [0, \infty)$ the set $\mathbb{X}(t)$ is nonempty.

As discussed later in the paper, the next technical assumption is required in order to set up a proper gain α in our algorithm.

Assumption 6: Matrix $Q - A^T A$ is positive definite.

Assumption 6 is given without loss of generality. In particular, as discussed later in Remark 4, if the assumption is not satisfied, it is sufficient to scale the objective function by a constant $\beta > 0$, which can be computed in a distributed fashion, obtaining an equivalent problem which satisfies the assumption.

The next assumption characterizes the information available to each agent

Assumption 7: Each agent i knows:

- The total number n of agents
- The total number m of constraints
- the entries Q_{ij} for all $j \in \mathcal{N}_i$
- the entries A_{ji} for all $j \in \mathcal{N}_i$
- the entries $A_{\ell j}$ for all $j \in \mathcal{N}_i$, if the agent is responsible for the l-th constraint
- $\phi_i(t)$ and $b_i(t)$
- the constants κ_φ and κ_b.
- the minimum eigenvalue $\lambda_{\min}(Q)$ of Q
- the minimum singular value $\sigma_{\min}(C)$ of

$$C = \begin{bmatrix} O_{m\times m} & -A \\ A^T & Q - A^T A \end{bmatrix} \tag{2}$$

Remark 1: Knowledge of $\lambda_{\min}(Q)$ and $\sigma_{\min}(C)$ is required in order to adequately choose the gain α within the proposed algorithm. Notice that, as it will be shown later in the paper, under Assumption 6, C is guaranteed to be nonsingular and thus $\sigma_{\min}(C) > 0$. Notice further that, as discussed later in the paper, the requirement to know $\sigma_{\min}(C)$ can be relaxed, as it is possible to derive a positive lower bound on $\sigma_{\min}(C)$ that only depends on $\lambda_{\min}(Q), \|Q\|_F, \sigma_{\min}(A), \|A\|_F$ and n. This requirement could be lifted resorting to adaptive gains, which represent a valuable future work direction.

Remark 2: The objective function of Problem 1 is convex by construction and the constraints are linear. Therefore, assuming a feasible solution exists at all time instants, the Slater's Constraint qualification holds true at all times [32].

Remark 3: In this paper, for the sake of simplicity, we assume each agent is associated to a scalar choice variable. However, the approach can easily be extended to the vectorial case where each agent is associated to vectorial variables $\boldsymbol{x}_i(t) \in \mathbb{R}^h$ and time-varying signals $\boldsymbol{b}_i(t), \boldsymbol{\phi}_i(t) \in \mathbb{R}^m$ and the aim is to solve a problem where $Q \in \mathbb{R}^{nh\times nh}$ is positive definite and $A \in \mathbb{R}^{\ell\times nh}$, with $\ell \leq nh$ is full row rank. In particular, in order to generalize the approach, matrices Q and A should exhibit the same sparsity pattern as the graph. For instance, it is possible to consider matrices Q with the following structure

$$Q = I_n \otimes Q^{\texttt{local}} + Q^{\texttt{interaction}} \otimes Q^{\texttt{coupling}},$$

where $Q^{\texttt{local}} \in \mathbb{R}^{m\times m}$ models a local term that only depends on the choice variables available at each agent, while the second term is the combination of $Q^{\texttt{interaction}} \in \mathbb{R}^{n\times n}$, which accounts for the agents' interaction and has the same structure as the communication graph, and $Q^{\texttt{coupling}} \in \mathbb{R}^{m\times m}$, which models the influence among pairs of agents.

IV. Frozen-Time Global Optimal Solution

Let us now characterize the structure of the global optimal solution at a given time instant t, which we refer to as the *frozen-time solution* at time t. The result will be a system of nonsmooth algebraic equations in the Lagrangian multipliers, which will be the basis for the proposed algorithm.

Theorem 3: Consider Problem 1 under Assumptions 1–5; the frozen-time formulation at any time instant $t \geq 0$ has a unique global optimal solution $\boldsymbol{x}^*(t)$ and Lagrange multipliers $\boldsymbol{\zeta}^*(t) \in \mathbb{R}^m$ for the inequality constraint that satisfy

$$\underbrace{\begin{bmatrix} \min\{\boldsymbol{\zeta}^*(t), A\boldsymbol{x}^*(t) - \boldsymbol{b}(t)\} \\ Q\boldsymbol{x}^*(t) + \boldsymbol{\varphi}(t) - A^T\boldsymbol{\zeta}^*(t) \end{bmatrix}}_{\boldsymbol{h}(\boldsymbol{x}^*(t),\boldsymbol{\zeta}^*(t),t)} = \mathbf{0}_{n+m}. \tag{3}$$

Proof: The proof follows by classical KKT theory (e.g., see [32], [33]). In particular, we have that the Lagrangian function is

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\zeta}) = \frac{1}{2}\boldsymbol{x}^T Q\,\boldsymbol{x} + \boldsymbol{\varphi}^T(t)\,\boldsymbol{x} + (\boldsymbol{b}(t) - A\boldsymbol{x})^T\,\boldsymbol{\zeta},$$

$\boldsymbol{\zeta}$ being the vector of Lagrange multipliers associated to the constraint. Moreover, $\boldsymbol{x}^*(t), \boldsymbol{\zeta}^*(t)$ are globally optimal and, in particular, $\boldsymbol{x}^*(t)$ is unique, if and only if the following conditions hold true:

1) $\nabla_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x}^*(t), \boldsymbol{\zeta}^*(t)) = \mathbf{0}_n$;
2) $A\boldsymbol{x}^*(t) - \boldsymbol{b}(t) \geq \mathbf{0}_m$;
3) $\boldsymbol{\zeta}^*(t) \geq \mathbf{0}_m$;
4) $\boldsymbol{\zeta}^*(t) \odot (A\boldsymbol{x}^*(t) - \boldsymbol{b}(t)) = \mathbf{0}_m$,

where $\odot$ denotes the Hadamard product.The proof follows noting that points 2)–4) and 1) correspond, respectively, to the first and second block of $\boldsymbol{h}(\cdot)$ in Eq. (3). ∎

We now establish two results on the Lagrange multiplier vector corresponding to the optimal solution of the frozen-time problem. We will utilize these results later to prove the finite-time convergence and tracking properties of our protocol.

Proposition 2: Under Assumptions 1–5, the set of Lagrange multiplier vector $\boldsymbol{\zeta}^*(t)$ corresponding to the global optimal solution $\boldsymbol{x}^*(t)$ is unique for all $t \geq 0$.

Proof: To prove the result, as described in [34] and references therein, it is sufficient to show that the *Linear-Independence Constraint Qualification* (LICQ) holds true, i.e., the fact the gradients of the constraints evaluated at $\boldsymbol{x}^*(t)$ are linearly independent. Notably, in our case the matrix having such gradients as columns corresponds to A^T. Therefore, by Assumption 3, LICQ is verified. The proof is complete. ∎

V. Distributed Optimization Algorithm Design

In this section, we develop a distributed algorithm to solve the time-varying optimization problem illustrated in Problem 2. Note that, from Theorem 3, at each time instant t the optimal solution $\boldsymbol{x}^*(t)$ and the optimal Lagrange multipliers $\boldsymbol{\zeta}^*(t)$ satisfy Eq. (3). Therefore, our goal is to enforce this condition for any $t \geq T$, with $T > 0$. To achieve this goal, let us introduce the stacked vector $\boldsymbol{z}(t) \in \mathbb{R}^{n+m}$ collecting the set of Lagrangian multipliers $\boldsymbol{\zeta}(t)$ and the state $\boldsymbol{x}(t)$ as

$$\boldsymbol{z}(t) = \left[\boldsymbol{\zeta}(t)^T,\ \boldsymbol{x}(t)^T\right]^T.$$

Notably, based on the definition of A in Problem 2, we have that only the subset $\mathcal{H} \subseteq \mathcal{V}$ of agents is in charge of handling a constraint and it is thus associated to a variable $\zeta_i(t)$, which models the corresponding Lagrange multiplier. Conversely, each agent $i \in \mathcal{V}$ is associated to a variable $x_i(t)$.

For the sake of readability let us introduce the functions

$$\boldsymbol{w} : \mathbb{R}^{n+m} \times \mathbb{R} \to \mathbb{R}^m \quad \text{and} \quad \boldsymbol{g} : \mathbb{R}^{n+m} \times \mathbb{R} \to \mathbb{R}^n,$$

defined as

$$\boldsymbol{w}(\boldsymbol{z}, t) = \min\{\boldsymbol{\zeta}(t), A\boldsymbol{x}(t) - \boldsymbol{b}(t)\}, \tag{4}$$

$$\boldsymbol{g}(\boldsymbol{z}, t) = Q\boldsymbol{x}(t) + \boldsymbol{\varphi}(t) - A^T\boldsymbol{\zeta}(t), \tag{5}$$

and the function $\boldsymbol{y} : \mathbb{R}^{n+m} \times \mathbb{R} \to \mathbb{R}^m$ defined as

$$\boldsymbol{y}(\boldsymbol{z}, t) = A\boldsymbol{x}(t) - \boldsymbol{b}(t) - \boldsymbol{\zeta}(t).$$

Furthermore, let $S(\boldsymbol{z}, t)$ be the diagonal $m \times m$ matrix such that $S_{ii}(\boldsymbol{z}, t) = 1$ if $y_i(\boldsymbol{z}, t) \leq 0$, 0 otherwise. and let also $\mathbb{S}(\boldsymbol{z}, t)$ be the set of diagonal matrices $\mathcal{S}(\boldsymbol{z}, t) \in \mathbb{R}^{m\times m}$ with structure

$$\mathcal{S}_{ii}(\boldsymbol{z}, t) \in \begin{cases} \{1\} & \text{if } y_i(\boldsymbol{z}, t) < 0, \\ [0, 1] & \text{if } y_i(\boldsymbol{z}, t) = 0, \\ \{0\} & \text{if } y_i(\boldsymbol{z}, t) > 0, \end{cases} \quad \forall\, i = 1, \ldots, m. \tag{6}$$

Matrices $S(\boldsymbol{z}, t)$ and $\mathcal{S}(\boldsymbol{z}, t)$ will be used later to compute the derivative of the minimum function in Eq. (4).

We now outline our distributed protocol. Specifically, from the perspective of the i-th agent, the proposed algorithm reads as follows (we omit dependencies on the state $\boldsymbol{z}$ and on time t for the sake of readability)

$$\begin{aligned} \dot{\zeta}_i &= -\,\alpha\,(1 - S_{ii})\,\text{sign}(w_i) + \alpha \sum_{j=1}^{n} A_{ij}\,\text{sign}(g_j), \quad \forall\, i \in \mathcal{H}, \\ \dot{x}_i &= -\,\alpha \sum_{j=1}^{m} A_{ji} S_{jj}\,\text{sign}(w_j) - \alpha \sum_{j=1}^{n} Q_{ij}\,\text{sign}(g_j), \quad \forall\, i \in \mathcal{V}. \end{aligned} \tag{7}$$

Note that, in order to implement Eq. (7), each agent i is required to collect the following information:

i) the state variables $z_l(t) = [\zeta_l(t)\; x_l(t)]^T$ of the agents $l \in \mathcal{N}_i^2$ belonging to its 2-hop neighborhood (i.e., in order to compute the functions w_j and g_j for each 1-hop neighbor $j \in \mathcal{N}_i^1$);
ii) the elements A_{ji} for the agents $j \in \mathcal{N}_i^1$ belonging to its 1-hop neighborhood;
iii) the elements A_{jl}, A_{lj}, and Q_{jl} for the 2-hop neighbors $l \in \mathcal{N}_j^1$ such that $j \in \mathcal{N}_i^1$;
iv) the time-varying values $\varphi_j(t)$, $b_j(t)$ for the agents $j \in \mathcal{N}_i^1$ belonging to its 1-hop neighborhood.

For point i) we notice that the states $z_l(t)$ of the 2-hop neighborhood can be locally estimated by agent i through 1-hop local interactions by resorting to the state of the art finite-time k-hop distributed observer proposed in [24]; while for points ii) and iii) we observe that since the required elements are constant, they can be exchanged once before the execution of the proposed algorithm; finally, for point iv), we assume that the values $\varphi_j(t)$, $b_j(t)$ of the 1-hop neighborhood can be exchanged through 1-hop communication.

Stacking Eq. (7) $\forall\, i \in \mathcal{V}$ yields the following matrix form

$$\begin{aligned} \dot{\boldsymbol{z}}(t) = \boldsymbol{f}(\boldsymbol{z}, t) &= -\alpha\, M(\boldsymbol{z}, t) \begin{bmatrix} \textbf{sign}(\boldsymbol{w}(\boldsymbol{z}, t)) \\ \textbf{sign}(\boldsymbol{g}(\boldsymbol{z}, t)) \end{bmatrix} \\ &= -\alpha\, M(\boldsymbol{z}, t)\, \textbf{sign}(\boldsymbol{h}(\boldsymbol{z}, t)), \end{aligned} \tag{8}$$

with $\boldsymbol{h}(\boldsymbol{z}, t)$ defined in Eq. (3) and $M(\boldsymbol{z}, t) \in \mathbb{R}^{(n+m)\times(n+m)}$ the matrix defined as

$$M(\boldsymbol{z}, t) = \begin{bmatrix} I_m - S(\boldsymbol{z}, t) & -A \\ A^T S(\boldsymbol{z}, t) & Q \end{bmatrix}. \tag{9}$$

In the sequel, we will analyze the convergence properties of our proposed algorithm given in Eq. (7) considering its equivalent matrix version given in Eq. (8). Moreover, we will show how to choose the gain α in order to guarantee convergence, based on the information available to the agents as per Assumption 7. Finally, we will also demonstrate how to relax the assumption that the agents need to know $\sigma_{\min}(C)$.

A. Convergence Analysis

In order to establish convergence of the proposed algorithm, let us first introduce a preliminary result. In this view, let

$$\mathbb{M}(\boldsymbol{z}, t) = \{M_{\mathcal{S}}(\boldsymbol{z}, t) : \mathcal{S}(\boldsymbol{z}, t) \in \mathbb{S}\}$$

be the set collecting all matrices $M_{\mathcal{S}}(\boldsymbol{z}, t) \in \mathbb{R}^{(n+m)\times(n+m)}$ defined as

$$M_{\mathcal{S}}(\boldsymbol{z}, t) = \begin{bmatrix} I_m - \mathcal{S}(\boldsymbol{z}, t) & -A \\ A^T \mathcal{S}(\boldsymbol{z}, t) & Q \end{bmatrix}. \tag{10}$$

Lemma 1: Let Assumptions 2 and 3 hold. Then, every matrix $M_{\mathcal{S}}(\boldsymbol{z}, t) \in \mathbb{M}(\boldsymbol{z}, t)$ defined as in Eq. (10) is nonsingular.

Proof: In order to prove the result we observe that, by Assumption 2, the lower diagonal block of $M_{\mathcal{S}}$ is invertible and its Schur complement with respect to such a block is $M_{\mathcal{S}}/Q = I_m - \mathcal{S} + A\,Q^{-1}A^T\mathcal{S}$. It is well known, e.g., [35], that it holds $\det(M_{\mathcal{S}}) = \det(Q)\,\det(M_{\mathcal{S}}/Q)$. Since by construction $\det(Q) \neq 0$, we have that $\det(M_{\mathcal{S}}) \neq 0$ if and only if $\det(M_{\mathcal{S}}/Q) \neq 0$. In view of a contradiction, suppose $\det(M_{\mathcal{S}}/Q) = 0$. This means there is $\boldsymbol{v} \neq \boldsymbol{0}_m$ such that

$$(M_{\mathcal{S}}/Q)^T\boldsymbol{v} = (I_m - \mathcal{S} + \mathcal{S}A\,Q^{-1}A^T)\,\boldsymbol{v} = \boldsymbol{0}_m. \tag{11}$$

Notice that if the diagonal entries of $\mathcal{S}$ are all zero we have that Eq. (11) is satisfied only for $\boldsymbol{v} = \boldsymbol{0}_m$, i.e., we reach a contradiction. Hence, let us assume that $\mathcal{S}$ has $0 < \ell \leq m$ nonzero diagonal entries and let P be the $m \times m$ permutation matrix such that

$$\widehat{\mathcal{S}} = P\mathcal{S}P^T = \begin{bmatrix} \widehat{\mathcal{S}}_1 & O_{\ell\times(m-\ell)} \\ O_{(m-\ell)\times\ell} & O_{(m-\ell)\times(m-\ell)} \end{bmatrix}, \tag{12}$$

has the first ℓ diagonal entries that are positive, i.e., $\widehat{\mathcal{S}}_{11} \in (0, 1], \ldots, \widehat{\mathcal{S}}_{\ell\ell} \in (0, 1]$. Since $\boldsymbol{v} \neq \boldsymbol{0}_m$, we can write $\boldsymbol{v} = P^T\widehat{\boldsymbol{v}}$, for some $\widehat{\boldsymbol{v}} \neq \boldsymbol{0}_m$; therefore, noting that $PP^T = I_m$, by premultiplying Eq. (11) by P we obtain

$$\begin{aligned} \boldsymbol{0}_m &= P\,(I_m - \mathcal{S} + \mathcal{S}A\,Q^{-1}A^T)P^T\widehat{\boldsymbol{v}} \\ &= (I_m - \widehat{\mathcal{S}} + P\mathcal{S}AQ^{-1}A^TP^T)\,\widehat{\boldsymbol{v}} = (I_m - \widehat{\mathcal{S}} + \widehat{\mathcal{S}}\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)\,\widehat{\boldsymbol{v}}, \end{aligned} \tag{13}$$

where $\widehat{A}$ and $\widehat{Q}$ are obtained by permutation of rows and columns of A and Q, respectively. Let us now partition $\widehat{\boldsymbol{v}}$ as $\widehat{\boldsymbol{v}} = [\widehat{\boldsymbol{v}}_1^T, \widehat{\boldsymbol{v}}_2^T]^T$, with $\widehat{\boldsymbol{v}}_1 \in \mathbb{R}^\ell$ and $\widehat{\boldsymbol{v}}_2 \in \mathbb{R}^{m-\ell}$. Such a decomposition induces the following block decomposition for the matrix $\widehat{A}\widehat{Q}^{-1}\widehat{A}^T$

$$\widehat{A}\widehat{Q}^{-1}\widehat{A}^T = \begin{bmatrix} (\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{11} & (\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{12} \\ (\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{21} & (\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{22} \end{bmatrix}, \tag{14}$$

With this decomposition, Eq. (13) corresponds to

$$(I_\ell - \widehat{\mathcal{S}}_1)\widehat{\boldsymbol{v}}_1 + \widehat{\mathcal{S}}_1(\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{11}\widehat{\boldsymbol{v}}_1 + \widehat{\mathcal{S}}_1(\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{12}\widehat{\boldsymbol{v}}_2 = \mathbf{0}_\ell, \tag{15}$$

and $\widehat{\boldsymbol{v}}_2 = \mathbf{0}_{m-\ell}$. Therefore, Eq. (15) corresponds to

$$(I_\ell - \widehat{\mathcal{S}}_1 + \widehat{\mathcal{S}}_1(\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{11})\widehat{\boldsymbol{v}}_1 = \mathbf{0}_\ell. \tag{16}$$

Since $\widehat{\mathcal{S}}_1$ is nonsingular by construction, Eq. (16) can be rearranged as

$$\widehat{\mathcal{S}}_1\big(\widehat{\mathcal{S}}_1^{-1} - I_\ell + (\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{11}\big)\widehat{\boldsymbol{v}}_1 = \widehat{\mathcal{S}}_1\widehat{H}\widehat{\boldsymbol{v}}_1 = \mathbf{0}_\ell.$$

At this point we observe that $(\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{11}$ is the leading principal minor of $\widehat{A}\widehat{Q}^{-1}\widehat{A}^T$ of size ℓ. Notice that by Assumptions 2 and 3, $AQ^{-1}A^T$ is symmetric and positive definite; therefore, by construction, $\widehat{A}\widehat{Q}^{-1}\widehat{A}^T$ is symmetric and positive definite. As a consequence, by the Sylvester's Criterion [36], also $(\widehat{A}\widehat{Q}^{-1}\widehat{A}^T)_{11}$ is symmetric and positive definite. Furthermore, we observe that $\widehat{\mathcal{S}}_1^{-1} - I_\ell$ is diagonal and positive semidefinite since its diagonal entries are

$$(\widehat{\mathcal{S}}_1^{-1} - I_\ell)_{ii} = (\widehat{\mathcal{S}}_1)_{ii}^{-1} - 1 \geq 0, \tag{17}$$

being $0 < (\widehat{\mathcal{S}}_1)_{ii} \leq 1$. Therefore $\widehat{H}$ is the sum of a positive semidefinite and a positive definite matrix and is thus positive definite. Moreover, since $\widehat{\mathcal{S}}_1$ is positive definite, $\widehat{\mathcal{S}}_1\widehat{H}$ is nonsingular. Hence, Eq. (16) is satisfied only for $\widehat{\boldsymbol{v}}_1 = \mathbf{0}_\ell$. This implies that Eq. (11) is satisfied only for $\boldsymbol{v} = \mathbf{0}_m$. We reached a contradiction; therefore, the Schur complement $M_{\mathcal{S}}/Q$ is nonsingular, and this implies that $M_{\mathcal{S}}$ is nonsingular. ■

Let us now to prove that our distributed algorithm introduced in Eq. (8) allows our system to reach the compact set $\mathcal{C}(t) = \{\boldsymbol{z}^*(t)\}$, i.e., the singleton corresponding to the unique optimal solution at time t, in finite-time T and then remain contained therein for $t \geq T$, i.e., is able to solve Problem 2 in finite-time.

Theorem 4: Consider the settings of Problem 2 and let the agents run the proposed protocol in Eq. (8). Let Assumptions 1–7 hold and suppose that the coefficient $\epsilon > 0$ and ρ are known to the agents. Assume also that the gain α satisfies

$$\alpha > \frac{\sqrt{n}\,(\kappa_b + \kappa_\varphi) + \epsilon}{\rho^2}, \tag{18}$$

where $\epsilon > 0$ is a design parameter used to impose arbitrary convergence time and $\rho > 0$ is

$$\rho = \min_{M_{\mathcal{S}} \in \mathbb{M}} \left(\sigma_{\min}(M_{\mathcal{S}})\right), \tag{19}$$

with $\sigma_{\min}(M_{\mathcal{S}})$ the smallest singular value of the matrix $M_{\mathcal{S}}$ introduced in Eq. (10). Then, there exits $T(\epsilon) > 0$ such that the stacked vector $\boldsymbol{h}(\boldsymbol{z}, t)$ introduced in Eq. (3) converges to zero in finite-time, that is, $\|\boldsymbol{h}(\boldsymbol{z}, t)\|_1 = 0$, for all $t \geq T(\epsilon)$, where the convergence time $T(\epsilon)$ is upper bounded by a positive finite value $T_{\max} = \frac{1}{\epsilon}\,\|\boldsymbol{h}(\boldsymbol{z}(0), 0)\|_1$.

Proof: Consider the following generalized time-varying Lyapunov-like function $V(\boldsymbol{z}, t) : \mathbb{R}^{n+m} \times \mathbb{R} \to \mathbb{R}$

$$V(\boldsymbol{z}, t) = \|\boldsymbol{h}(\boldsymbol{z}, t)\|_1\,, \tag{20}$$

which, by Proposition 2, satisfies $V(\boldsymbol{z}, t) = 0$ for $\boldsymbol{z} = \boldsymbol{z}^*(t)$ and $V(\boldsymbol{z}, t) > 0$ for all $\boldsymbol{z} \neq \boldsymbol{z}^*(t)$. We now prove that the Lyapunov-like function introduced in Eq. (20) reaches zero in finite-time $T(\epsilon)$ and remains zero $\forall\, t \geq T(\epsilon)$.

In order to apply Theorem 1, let us now compute the generalized gradient $\partial V(\boldsymbol{z}, t)$ as

$$\partial V(\boldsymbol{z}, t) \subset \begin{bmatrix} \dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{z}} & \dfrac{\partial \boldsymbol{h}}{\partial t} \end{bmatrix}^T \mathbf{SIGN}(\boldsymbol{h}(\boldsymbol{z}, t)) = \begin{bmatrix} \partial_{\boldsymbol{z}} V \\ \partial_t V \end{bmatrix}, \tag{21}$$

from the application of [29, Thereom 2.6.6] and [26, Theorem 1]. It can be noticed how the $\mathrm{co}\{\cdot\}$ in Eq. (21) is superfluous, since $[\partial \boldsymbol{h}/\partial \boldsymbol{z}\ \partial \boldsymbol{h}/\partial t]^T\, \mathbf{SIGN}(\boldsymbol{h})$ is a vector of closed intervals [37] and thus convex by construction.

Since $\boldsymbol{h}(\boldsymbol{z}, t)$ can be discontinuous for some i when $y_i(\boldsymbol{z}, t) = 0$, in general the terms $\partial \boldsymbol{h}/\partial \boldsymbol{z}$ and $\partial \boldsymbol{h}/\partial t$ are not singletons and generate the following structure for $\partial V(\boldsymbol{z}, t)$

$$\partial V(\boldsymbol{z}, t) \subset \left\{ \begin{bmatrix} I_n - \mathcal{S} & -A \\ A^T\mathcal{S} & Q \\ -\boldsymbol{\beta}^T\mathcal{S} & \boldsymbol{\psi}^T \end{bmatrix} \mathbf{SIGN}(\boldsymbol{h}) : \begin{matrix} \mathcal{S} \in \mathbb{S}, \\ \boldsymbol{\beta} \in K[\chi](t), \\ \boldsymbol{\psi} \in K[\omega](t) \end{matrix} \right\}. \tag{22}$$

In order to analyze the structure of a generic element $\boldsymbol{\xi} \in \partial V$, let us introduce the matrix $M_{\mathcal{S}}^{\boldsymbol{z}} \in \mathbb{R}^{n+m \times n+m}$ and the vector $M_{\mathcal{S}}^{t} \in \mathbb{R}^{1 \times n+m}$ defined as

$$M_{\mathcal{S}}^{\boldsymbol{z}} = \begin{bmatrix} I_m - \mathcal{S} & -A \\ A^T\mathcal{S} & Q \end{bmatrix} \quad \text{and} \quad M_{\mathcal{S}}^{t} = \begin{bmatrix} -\boldsymbol{\beta}^T\mathcal{S} & \boldsymbol{\psi}^T \end{bmatrix}, \tag{23}$$

with $\mathcal{S} \in \mathbb{S}$, $\boldsymbol{\beta} \in K[\chi](t)$, and $\boldsymbol{\psi} \in K[\omega](t)$. An element $\boldsymbol{\xi} \in \partial V$ can then be expressed as

$$\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\xi}_{\boldsymbol{z}} \\ \boldsymbol{\xi}_t \end{bmatrix} = \begin{bmatrix} M_{\mathcal{S}}^{\boldsymbol{z}} \\ M_{\mathcal{S}}^{t} \end{bmatrix} \boldsymbol{\eta}, \qquad \text{with } \boldsymbol{\eta} \in \mathbf{SIGN}(\boldsymbol{h}). \tag{24}$$

In virtue of the above equations, we can now restate $\dot{\tilde{V}}(\boldsymbol{z}, t)$ as

$$\dot{\tilde{V}}(\boldsymbol{z}, t) = \bigcap_{[\boldsymbol{\xi}_{\boldsymbol{z}}^T\ \boldsymbol{\xi}_t^T]^T \in \partial V} \boldsymbol{\xi}_{\boldsymbol{z}}^T\, K[\boldsymbol{f}](\boldsymbol{z}, t) + \boldsymbol{\xi}_t^T\, 1\,.$$

Now, since the proposed control law is the nonsmooth version of the classical gradient descent flow of a differentiable function, an element $\boldsymbol{v} \in K[\boldsymbol{f}](\boldsymbol{z}, t)$ is in the form

$$\boldsymbol{v} \in -\alpha\, \partial_{\boldsymbol{z}} V,$$

thus $\dot{\tilde{V}}(\boldsymbol{z}, t)$ can be further developed as

$$\dot{\tilde{V}}(\boldsymbol{z}, t) = \bigcap_{[\boldsymbol{\xi}_{\boldsymbol{z}}^T\ \boldsymbol{\xi}_t]^T \in \partial V} -\alpha\, \boldsymbol{\xi}_{\boldsymbol{z}}^T\, \partial_{\boldsymbol{z}} V + \boldsymbol{\xi}_t,$$

where we point out that $\boldsymbol{\xi}_t$ is a scalar in virtue of Eqs. (23) and (24). At this point we can proceed by applying a similar reasoning as in [27]. In particular, since $\partial_{\boldsymbol{z}} V$ is convex, it follows that for all $\boldsymbol{z} \neq \boldsymbol{z}^*(t)$ there exists

$$\widehat{\boldsymbol{\xi}} = [\widehat{\boldsymbol{\xi}}_{\boldsymbol{z}}^T\ \boldsymbol{\xi}_t]^T \in \partial V$$

such that

$$-\alpha\, \boldsymbol{\xi}_{\boldsymbol{z}}^T\, \widehat{\boldsymbol{\xi}}_{\boldsymbol{z}} + \boldsymbol{\xi}_t\ \leq -\alpha\, \|\widehat{\boldsymbol{\xi}}_{\boldsymbol{z}}\|^2 + \boldsymbol{\xi}_t,\ \forall\, \boldsymbol{\xi} \in \partial V.$$

Considering now $\widehat{\boldsymbol{\xi}}_{\boldsymbol{z}} = M_{\widehat{\mathcal{S}}}^{\boldsymbol{z}}\, \widehat{\boldsymbol{\eta}}$ and $\boldsymbol{\xi}_t = M_{\mathcal{S}}^{t}\, \boldsymbol{\eta}$, we obtain the following bound on the generalized time-derivative

$$\frac{d}{dt}(V(\boldsymbol{z}, t)) \in^{\text{a.e.}} \dot{\tilde{V}}(\boldsymbol{z}, t)$$

as

$$\begin{aligned}\frac{d}{dt}\left(V(\boldsymbol{z},t)\right) &\leq -\alpha\,\widehat{\boldsymbol{\eta}}^T M_{\mathcal{S}}^{\boldsymbol{z}\,T} M_{\mathcal{S}}^{\boldsymbol{z}}\,\widehat{\boldsymbol{\eta}} + M_{\mathcal{S}}^{t}\,\boldsymbol{\eta} \\ &\leq -\alpha\,\rho^2\,\|\widehat{\boldsymbol{\eta}}\|_\infty^2 + \|\boldsymbol{\beta}\|\|\mathcal{S}\|\|\boldsymbol{\eta}_1\| + \|\boldsymbol{\psi}\|\|\boldsymbol{\eta}_2\| \\ &\leq -\alpha\,\rho^2 + \sqrt{n}\,(\kappa_b+\kappa_\varphi)\,, \end{aligned} \tag{25}$$

where ρ defined as in Eq. (19) is positive in virtue of Lemma 1, κ_b, κ_φ are the positive bounds on the possible values of the derivatives (whose knowledge is required to set an adequate gain α) of the signals $\boldsymbol{b}(t)$ and $\boldsymbol{\varphi}(t)$, respectively, $\|\mathcal{S}\| \leq 1$ from its structure detailed in Eq. (6), and we used the fact that whenever $\boldsymbol{h} \neq \mathbf{0}_{n+m}$, i.e., $\boldsymbol{z} \neq \boldsymbol{z}^*(t)$, $\widehat{\boldsymbol{\eta}}$ has at least one component with $|\widehat{\eta}_i| = 1$ while in general all other components satisfy $|\widehat{\eta}_j| \leq 1$ and thus it holds $\|\widehat{\boldsymbol{\eta}}\|_\infty^2 = 1$. At this point, by choosing α according to Eq. (18) the following holds true

$$\frac{d}{dt}\left(V(\boldsymbol{z},t)\right) < -\epsilon < 0,$$

thus from Theorem 2, noting that in our case $\mathcal{C}(t) = \{\boldsymbol{z}^*(t)\}$ is compact by definition (i.e., by Proposition 2 $\boldsymbol{z}^*(t)$ is unique and thus $\{\boldsymbol{z}^*(t)\}$ is compact at each t), it follows that $V(\boldsymbol{z},t)$ converges to 0 in finite-time and $\|\boldsymbol{z}(t) - \boldsymbol{z}^*(t)\|$ reaches zero in finite-time too (and remains equal to zero). A characterization on the convergence time $T(\epsilon)$ can be computed as

$$\begin{aligned} V(\boldsymbol{z}(t),t) &= V(\boldsymbol{z}(0),0) + \int_0^t \frac{d}{dt}(V(\boldsymbol{z}(\tau),\tau))d\tau \\ &< V(\boldsymbol{z}(0),0) - \int_0^t \epsilon\, d\tau = V(\boldsymbol{z}(0),0) - \epsilon\, t. \end{aligned}$$

An upper bound on the convergence time $T(\epsilon) \leq T_{\max}$ is

$$T_{\max} = \frac{1}{\epsilon}\, V(\boldsymbol{z}(0),0).$$

The result follows. ∎

Having proven finite-time convergence of the function $\boldsymbol{h}(\boldsymbol{z},t)$ to the origin, we can now prove that Problem 2 is solved in finite-time as well.

Corollary 1: Let the conditions of Theorem 4 hold. Then Problem 2 is solved for all $t \geq T(\epsilon)$.

Proof: The results follows from the application of Theorem 4. In particular, for $\boldsymbol{z}(t) = \boldsymbol{z}^*(t)$ it holds $\boldsymbol{h}(\boldsymbol{z}^*(t),t) = \mathbf{0}_{n+m}$ and thus the optimality condition in Eq. (3) is satisfied, proving that $\boldsymbol{x}(t) = \boldsymbol{x}^*(t)$ for each time instant $t \geq T(\epsilon)$. ∎

B. Choice of the gain α

According to Theorem 4, for each given $T_{\max}$ there is a sufficiently large choice of α such that the system achieves convergence in finite-time that is upper bounded by $T_{\max}$. However, choosing α via Eq. (18) may look very hard at first glance, since computing ρ requires to evaluate an infinity of matrices $M_{\mathcal{S}} \in \mathbb{M}$.

In this subsection, in order to simplify this endeavor, we first provide a practical way to choose α based on the information available to the agents as per Assumption 7, without the need to consider the different matrices in $\mathbb{M}$; then, with the aim to further simplify the task of choosing α, we show how to lift the assumption that the agents need to know $\sigma_{\min}(C)$.

In the following proposition we derive a lower bound for ρ, trading easiness of computation for: i) a slight increase in the magnitude of the resulting gain α and ii) an additional assumption which, as it will be shown later in Remark 4, can always be satisfied by considering an equivalent formulation of the problem at hand.

Proposition 3: Let Assumptions 2, 3 and 6 hold true. Then, we have that $\rho \geq \overline{\rho} > 0$, where ρ is defined as in Eq. (19) whereas $\overline{\rho}$ is

$$\overline{\rho} = \frac{\sigma_{\min}(C)}{\sqrt{n+m+\|A\|_F^2}},$$

with C defined in Eq. (2).

Proof: In order to prove the result, let us introduce the matrices $C_{\mathcal{S}}, X \in \mathbb{R}^{n+m\times n+m}$ defined as

$$C_{\mathcal{S}} = \begin{bmatrix} I_m - \mathcal{S} & -A \\ A^T & Q - A^T A \end{bmatrix}, \quad X = \begin{bmatrix} I_m & 0_{m\times n} \\ A^T & I_n \end{bmatrix}. \tag{26}$$

By construction, it follows that $C_{\mathcal{S}} = X\, M_{\mathcal{S}}$. Therefore, as shown in [38], it holds $\sigma_{\min}(C_{\mathcal{S}}) \leq \|X\|\sigma_{\min}(M_{\mathcal{S}})$, and, thus,

$$\sigma_{\min}(M_{\mathcal{S}}) \geq \sigma_{\min}(C_{\mathcal{S}})\|X\|^{-1}.$$

At this point, by resorting to the properties of the Shur complement of block matrices, since $\mathcal{S}_{ii} \in [0,1]$ for all i and due to the fact, by Assumption 6, $Q - A^T A$ is positive definite, we have that

$$\begin{aligned} \mathtt{det}(C_{\mathcal{S}}) &= \underbrace{\mathtt{det}(Q - A^T A)}_{>0} \times \\ &\times \mathtt{det}\left(I_m - \mathcal{S} + \underbrace{A(Q-A^T A)^{-1}A^T}_{R}\right) \\ &= \mathtt{det}(Q - A^T A)\prod_{i=1}^{m}\left(\underbrace{1 - \mathcal{S}_{ii}}_{\geq 0} + \lambda_i(R)\right) \\ &\geq \mathtt{det}(Q - A^T A)\prod_{i=1}^{m}\lambda_i(R) \\ &= \mathtt{det}(Q - A^T A)\mathtt{det}\,(R) \\ &= \mathtt{det}(Q - A^T A)\mathtt{det}\left(A(Q - A^T A)^{-1}A^T\right) \\ &= \mathtt{det}(C), \end{aligned}$$

where the last equation holds in virtue of the properties of the Schur complement [35] $A(Q - A^T A)^{-1}A^T$ of C. To conclude our proof we observe that, since A is full row rank and $Q - A^T A$ is positive definite, also $A(Q - A^T A)^{-1}A^T$ is positive definite: in fact, for all $\boldsymbol{x} \in \mathbb{R}^m$ with $\boldsymbol{x} \neq \mathbf{0}_m$ we have $\boldsymbol{y} = A^T\boldsymbol{x} \neq \mathbf{0}_n$ and thus

$$\boldsymbol{x}^T A(Q - A^T A)^{-1}A^T\boldsymbol{x} = \boldsymbol{y}^T(Q - A^T A)^{-1}\boldsymbol{y} > 0,$$

where the latter inequality holds since $\boldsymbol{y} \neq \mathbf{0}_n$ and since $Q - A^T A$ is positive definite (which implies that also $(Q - A^T A)^{-1}$ is positive definite). Therefore, we have that $\sigma_{\min}(M_{\mathcal{S}}) \geq \sigma_{\min}(C)/\|X\| > 0$. The proof is complete since

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2023.3272220

it holds

$$\|X\| \leq \|X\|_F = \sqrt{\|I_n\|_F^2 + \|I_m\|_F^2 + \|A\|_F^2} = \sqrt{n + m + \|A\|_F^2}.$$

■

We point out the above lower bound is remarkably easier to compute than ρ, as it requires only knowledge of just $\sigma_{\min}(C), n, m$ and $\|A\|_F$ without the need to inspect the set of all $M_{\mathcal{S}} \in \mathbb{M}$. Interestingly, as discussed later in this section, $\|A\|_F$ can be computed in a distributed way, based on the information available to the agents as per Assumption 7.

Remark 4: Proposition 3 requires $Q - A^T A$ to be positive definite. Since for symmetric matrices $U, V \in \mathbb{R}^{n \times n}$, it is well known [39] that $\lambda_{\min}(U + V) \geq \lambda_{\min}(U) - \|V\|$, we have that, scaling Q by $\beta > \|A\|_F^2 / \lambda_{\min}(Q)$, it holds

$$\begin{aligned}\lambda_{\min}\left(\beta Q - A^T A\right) &\geq \beta \lambda_{\min}(Q) - \|A\|^2 \\ &> \|A\|_F^2 - \|A\|^2 \geq \|A\|_F^2 - \|A\|_F^2 = 0,\end{aligned}$$

where we used the property that $\|A\|_F \geq \|A\|$. Notably, an optimization problem is equivalent under positive scaling of the objective function; hence given Q, A we can consider the equivalent formulation

$$\boldsymbol{x}^*(t) = \arg\min_{\boldsymbol{x} \in \mathbb{R}^n} \beta \left(\frac{1}{2}\boldsymbol{x}^T Q \boldsymbol{x} + \boldsymbol{\varphi}^T(t)\boldsymbol{x}\right) \text{ s.t. } A\boldsymbol{x} \geq \boldsymbol{b}(t),$$

which satisfies Assumption 6.

Interestingly, the Frobenius norm of A can be computed in a distributed way in finite time. In particular, assuming each agent i knows the entries A_{ij} corresponding to its neighbors (if an agent is not in charge of a constraint, it simply assumes $A_{ij} = 0$ for all its neighbors), it is sufficient to run a finite-time distributed average consensus procedure [40] with initial condition $y_i(0) = \sum_{j \in \mathcal{N}_i} A_{ij}^2$, which converges to $\overline{y} = \frac{1}{n}\sum_i y_i(0) = \frac{1}{n}\|A\|_F^2$. Then, based on knowledge on n, each agent can compute $\|A\|_F = \sqrt{n\overline{y}}$. Similarly, the agents are able to compute $\|Q\|_F$ in a distributed way. Therefore, $\|A\|_F, \|Q\|_F$ can be computed during an initialization phase.

To conclude the section, with the aim to further reduce the amount of information required for the agents in order to choose the gain α, let us discuss a way to relax the assumption that the agents need to know $\sigma_{\min}(C)$. Specifically, we now show that there is a positive lower bound of $\sigma_{\min}(C)$ which is based on $\|A\|_F, \|Q\|_F, n, m, \sigma_{min}(A), \lambda_{\min}(Q)$.

Proposition 4: Let Assumptions 1–6 hold true. Then, it holds

$$\sigma_{\min}(C) \geq \left(\lambda_{\min}(Q) - \|A\|_F^2\right)^{n-m} \sigma_{\min}(A)^{2m} \times \left(\frac{n+m-1}{\|Q\|_F^2 + 4\|A\|_F^2}\right)^{\frac{n+m-1}{2}}.$$

Proof: In order to prove the result we resort to the lower bound in [41], where it is shown that, for a given nonsingular square matrix $U \in \mathbb{R}^{n \times n}$ it holds

$$\sigma_{\min}(U) \geq |\mathtt{det}(U)| \left(\frac{n-1}{\|U\|_F^2}\right)^{\frac{n-1}{2}} > 0,$$

from which we have that

$$\sigma_{\min}(C) \geq |\mathtt{det}(C)| \left(\frac{n+m-1}{\|C\|_F^2}\right)^{\frac{n+m-1}{2}} > 0. \tag{27}$$

Notice that, since $Q - A^T A$ is nonsingular, by resorting to the properties of the Schur complement of block matrices [35], we have that

$$\begin{aligned}\mathtt{det}(C) &= \mathtt{det}(Q - A^T A)\mathtt{det}(A(Q - A^T A)^{-1}A^T) \\ &= \prod_{i=1}^{n} \lambda_i(Q - A^T A) \prod_{j=1}^{m} \sigma_i^2((Q - A^T A)^{-1/2}A^T),\end{aligned}$$

where we used the well known properties that, for $Y \in \mathbb{R}^{n \times n}$ and $G \in \mathbb{R}^{m \times n}$ with $m \leq n$, it holds $\mathtt{det}(Y) = \prod_{i=1}^{n} \lambda_i(Y)$ and $\mathtt{det}(GG^T) = \prod_{i=1}^{m} \sigma_i^2(G)$ (see, for instance, [42]). At this point, noting that $\lambda_i(Q - A^T A) \geq \lambda_{\min}(Q - A^T A)$ and that [43]

$$\sigma_i(UV) \geq \sigma_{\min}(U)\sigma_{\min}(V),$$

we have that

$$\begin{aligned}\mathtt{det}(C) &\geq \lambda_{\min}^n(Q - A^T A)\underbrace{\sigma_{\min}^{2m}((Q - A^T A)^{-1/2})}_{=\lambda_{\min}^{-m}(Q - A^T A)}\sigma_{\min}^{2m}(A) \\ &= \lambda_{\min}^{n-m}(Q - A^T A)\sigma_{\min}^{2m}(A) \\ &\geq \left(\lambda_{\min}(Q) - \|A\|^2\right)^{n-m} \sigma_{\min}^{2m}(A) \\ &\geq \left(\lambda_{\min}(Q) - \|A\|_F^2\right)^{n-m} \sigma_{\min}^{2m}(A).\end{aligned} \tag{28}$$

Moreover, it holds

$$\|C\|_F^2 = 2\|A\|_F^2 + \|Q - A^T A\|_F^2 \leq 4\|A\|_F^2 + \|Q\|_F^2. \tag{29}$$

The proof follows by plugging Eqs. (28) and (29) into Eq. (27).

■

VI. SIMULATIONS

For the numerical validation of the proposed protocol, we considered a multi-agent system with $n = 10$ agents interacting over an undirected graph with $|\mathcal{E}| = 14$ edges. Moreover, we consider uniformly random $Q \in \mathbb{R}^{10 \times 10}$ and $A \in \mathbb{R}^{7 \times 10}$, i.e., $|\mathcal{H}| = 7$. The time-varying vectors $\boldsymbol{\varphi}(t)$ and $\boldsymbol{b}(t)$ are depicted in Fig. 1d and satisfy Assumption 4 with $\kappa_\varphi = 1.755$ and $\kappa_b = 1.5$. The considered matrices A, Q are reported in the following

$$A = \begin{bmatrix} 4.73 & -0.06 & 0.36 & 0 & 0 & 0 & 0 & -0.32 & 0 & 0 \\ 0.79 & 4.02 & 0 & 0 & 0 & 0 & 0.47 & 0 & 0 & 0.59 \\ -0.15 & 0 & 4.64 & 0 & 0 & 0 & 0 & 0 & 0.26 & 0.08 \\ 0 & 0 & 0 & 3.63 & 0 & 0.08 & 0.70 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.74 & 0.01 & 0 & 0 & -0.19 & 0 \\ 0 & 0 & 0 & 0.01 & 0.20 & 4.6 & 0.45 & 0 & -0.08 & 0 \\ 0 & -0.08 & 0 & 0.01 & 0 & 0.29 & 3.94 & 0 & 0 & 0 \end{bmatrix}$$

$$Q = \begin{bmatrix} 7.02 & 1.41 & 1.02 & 0 & 0 & 0 & 0 & 0.13 & 0 & 0 \\ 1.41 & 6.54 & 0 & 0 & 0 & 0 & 0.24 & 0 & 0 & 0.63 \\ 1.02 & 0 & 6.86 & 0 & 0 & 0 & 0 & 0 & 1.14 & 0.14 \\ 0 & 0 & 0 & 4.62 & 0 & 0.55 & 0.83 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5.26 & 0.74 & 0 & 0 & 0.52 & 0 \\ 0 & 0 & 0 & 0.55 & 0.74 & 7.51 & 1.79 & 0 & 1.01 & 0 \\ 0 & 0.24 & 0 & 0.83 & 0 & 1.79 & 7. & 0 & 0 & 0 \\ 0.13 & 0 & 0 & 0 & 0 & 0 & 0 & 5.25 & 0.62 & 0 \\ 0 & 0 & 1.14 & 0 & 0.52 & 1.01 & 0 & 0.62 & 8.41 & 0 \\ 0 & 0.63 & 0.14 & 0 & 0 & 0 & 0 & 0 & 0 & 7.03 \end{bmatrix}.$$

$$\boldsymbol{\varphi}(0) = [0.93, 0.92, -0.5, 0.05, -0.49, -0.6, 0.08, -0.1, -0.02, -0.19]^T,$$

$$\boldsymbol{b}(0)=[-0.48, 0.46, 0.18, 0.72, -0.33, 0.17, 1, 0.76, 0.36, 0.14]^T.$$

In order to correctly tune the gain α the results of Proposition 3 can be exploited. However, since Q and A do not satisfy the condition on the positive definiteness of $Q - A^T A$, the method explained in Remark 4 can be applied to scale the matrix Q such that $\beta Q - A^T A$ is positive definite. In particular, for this example, we chose $\beta = 1.83$.

The proposed algorithm was implemented in discrete time using the forward Euler method with sampling time $\tau = 10^{-8}$. Agents implement the local interaction rule given in Eq. (7) with gain $\alpha = 1706$ according to the results of Theorem 4 and Proposition 3. Notably, the estimated bound for the minimum singular value of the matrix $M_{\mathcal{S}}$ obtained exploiting the results of Proposition 3, i.e., $\overline{\rho} = 0.0781$, is not far from the best value numerically obtained via a Monte Carlo simulation campaign featuring 10^6 trials. The minimum singular value of the matrix $M_{\mathcal{S}}$, according to Monte Carlo evaluation corresponds to a value of $\rho = 0.323$. The ratio between the bound computed utilizing Proposition 3 and the Monte Carlo minimum value is $\rho/\overline{\rho} = 4.1379$. This implies that our control gain α is about 17 times larger than the minimum value required by the conditions of Theorem 4.

We remind the reader that the proposed algorithm requires 2-hop information which can be estimated in finite-time implementing a 2-hop distributed observer as the one given in [24] which requires only 1-hop information to work. For the sake of simplicity and with no lack of generality, we assume that at the initial time t_0 the local observer tracking error has already reached zero, i.e., all the agents for time $t \geq t_0$ possess the 2-hop state information required to implement the proposed distributed strategy, and we focus only on illustrating the properties of the proposed algorithm.

Fig. 1a and Fig. 1b show the evolution of $\boldsymbol{x}(t)$ and $\boldsymbol{\zeta}(t)$, respectively. Fig. 1c depicts the evolution of the Lyapunov-like function $V(\boldsymbol{z}, t)$ introduced in Eq. (20) with a small frame in the top right side of the picture showing a detail of its convergence during the first 0.0025 seconds of the simulation. Fig. 1d shows the evolution of the time-varying vectors $\boldsymbol{\varphi}(t)$ and $\boldsymbol{b}(t)$. Fig. 1e shows the evolution of the time-varying constraint $A\boldsymbol{x}(t) - \boldsymbol{b}(t) \geq 0$. Finally, Fig. 1f show the evolution of the errors between $\boldsymbol{x}(t)$, $\boldsymbol{\zeta}(t)$ and the optimal solution $\boldsymbol{x}^*(t)$, $\boldsymbol{\zeta}^*(t)$ obtained via a centralized solver where in the top right side a detail of their convergence during the first 1.5 seconds of the simulation is shown. According to the figures, the proposed algorithm is able to track the global optimal solution of the time-varying optimization problem as expected from the results of Corollary 1.

VII. Conclusions and Future Work

In this paper we considered a class of quadratic optimization problems with time-varying linear objective term and time-varying linear constraints with the same sparsity pattern as the static graph encoding the undirected network topology over which the multi-agent systems interacts. Our contribution is twofold. First we exploited the Karush-Kuhn-Tucker conditions to derive a necessary and sufficient global optimality condition for the frozen-time problem. Since the derived optimality condition is in the form of a system of nonsmooth equations, we developed a nonsmooth distributed algorithm to achieve finite-time convergence and track to the optimal time-varying solution. Furthermore, we derived a lower bound for the minimum singular value of the family of matrices $M_{\mathcal{S}} \in \mathbb{M}$, providing a method to practically compute the gain α required to solve optimization problem. Future work will aim to consider more general time-varying problems, e.g., quadratic problems with time-varying Hessian and constraint matrix; in this context a challenge to overcome is that the time-variability of the aforementioned matrices would not be dominated by a static gain, thus calling for an adaptive gain approach. Furthermore, the introduction of adaptive gains will also allow to lift the requirements on the information that must be available to the nodes, e.g., the number of agents, the bounds of the derivatives of the time-varying signals, and $\sigma_{\min}(C)$.

References

[1] M. Akbari, B. Gharesifard, and T. Linder, "Distributed online convex optimization on time-varying directed graphs," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 3, pp. 417–428, 2015.

[2] X. Wu and J. Lu, "Fenchel dual gradient methods for distributed convex optimization over time-varying networks," *IEEE Transactions on Automatic Control*, vol. 64, no. 11, pp. 4629–4636, 2019.

[3] A. Rogozin, C. Uribe, A. Gasnikov, N. Malkovskii, and A. Nedich, "Optimal distributed convex optimization on slowly time-varying graphs," *IEEE Transactions on Control of Network Systems*, 2019.

[4] X. Wu and J. Lu, "Distributed optimization over time-varying networks with minimal connectivity," *IEEE Control Systems Letters*, vol. 4, no. 3, pp. 536–541, 2020.

[5] S. Rahili and W. Ren, "Distributed continuous-time convex optimization with time-varying cost functions," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1590–1605, 2016.

[6] C. Sun, M. Ye, and G. Hu, "Distributed time-varying quadratic optimization for multiple agents under undirected graphs," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, 2017.

[7] S. Shahrampour and A. Jadbabaie, "Distributed online optimization in dynamic environments using mirror descent," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 714–725, 2017.

[8] L. Bai, C. Sun, Z. Feng, and G. Hu, "Distributed continuous-time resource allocation with time-varying resources under quadratic cost functions," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 823–828.

[9] M. Vaquero and J. Cortés, "Distributed augmentation-regularization for robust online convex optimization," *IFAC-PapersOnLine*, vol. 51, no. 23, pp. 230–235, 2018.

[10] A. Simonetto, "Dual prediction–correction methods for linearly constrained time-varying convex programs," *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3355–3361, 2018.

[11] M. Hosseinzadeh, E. Garone, and L. Schenato, "A distributed method for linear programming problems with box constraints and time varying inequalities," *IEEE Control Systems Letters*, vol. 3, no. 2, 2018.

[12] S. Sun and W. Ren, "Distributed continuous-time optimization with time-varying objective functions and inequality constraints," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 5622–5627.

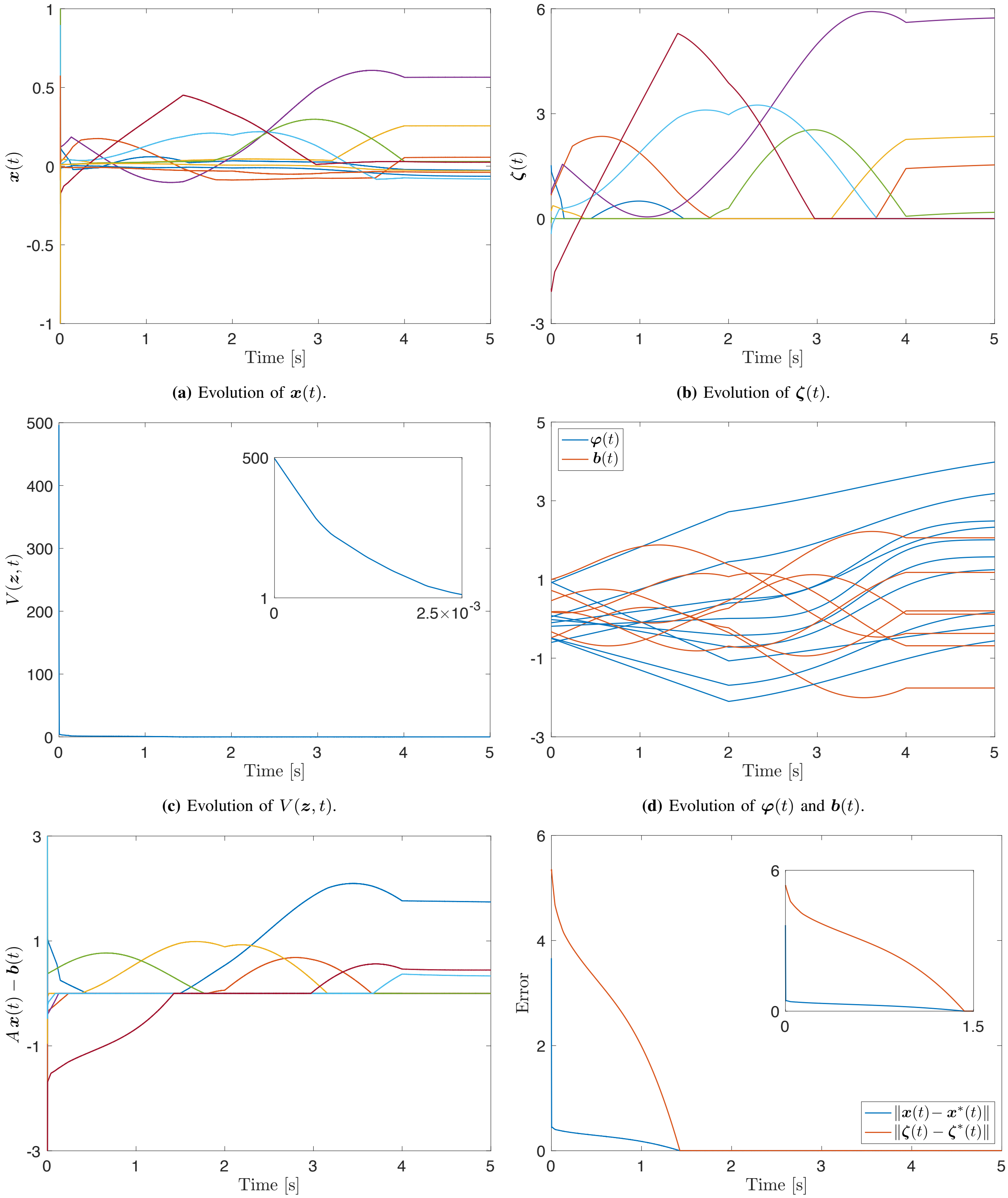
[13] M. Santilli, G. Oliva, and A. Gasparri, "Distributed finite-time algorithm for a class of quadratic optimization problems with time-varying linear constraints," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 4380–4386.

[14] S. Sun, J. Xu, and W. Ren, "Distributed continuous-time algorithms for time-varying constrained convex optimization," *IEEE Transactions on Automatic Control*, 2022.

[15] G. Oliva, A. I. Rikos, C. N. Hadjicostis, and A. Gasparri, "Distributed flow network balancing with minimal effort," *IEEE Transactions on Automatic Control*, vol. 64, no. 9, pp. 3529–3543, 2019.

[16] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2023.3272220



(a) Evolution of $\boldsymbol{x}(t)$.

(b) Evolution of $\boldsymbol{\zeta}(t)$.

(c) Evolution of $V(\boldsymbol{z},t)$.

(d) Evolution of $\boldsymbol{\varphi}(t)$ and $\boldsymbol{b}(t)$.

(e) Evolution of $A\boldsymbol{x}(t)-\boldsymbol{b}(t)$.

(f) Evolution of the errors $\|\boldsymbol{x}(t)-\boldsymbol{x}^*(t)\|$ and $\|\boldsymbol{\zeta}(t)-\boldsymbol{\zeta}^*(t)\|$.

Fig. 1**:** Results of the numerical simulations involving a team of $n=10$ agents solving distributively Problem 1 with $m=7$ in finite-time.

[17] X. You, X. Li, Y. Xu, H. Feng, J. Zhao, and H. Yan, "Toward packet routing with fully distributed multiagent deep reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020.
[18] A. Nedić and J. Liu, "Distributed optimization for control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, 2018.
[19] B. Ning, Q.-L. Han, and Z. Zuo, "Distributed optimization for multi-agent systems: An edge-based fixed-time consensus approach," *IEEE transactions on cybernetics*, vol. 49, no. 1, pp. 122–132, 2017.
[20] M. Fazlyab, S. Paternain, V. M. Preciado, and A. Ribeiro, "Prediction-correction interior-point method for time-varying convex optimization," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, 2018.
[21] M. Colombino, E. Dall'Anese, and A. Bernstein, "Online optimization as a feedback controller: Stability and tracking," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 422–432, 2020.
[22] A. Simonetto and E. Dall'Anese, "Prediction-correction algorithms for time-varying constrained optimization," *IEEE Transactions on Signal Processing*, vol. 65, no. 20, pp. 5481–5494, 2017.
[23] M. Maros and J. Jaldén, "A geometrically converging dual method for distributed optimization over time-varying graphs," *IEEE Transactions on Automatic Control*, vol. 66, no. 6, pp. 2465–2479, 2021.
[24] A. Gasparri and A. Marino, "A distributed framework for k-hop control strategies in large-scale networks based on local interactions," *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 1825–1840, 2020.
[25] C. Godsil and G. F. Royle, *Algebraic graph theory*. Springer Science & Business Media, 2013, vol. 207.
[26] B. Paden and S. Sastry, "A calculus for computing Filippov's differential inclusion with application to the variable structure control of robot manipulators," *IEEE Transactions on Circuits and Systems*, vol. 34, no. 1, pp. 73–82, Jan 1987.
[27] D. Shevitz and B. Paden, "Lyapunov stability theory of nonsmooth systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 9, pp. 1910–1914, Sep 1994.
[28] J. Cortés, "Discontinuous dynamical systems," *IEEE Control Systems*, vol. 28, no. 3, pp. 36–73, June 2008.
[29] F. Clarke, *Optimization and Nonsmooth Analysis*. Wiley & Sons, New York, 1983.
[30] N. Basilico and F. Amigoni, "Exploration strategies based on multi-criteria decision making for searching environments in rescue operations," *Autonomous Robots*, vol. 31, pp. 401–417, 2011.
[31] A. G. Azar, H. Nazaripouya, B. Khaki, C.-C. Chu, R. Gadh, and R. H. Jacobsen, "A non-cooperative framework for coordinating a neighborhood of distributed prosumers," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2523–2534, 2018.
[32] W. I. Zangwill, *Nonlinear Programming: a Unified Approach*. Prentice-Hall, 1969.
[33] Wenyu Sun and Ya-Xiang Yuan, *Theory of Constrained Optimization*. Boston, MA: Springer US, 2006, pp. 385–410.
[34] J. Kyparisis, "On uniqueness of Kuhn-Tucker multipliers in nonlinear programming," *Mathematical Programming*, vol. 32, no. 2, 1985.
[35] F. Zhang, *The Schur complement and its applications*. Springer Science & Business Media, 2006, vol. 4.
[36] G. T. Gilbert, "Positive definite matrices and Sylvester's Criterion," *The American Mathematical Monthly*, vol. 98, no. 1, pp. 44–46, 1991.
[37] O. Kosheleva and P. G. Vroegindeweij, "When is the product of intervals also an interval?" *Reliable Computing*, vol. 4, no. 2, pp. 179–190, 1998.
[38] W. Govaerts and J. Pryce, "A singular value inequality for block matrices," *Linear algebra and its applications*, vol. 125, 1989.
[39] H. Weyl, "Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung)," *Mathematische Annalen*, vol. 71, no. 4, pp. 441–479, 1912.
[40] L. Wang and F. Xiao, "Finite-time consensus problems for networks of dynamic agents," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 950–955, 2010.
[41] A. D. Güngör, "Erratum to "an upper bound for the condition number of a matrix in spectral norm"[j. comput. appl. math. 143 (2002) 141–144]," *Journal of Computational and Applied Mathematics*, vol. 234, no. 1, p. 316, 2010.
[42] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.
[43] S. Loyka and C. D. Charalambous, "Novel matrix singular value inequalities and their applications to uncertain mimo channels," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6623–6634, 2015.

Matteo Santilli Matteo Santilli received the B.Sc. and M.Sc. degrees (cum laude) in computer science and automation engineering from Roma Tre University, Italy, in 2015 and 2017, respectively. In 2021 he received the Ph.D. degree in automation and computer science with the Department of Engineering at Roma Tre University. His research interests include nonlinear multi-agent systems, resilient control problems, and distributed optimization.

Antonio Furchí received the M.sc in computer science and automation engineering from the University Roma Tre of Rome, Italy, in 2020 and is currently pursuing a Ph.D. degree in computer science and automation at University Roma Tre of Rome, Italy, under the supervision of prof. Gasparri. His main research interests include distributed systems, distributed optimization, and precision agriculture.

Gabriele Oliva (M'11, SM'19) received the M.sc and the Ph.D. degrees in computer science and automation engineering from the University Roma Tre of Rome, Italy, in 2008 and 2012, respectively. He is currently an Associate Professor in automatic control with the University Campus Bio-Medico of Rome, Italy, where he directs the Complex Systems & Security Laboratory (CoserityLab). Since 2019, he serves as an Associate Editor for the Conference Editorial Board of the IEEE Control Systems Society. Moreover, since 2020, he serves as an Academic Editor for the journal PLOS ONE on subject areas such as Systems Science, Optimization and Decision Theory. Finally, since 2022 he is an Associate Editor for the IEEE Control Systems Letters Journal. His main research interests include distributed multi-agent systems, optimization, decision-making and critical infrastructure protection.

Andrea Gasparri (M'09, SM'19) received the Laurea (cum laude) degree in computer science and the Ph.D. degree in computer science and automation from Roma Tre University, Rome, Italy, in 2004 and 2008, respectively. He is currently a Professor with the Department of Civil, Computer Science and Aeronautical Technologies Engineering, Roma Tre University. His research interests include robotics, sensor networks, networked multiagent systems, and precision agriculture. Dr. Gasparri has been a member of the Steering Committee for the IEEE RAS Technical Committee on Multirobot Systems since 2014, the IEEE CSS Technical Committee on Networks and Communications since 2015, and the IEEE Technical Committee on Agricultural Robotics since 2021. From 2017 to 2021, he was an Associate Editor for the IEEE Transactions on Cybernetics. Since 2021, he has been an Associate Editor for the IEEE Transactions in Control of Network Systems. He was the recipient of the Italian Grant FIRB Futuro in Ricerca 2008 for the project Networked Collaborative Team of Autonomous Robots funded by the Italian Ministry of Research and Education. He was the Coordinator of the project "PANTHEON" focused on robotics for precision farming supported by the European Community within the H2020 framework (under grant agreement number 774571) and is currently the Coordinator of the project "CANOPIES" focused on the development of a collaborative paradigm for human workers and multirobot teams in precision agriculture systems within the H2020 framework (under grant agreement number 101016906).