A Finite-Time Protocol for Distributed Continuous-Time Optimization of Sum of Locally Coupled Strictly Convex Functions

Matteo Santilli, Alessandro Marino, *IEEE Senior Member* Andrea Gasparri, *IEEE Member*

***Abstract*— In this paper we study a distributed optimization problem for continuous time multi-agent systems. In our setting, the global objective for the multi-agent system is to minimize the sum of locally coupled strictly convex cost functions. Notably, this class of optimization objectives can be used to encode several important problems such as distributed estimation. For this problem setting, we propose a distributed signed gradient descent algorithm, which relies on local observers to retrieve 2-hop state information that are required to compute the descent direction. Adaptive gains for the local observer are introduced to render the convergence independent from: i) the structure of the network topology and ii) the local gains of the per-agent signed gradient-descent update law. The finite-time convergence of the local observer and of the proposed signed gradient descent method is demonstrated. Numerical simulations involving a distributed weighted least-square (WLS) estimation problem, with the aim of identifying in the context of an advanced water management system for precision-farming the soil thermal properties in a large-scale hazelnut orchard, have been proposed to corroborate the theoretical findings.**

I. Introduction

Distributed optimization has been a very active research field over the last few decades. As a matter of fact, current technological advances in the design and realization of micro-electro-mechanical systems have motivated the development of distributed methodologies that fit well into the multi-agent coordination paradigm for cooperative optimization [1]. In this work, we focus our attention on a popular family of optimization algorithms, referred to as *coordinate gradient-descent methods*, which substantially involves a local update of a subset of the decision variable to globally reach an optimum under reasonable assumption of the cost function [2]–[14]. This collaborative optimization paradigm offers several interesting advantages, such a substantial reduction of the computational effort, due to the fact that every agent is only required to steer its local state towards the optimal solution according to its partial derivative.

Nowadays, a great amount of coordinate gradient-descent methods can be found at the state of the art. A typical optimization formulation involves a sum of locally convex functions, where the objective is to steer a common variable towards the optimal value. Although the majority of the works are focused on distributed discrete-time protocols [2]–[7], more recently, as pointed out in [8], continuous-time multi-agent approaches have been applied to distributed optimization problems as promising and useful techniques [8]–[14].

In our setting, we consider an alternative problem formulation which instead involves the sum of locally *coupled* convex functions. As pointed out in several works [4], [5], [15], this setting is fundamentally different from the previous one as the pairwise state coupling between the agents naturally leads to a coupling in the definition of the gradient components, thus rendering the design of fully distributed algorithms more challenging. A typical solution at the state of the art to solve this problem is to rely on discrete-time implementations where the 2-hop information is recovered by 2-rounds of communications for each iteration of the optimization algorithm [4], [15]. More recently, in [5] a solution amenable for distribution which, under a suitable condition on the step size, is provable locally resilient to communication failures at the price of reduced converge speed has been proposed.

In this work, we propose a novel alternative continuous-time solution which does not require 2-rounds of communication but, instead, exploits an estimate of the 2-hop neighborhood provided by local observers along with the (one-time) exchange of the structure of the local components of the gradients for pairs of neighboring agents. The major novelty of this work, which is based on the general k-hop graph-based local observer proposed in [16], is the introduction of adaptive gains to render the convergence independent from: i) the structure of the network topology and ii) the local gains of the signed gradient-descent of each agent. In addition, we demonstrate the finite-time convergence of the local observer with adaptive gains and the finite-time convergence of the signed gradient descent method where the outcome of the local observer is used to compute the descent direction.

II. Preliminaries

A. Agent and Network Modeling

Let us consider a system composed of n agents where each agent i satisfies the continuous-time single integrator dynamics

$$\dot{x}_i = u_i, \quad u_i \in \mathbb{R}^d \tag{1}$$

and let us denote with $\boldsymbol{x}$ the stacked vector collecting the state of all agents, that is $\boldsymbol{x} = \left[x_1^T, \dots, x_n^T\right]^T \in \mathbb{R}^{nd}$.

Assume the interaction among the agents to be described by an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1, \dots, n\}$ is the set of agents and $\mathcal{E} = \{(i, j)\}$ is the set of pairwise interactions among agents i and j. Note that since the graph is undirected the existence of an edge (i, j) implies that an edge (j, i) exists as well. Thus, the indexes i and j can be arbitrarily switched. Let us denote with $\mathcal{N}_i^1 = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$ the 1-hop neighborhood of agent i. In addition, let us denote with

M. Santilli and A. Gasparri are with the Department of Engineering, Roma Tre University, Rome, Italy, 00146 e-mail: (see http://gasparri.dia.uniroma3.it).

A. Marino is with the Department of Information and Electrical Engineering and Applied Mathematics, University of Salerno, Salerno, Italy, 84084 email: (see http://docenti.unisa.it/alessandro.marino).

This work has been supported by the European Commission under the Grant Agreement number 774571 (Project PANTHEON – "Precision farming of hazelnut orchards").

$\mathcal{N}_i^2 = \{j \in \mathcal{V} : (i,j) \in \mathcal{E} \vee (k,j) \in \mathcal{E},\, k \in \mathcal{N}_i\}$ the 2-hop neighborhood of agent i with no repetitions, that is only one copy of the index of common agents is included in such neighborhood. Furthermore, let us denote with $\bar{\mathcal{N}}_i^\alpha = \mathcal{N}_i^\alpha \cup \{i\}$, that is the augmented α-hop neighborhood of agent i, including agent i itself, with $\alpha \in \{1,\, 2\}$. Note that, the knowledge of the 2-hop neighborhood can be still assumed local as it can be provided by 1-hop agents. In addition, we point out that we do not require any knowledge of the network topology underlying the 2-hop neighborhood.

Let us now introduce the (sub)-graph $\mathcal{G}_i^2$ obtained by reducing $\mathcal{G}$ with respect to $\bar{\mathcal{N}}_i^2$, that is $\mathcal{G}_i^2 = \{\mathcal{V}_i^2, \mathcal{E}_i^2\}$ where $\mathcal{V}_i^2 = \bar{\mathcal{N}}_i^2$ and $\mathcal{E}_i^2 \subseteq \mathcal{E} : (p,q) \in \mathcal{E}_i^2 \leftrightarrow \{p,\, q\} \in \mathcal{V}_i^2$. In addition, let us define the Laplacian matrix $\mathcal{L}_i^2$ associated to $\mathcal{G}_i^2$ as $\mathcal{L}_i^2 = \mathcal{D}_i^2 - \mathcal{A}_i^2$ where $\mathcal{D}_i^2$ and $\mathcal{A}_i^2$ are the related degree matrix and adjacency matrix, respectively.

Let us denote with I_d and O_d an identity square matrix and a zero square matrix of dimension d, respectively. Similarly, let us denote with $O_{d\times p}$ a zero matrix with d rows and p columns.

Let us denote with $\boldsymbol{x}_i$ the vector containing the actual state x_j of the agents belonging to the augmented 2-hop neighborhood $\bar{\mathcal{N}}_i^2$ of an agent i, that is $\boldsymbol{x}_i = \bar{P}_i\, \boldsymbol{x} = \left[{x_{f_i(1)}}^T, \ldots, {x_{f_i(n_i)}}^T \right]^T$ with $n_i = \left|\bar{\mathcal{N}}_i^2\right|$ the cardinality of the augmented 2-hop neighborhood, $f_i(\cdot)$ a monotonically increasing direct mapping function defined as $f_i : [1, \ldots, n_i] \rightarrow \bar{\mathcal{N}}_i^2$ that maps the local indexes $\{1,\, \ldots,\, n_i\}$ into the subset of global indexes $\{1,\, \ldots,\, n\}$ belonging to $\bar{\mathcal{N}}_i^2$, and $\bar{P}_i = P_i \otimes I_d$ a selection matrix obtained according to the given graph structure $\mathcal{G}$. In particular, P_i is a $n_i \times n$ binary matrix with $P_i(m,j) = 1$ of $j \in \bar{\mathcal{N}}_i^2$ with $j \geq m$ or $P_i(m,j) = 0$ otherwise. Note that, $\bar{P}_i^T\, \bar{P}_i\, \boldsymbol{x} = \bar{P}_i^T\, \boldsymbol{x}_i$, that is it returns a vector which has the same size as the vector $\boldsymbol{x}$ but it contains only the components of the agents $j \in \bar{\mathcal{N}}_i^2$ by preserving the original index ordering.

Let us also define the inverse mapping $g_i(\cdot)$ of the function $f_i(\cdot)$ as $g_i : \bar{\mathcal{N}}_i^2 \rightarrow [1, \ldots, n_i]$ that maps the subset of global indexes $\{1,\, \ldots,\, n\}$ for which $j \in \bar{\mathcal{N}}_i^2$ in the local indexes $\{1,\, \ldots,\, n_i\}$. Let us now define the state estimate $\hat{\boldsymbol{x}}_i$ of an agent i as $\hat{\boldsymbol{x}}_i = \left[\hat{x}_{i,f_i(1)}^T, \ldots, \hat{x}_{i,f_i(n_i)}^T \right]^T$ where $\hat{x}_{i,f_i(p)}$ is the estimate of the state $x_{f_i(p)}$ carried out by the agent i.

Finally, let us denote with $\otimes$ the Kronecker product, and given a matrix Q with $\lambda_j(Q)$, $\lambda_{min}(Q)$ and $\lambda_{max}(Q)$ the generic, the minimum and the maximum eigenvalues of Q.

B. Problem Setting

Our objective is to design $u_i(t)$ using only 1-hop local interactions for the agent i dynamics given in eq. (1), such that the agents can collaborate to find the optimal state $\boldsymbol{x}^\star = [{x_1^\star}^T, \ldots, {x_n^\star}^T]^T$ that solves the following optimization problem

$$\min_{\boldsymbol{x}} J(\boldsymbol{x}) \triangleq \min_{\boldsymbol{x}} \sum_{i=1}^{n} J_i\left(\{x_j\}_{j \in \bar{\mathcal{N}}_i^1}\right) \tag{2}$$

We reiterate that the global objective given in eq. (2) is the sum of locally coupled cost functions, that is each J_i depends only on information concerning the 1-hop augmented neighborhood $\bar{\mathcal{N}}_i^1$ of agent i.

The following technical assumption concerning the structure of the local objective is now introduced to define uniqueness of the optimal solution $\boldsymbol{x}^\star$.

Assumption 1 (Strict Convexity)**.** The local objective functions $J_i(\boldsymbol{x})$ with $i \in \{1,\, \ldots,\, n\}$ are strictly convex.

Note that, under Assumption 1 if an optimal solution $\boldsymbol{x}^\star$ exists for the optimization problem given in eq. (2), then it is also unique [17], that is

$$\boldsymbol{x}^\star = \operatorname{argmin} J(\boldsymbol{x}) \tag{3}$$

We are now ready to formally state our problem.

Problem 1. Consider an optimization problem of the form (2) where the local objective functions satisfy Assumption 1. Assume an optimal solution $\boldsymbol{x}^\star$ for such optimization problem exists. In addition, consider a multi-agent system with dynamics (1) and design $u_i(t)$ such that $\boldsymbol{x}$ converge to the optimal state $\boldsymbol{x}^\star$ in finite-time, that is there exists $T > 0$ such that

$$\|\boldsymbol{x}(t) - \boldsymbol{x}^\star\| = 0, \quad \forall t \geq T \tag{4}$$

C. Continuous-Time Gradient Descent Optimization

In this section, we briefly overview continuous-time gradient descent optimization for the problem at hand. In particular, it can be proven that the following gradient descent control input $u_i(t)$ solves the optimization problem in eq. (2) under Assumption 1

$$u_i(t) = -\nabla_{x_i} J(\boldsymbol{x}) \tag{5}$$

Notably, for the class of global objectives given in (2), i.e., a sum of locally coupled cost function, as pointed out in several works [4], [5], [15], the gradient descent control input given in (5), cannot be naively implemented in a fully distributed manner. To better understand this point it suffices to further develop the gradient $\nabla_{x_i} J(\boldsymbol{x})$ to notice that it requires 2-hop informations, that is

$$\begin{aligned}\nabla_{x_i} J(\boldsymbol{x}) &= \nabla_{x_i} J_i\left(\{x_j\}_{j \in \bar{\mathcal{N}}_i^1}\right) + \sum_{k \in \mathcal{N}_i^1} \nabla_{x_i} J_k\left(\{x_j\}_{j \in \bar{\mathcal{N}}_k^1}\right) \\ &= \nabla_{x_i} J(\boldsymbol{x}_i)\end{aligned} \tag{6}$$

where the notation $\nabla_{x_i} J(\boldsymbol{x}_i)$ is used to emphasize the dependence of the gradient from the augmented 2-hop neighborhood of agent i.

III. Observer-based Signed Gradient Descent

In this section, we describe the proposed distributed continuous-time signed gradient descent which relies on local observers to compute the descent direction.

Let us consider a multi-agent system composed of n agents and assume each agent i is running the following control law

$$\dot{x}_i = -c_i \operatorname{sign}\left(\nabla_{x_i} J(\hat{\boldsymbol{x}}_i)\right) \tag{7}$$

where c_i is a positive scalar gain, $\operatorname{sign}(\cdot)$ is the component-wise signum function and $\hat{\boldsymbol{x}}_i$ is computed as follows

$$\begin{aligned}\dot{\hat{\boldsymbol{x}}}_i &= \Theta_i \operatorname{sign}\left(\boldsymbol{\xi}_i - \hat{\boldsymbol{x}}_i\right) \\ \boldsymbol{\xi}_i &= \bar{H}_i(\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i) + \sum_{j \in \mathcal{N}_i^1} \bar{P}_i\left(-\bar{P}_j^T\, \bar{P}_j \bar{P}_i^T \hat{\boldsymbol{x}}_i + \bar{P}_i^T\, \bar{P}_i \bar{P}_j^T\, \hat{\boldsymbol{x}}_j\right) + \hat{\boldsymbol{x}}_i\end{aligned} \tag{8}$$

with $\bar{H}_i = H_i \otimes I_d$ where $H_i \in \mathbb{R}^{n_i \times n_i}$ is a diagonal matrix with all the diagonal elements equal to zero but the entry $(g_i(i), g_i(i))$ that is equal to 1, and Θ_i a diagonal matrix of gains defined as $\Theta_i = \mathrm{diag}\{\theta_{i,f_i(1)}, \ldots, \theta_{i,f_i(n_i)}\} \otimes I_d$ where $\theta_{i,f_i(j)}$ is a scalar gain driven by the dynamics

$$\dot{\theta}_{i,f_i(j)} = \omega_{i,f_i(j)}(t) \tag{9}$$

where $\omega_{i,f_i(j)}$ is an update law to be defined.

A few remarks are now in order.

- From an algorithmic standpoint, each agent i can easily extract from $\hat{\boldsymbol{x}}_j$ (with $j \in \mathcal{N}_i^1$) information concerning its own 2-hop neighborhood ($\mathcal{N}_i^2$) given the fact all agents IDs are unique with no need to build each matrix P_j;
- Neighboring agents $(i,j) \in \mathcal{E}$ share the analytical functions encoding their local gradients $\nabla_{x_i} J_j$ and $\nabla_{x_j} J_i$, respectively. Indeed, this is a mild assumption as it requires a one-hop information to be shared once before running the algorithm.

A. *Finite-Time Convergence of Local Observers*

In this section, we focus on demonstrating the finite-time convergence of the proposed local observer given in (8). Note that, since this 2-hop local observer is inspired by the k-hop graph-based local observer proposed in [16], we focus our attention only on the main novelty introduced here, that is a proof of stability with a matrix Θ_i of adaptive gains, which does not require any a priori knowledge on the bound of the input u_i.

In particular, by following the logic of the analysis given in [16], instead of proving that each agent i is able to estimate the state of the 2-hop neighboring agents (namely, $\boldsymbol{x}_i$), we first show that this is equivalent to demonstrate that the local state x_i of each agent i is estimated by the agents belonging to its 2-hop augmented neighborhood $\bar{\mathcal{N}}_i^2$; and then we prove that our observer can achieve this in finite-time.

Let us consider an agent i and notice that for any index $l \in \bar{\mathcal{N}}_i^2$, it follows $\hat{x}_{i,l} = S_l \bar{P}_i^T \hat{\boldsymbol{x}}_i$, while for any index $l \notin \bar{\mathcal{N}}_i^2$ it follows that $S_l \bar{P}_l^T \hat{\boldsymbol{x}}_i$ is null, where the selection matrix S_l is defined as $S_l = \{O_d, \ldots, I_d, \ldots, O_{dn-th}\}$ with I_d the l-th block. Now, let us define the vector $\boldsymbol{\zeta}_i$ which collects n_i copies of the state x_i as $\boldsymbol{\zeta}_i = 1_{n_i} \otimes x_i$ and the vector $\hat{\boldsymbol{\zeta}}_i$ which collects the estimate of the state x_i computed by the 2-hop neighbors of agent i including itself, i.e., $\forall\, j \in \bar{\mathcal{N}}_i^2$, as

$$\hat{\boldsymbol{\zeta}}_i = \left[\hat{x}_{f_i(1),i}, \ldots, \hat{x}_{f_i(n_i),i}\right]^T \tag{10}$$

where $\hat{x}_{f_i(q),i} = S_i \bar{P}_{f_i(q)}^T \hat{\boldsymbol{x}}_{f_i(q)}$ with $q \in \{1, \ldots, n_i\}$.

In addition, let us denote with $\tilde{\boldsymbol{\zeta}}_i$ the collection of the estimation errors of the state x_i, made by the agents belonging to the augmented 2-hop neighborhood $\bar{\mathcal{N}}_i^2$ of an agent i, defined as $\tilde{\boldsymbol{\zeta}}_i = \boldsymbol{\zeta}_i - \hat{\boldsymbol{\zeta}}_i$.

We now state the equivalence between the fact that an agent i can estimate $\boldsymbol{x}_i$ and the fact that the agents belonging to the 2-hop augmented neighborhood of agent i, i.e., $\bar{\mathcal{N}}_i^2$, can estimate its state x_i.

Lemma 1. *Consider a multi-agent system running the local 2-hop observer given in eq.* (8)*. The following facts are equivalent:*

- $\hat{\boldsymbol{x}}_i \to \boldsymbol{x}_i, \quad \forall\, i \in \{1, \ldots, n\}$
- $\hat{x}_{j,l} \to x_l, \quad \forall\, l \in \bar{\mathcal{N}}_j^2, \quad \forall\, j \in \{1, \ldots, n\}$

Proof. The proof follows from [16]. □

From Lemma 1 it follows that, from a mathematical standpoint, proving the finite-time convergence of the local observer can be equivalently expressed in terms of the variables $\tilde{\boldsymbol{\zeta}}_i$ with $i \in \{1, \ldots, n\}$ as follows.

Problem 2. There exists $T > 0$ such that

$$\|\tilde{\boldsymbol{\zeta}}_i\| = \|\boldsymbol{\zeta}_i(t) - \hat{\boldsymbol{\zeta}}_i(t)\| = 0, \;\; \forall\, t \geq T, \;\; i \in \{1, \ldots, n\} \tag{11}$$

In the rest of this section, we focus on proving that the condition given in eq. (11) is achieved by the 2-hop local observer given in eq. (8).

At this point, our goal is to define the form of the error dynamics $\dot{\tilde{\boldsymbol{\zeta}}}_i$ for all $i \in \{1, \ldots, n\}$.

Lemma 2. *Consider a multi-agent system running the local* 2*-hop observer given in eq.* (8)*. The error dynamics of $\tilde{\boldsymbol{\zeta}}_i$ with $i \in \{1, \ldots, n\}$ has the following form*

$$\dot{\tilde{\boldsymbol{\zeta}}}_i = -\bar{\Theta}_i \,\mathrm{sign}\left((M_i^2 \otimes I_d)\tilde{\boldsymbol{\zeta}}_i\right) + \boldsymbol{\mu}_i \tag{12}$$

with $\bar{\Theta}_i = \mathrm{diag}\{\theta_{f_i(1),i}, \ldots, \theta_{f_i(n_i),i}\} \otimes I_d$, $M_i^2 = \mathcal{L}_i^2 + H_i$ and $\boldsymbol{\mu}_i = 1_{n_i} \otimes u_i$.

Proof. The proof follows from [16], with the matrix $\bar{A}_i = O_{dn_i}$ a matrix of zeros, the matrix $\bar{B}_i = I_{dn_i}$ an identity matrix, the gain $\omega_i = 0$, and the scalar constant gain θ_i is replaced by the matrix $\bar{\Theta}_i$ of adaptive gains. □

Let us now demonstrate the finite-time convergence of the proposed 2-hop graph-based observer, that will be achieved by proving the finite-time convergence of the error dynamics given in eq. (12), according to Lemma 1.

Theorem 1. *Let us consider a multi-agent system composed of n agents, where each agent i evolves according to* (7)*, with the finite-time* 2*-hop observer given in* (8)*, and adaptive gains evolving according to* (9)*. Assume, the input u_i to be bounded (with unknown bound), then $\tilde{\boldsymbol{\zeta}}_i$ reaches the origin in finite-time by choosing $\omega_{l,i}$ such as*

$$\omega_{l,i}(t) = \left(S_i \bar{P}_l^T (\boldsymbol{\xi}_l - \hat{\boldsymbol{x}}_l)\right)^{\mathrm{T}} \mathrm{sign}\left(S_i \bar{P}_l^T (\boldsymbol{\xi}_l - \hat{\boldsymbol{x}}_l)\right) \tag{13}$$

Proof. In order to prove the statement, let us consider the following Lyapunov candidate function

$$V_i = \frac{1}{2}\tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \tilde{\boldsymbol{\zeta}}_i + \frac{1}{2\alpha} \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})^2 \tag{14}$$

where $\gamma_{l,i}$ is an upper bound for $\theta_{l,i}$ (i.e., $\theta_{l,i} - \gamma_{l,i} \leq 0$) and α is a positive constant. Therefore, by exploiting eq. (12) and by defining $\bar{\Gamma}_i = \mathrm{diag}\{\gamma_{f_i(1),i}, \ldots, \gamma_{f_i(n_i),i}\} \otimes I_d$ the derivative of V_i becomes

$$\begin{aligned}
\dot{V}_i =\; & \tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \dot{\tilde{\boldsymbol{\zeta}}}_i + \frac{1}{\alpha} \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i}) \dot{\theta}_{l,i} \\
=& - \tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \bar{\Theta}_i \,\mathrm{sign}\left(\left(M_i^2 \otimes I_d\right) \tilde{\boldsymbol{\zeta}}_i\right) \\
&+ \tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \boldsymbol{\mu}_i + \frac{1}{\alpha} \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i}) \omega_{l,i} \pm \sum_{l \in \bar{\mathcal{N}}_i^2} \gamma_{l,i} \omega_{l,i}
\end{aligned} \tag{15}$$

At this point, by considering the expression of $\omega_{l,i}$ in eq. (13), the following chain of equalities holds (see [16] for details)

$$\tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \bar{\Gamma}_i \operatorname{sign}\left(\left(M_i^2 \otimes I_d\right) \tilde{\boldsymbol{\zeta}}_i\right) = \sum_{l \in \bar{\mathcal{N}}_i^2} \gamma_{l,i} \left(S_i \bar{P}_l^T (\boldsymbol{\xi}_l - \hat{\boldsymbol{x}}_l)\right)^{\mathrm{T}} \operatorname{sign}\left(S_i \bar{P}_l^T (\boldsymbol{\xi}_l - \hat{\boldsymbol{x}}_l)\right) = \sum_{l \in \bar{\mathcal{N}}_i^2} \gamma_{l,i}\omega_{l,i} \geq 0$$

and (which is obtained according to a similar reasoning as above)

$$\tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \bar{\Theta}_i \operatorname{sign}\left(\left(M_i^2 \otimes I_d\right) \tilde{\boldsymbol{\zeta}}_i\right) = \sum_{l \in \bar{\mathcal{N}}_i^2} \theta_{l,i}\omega_{l,i}$$

Thus, the Lyapunov time derivative $\dot{V}$ can be rewritten as

$$\begin{aligned}
\dot{V}_i =& - \sum_{l \in \bar{\mathcal{N}}_i^2} \theta_{l,i}\omega_{l,i} + \tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \boldsymbol{\mu}_i \\
&+ \frac{1}{\alpha} \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})\omega_{l,i} \pm \sum_{l \in \bar{\mathcal{N}}_i^2} \gamma_{l,i}\omega_{l,i} \\
&= \left(\frac{1}{\alpha} - 1\right) \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})\omega_{l,i} + \tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \boldsymbol{\mu}_i \\
&- \tilde{\boldsymbol{\zeta}}_i^T \left(M_i^2 \otimes I_d\right) \bar{\Gamma}_i \operatorname{sign}\left(\left(M_i^2 \otimes I_d\right) \tilde{\boldsymbol{\zeta}}_i\right) - \sum_{l \in \bar{\mathcal{N}}_i^2} \gamma_{l,i}\omega_{l,i} \\
&\leq \left(\frac{1}{\alpha} - 1\right) \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})\omega_{l,i} \\
&+ \|\boldsymbol{\mu}_i\| \| \left(M_i^2 \otimes I_d\right) \tilde{\boldsymbol{\zeta}}_i\| - \lambda_{\min}(\bar{\Gamma}_i) \| \left(M_i^2 \otimes I_d\right) \tilde{\boldsymbol{\zeta}}_i\|_1 \\
&= \left(\frac{1}{\alpha} - 1\right) \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})\omega_{l,i} - (\lambda_{\min}(\bar{\Gamma}_i) - \|\boldsymbol{\mu}_i\|) \|(M_i^2 \otimes I_d) \tilde{\boldsymbol{\zeta}}_i\|
\end{aligned} \tag{16}$$

Since, there always exist $\alpha \in (0,1)$ and $\bar{\Gamma}_i$ (i.e., $\gamma_{l,i}$) such that $(\lambda_{\min}(\bar{\Gamma}_i) - \|\boldsymbol{\mu}_i\|) > 0$, then $\dot{V}$ is negative definite by construction and the estimation error $\|\tilde{\boldsymbol{\zeta}}_i\|$ converges to zero.

Concerning the finite-time convergence, by recalling the definition of the Lyapunov function in eq. (14) and by exploiting the following relationships

$$\begin{cases} (M_i^2 \otimes I_d)^T (M_i^2 \otimes I_d) = (M_i^2 M_i^2 \otimes I_d) \\ \left\| \tilde{\boldsymbol{\zeta}}_i^{\mathrm{T}} (M_i^2 M_i^2 \otimes I_d) \tilde{\boldsymbol{\zeta}}_i \right\| \geq \lambda_{\min}(M_i^2) \|\tilde{\boldsymbol{\zeta}}_i\| \end{cases} \tag{17}$$

it holds

$$\begin{aligned}
V_i &\leq \frac{1}{2}\lambda_{\max}(M_i^2) \parallel \tilde{\boldsymbol{\zeta}}_i \parallel^2 + \frac{1}{2\alpha} \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})^2 \\
&\leq \phi_{M,i} \left(\parallel \tilde{\boldsymbol{\zeta}}_i \parallel^2 + \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})^2 \right)
\end{aligned} \tag{18}$$

with $\phi_{M,i} = \max\left\{ \frac{1}{2}\lambda_{\max}(M_i^2), \frac{1}{2\alpha} \right\}$.

Finally, the Lyapunov derivative in eq. (16) can be rewritten as

$$\begin{aligned}
\dot{V}_i \leq& \left(\frac{1}{\alpha} - 1\right) \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})\omega_{l,i} \\
&- (\lambda_{\min}(\bar{\Gamma}_i) - \|\boldsymbol{\mu}_i\|)\lambda_{\min}(M_i^2) \parallel \tilde{\boldsymbol{\zeta}}_i \parallel
\end{aligned} \tag{19}$$

from which it follows

$$\dot{V}_i \leq -\phi_{m,i} \left(\parallel \tilde{\boldsymbol{\zeta}}_i \parallel + \sum_{l \in \bar{\mathcal{N}}_i^2} |\theta_{l,i} - \gamma_{l,i}| \right) \tag{20}$$

with $\phi_{m,i} = \min\left\{ \left(\frac{1}{\alpha} - 1\right)\varepsilon, (\lambda_{\min}(\bar{\Gamma}_i) - \|\boldsymbol{\mu}_i\|)\lambda_{\min}(M_i^2) \right\}$ and $\varepsilon = \min_l \omega_{l,i}$.

At this point, by considering the following inequality

$$\begin{aligned}
&\left(\parallel \tilde{\boldsymbol{\zeta}}_i \parallel + \sum_{l \in \bar{\mathcal{N}}_i^2} |\theta_{l,i} - \gamma_{l,i}| \right)^2 \geq \parallel \tilde{\boldsymbol{\zeta}}_i \parallel^2 + \left(\sum_{l \in \bar{\mathcal{N}}_i^2} |\theta_{l,i} - \gamma_{l,i}| \right)^2 \\
&\geq \parallel \tilde{\boldsymbol{\zeta}}_i \parallel^2 + \sum_{l \in \bar{\mathcal{N}}_i^2} (\theta_{l,i} - \gamma_{l,i})^2 = \frac{V_i}{\phi_{M,i}}
\end{aligned} \tag{21}$$

it follows that eq. (20) can be finally rewritten as $\dot{V} \leq -\frac{\phi_{m,i}}{\sqrt{\phi_{M,i}}}\sqrt{V_i}$ which proves the finite-time convergence of V to zero, thus completing the proof. □

B. Finite-Time Convergence of the Signed Gradient Descent

In this section, we demonstrate the finite-time convergence of the signed gradient descent given in (7). Our analysis follows from [18], where the main technical difference relies on the fact that in our scenario the signed gradient descent protocol is fed with the estimated states rather than the actual ones. In particular, first we show that, by resorting to [18], under Assumption 1, a (centralized) signed gradient descent policy fed with the required 2-hop information $\boldsymbol{x}_i$ for each agent i would converge towards the optimal solution $\boldsymbol{x}^\star$ in finite-time; then we demonstrate that the proposed signed gradient descent policy fed with the estimated state $\hat{\boldsymbol{x}}_i$ provided by the local observer run by each agent i still converges towards the optimal solution $\boldsymbol{x}^\star$ in finite-time.

Let us now show the finite-time convergence of a (centralized) signed gradient descent policy where each agent i has full access to its 2-hop neighborhood information $\boldsymbol{x}_i$.

Proposition 1. *Consider a multi-agent system with agent dynamics* (1)*. Assume each agent i has instantaneous access to the* 2*-hop augmented neighborhood state $\boldsymbol{x}_i$. Then, the following control law $u_i(t)$ with $i \in \{1, \ldots, n\}$ solves Problem 1*

$$u_i(t) = -c_i \operatorname{sign}\left(\nabla_{x_i} J(\boldsymbol{x}_i)\right) \tag{22}$$

Proof. To prove this result, it suffices to notice that under Assumption 1, the Hessian $H_J(\boldsymbol{x})$ matrix of the global objective function $J(\boldsymbol{x})$ is positive definite by construction. Therefore, the finite-time convergence of the signed gradient descent control law given in (22) follows directly from [18, Theorem 8]. □

Theorem 2. *Consider a multi-agent system with agent dynamics* (1)*. Assume each agent i is running the update law* (7)*, the* 2*-hop state observer* (8) *and the gains $\theta_{i,f_i(j)}$ with $j \in \{1, \ldots, n_i\}$ are updated according to Theorem 1. Then, the multi-agent system reaches the optimal solution $\boldsymbol{x}^\star$ in finite-time and Problem 1 is solved.*

Proof. To prove this result, it suffices to notice that the update law (7) can be rewritten as

$$u_i(t) = -c_i \operatorname{sign}\left(\nabla_{x_i} J(\boldsymbol{x}_i - \boldsymbol{\varepsilon}_i)\right) \tag{23}$$

with the estimation error ε_i defined as $\varepsilon_i = \boldsymbol{x}_i - \hat{\boldsymbol{x}}_i$. At this point, from Theorem 1, we know that there exists $T \geq 0$ such that for all $t \geq T$ it holds $\|\varepsilon_i(t)\| = 0, \quad i \in \{1, \ldots, n\}$. It follows that eq. (23) becomes

$$u_i(t) = -c_i \operatorname{sign}\left(\nabla_{x_i} J(\boldsymbol{x}_i)\right), \quad t \geq T \tag{24}$$

thus, the condition given in Proposition 1 holds and the result follows from Assumption 1. □

IV. Simulation Results

In this section, we propose simulation results involving a weighted least-square (WLS) estimation problem formulation. In particular, motivated by the needs of our recent EU project PANTHEON – "Precision farming of hazelnut orchards" (Grant Agreement number 774571) and inspired by recent works in the field of precision farming [19], [20], we are interested in the application of distributed WLS techniques to identify the thermal properties of a soil.

As pointed out in several works in the field of precision farming, such as in [21], the investigation of the thermal properties of a soil can have significant practical consequences, e.g. evaluation of optimum conditions for plant growth and development, and can be utilized for the control of thermal-moisture regime of soil in the field. Specifically, in the context of the PANTHEON project, these parameters can be used to compute the coefficients of the dynamical model proposed in [22] in order to design an advanced water management system for large-scale hazelnut orchards. Briefly, this dynamical model considers a discretization of the field into small rectangular parcels, where for each parcel i the variation of the soil moisture is obtained according to a hydrological balance model. As a consequence, the finite-time distributed methodology proposed in this paper could be used within the context of an Internet of Thing (IoT) monitoring network deployed over the field, as the one that is being developed within this project, in order to identify the coefficients of the dynamical model of the different parcels in a collaborative fashion.

As pointed out in several works [4], [5], WLS problems follow into the family of objective functions $J(\boldsymbol{x})$ for which the local components $\nabla_{x_i} J(\boldsymbol{x}_i)$ of the gradient $\nabla_x J(\boldsymbol{x}_i)$ requires 2-hop information. Let us now briefly review the WLS problem formulation. To this end, let us denote with $y_i \in \mathbb{R}^q$ and $\nu_i \in \mathbb{R}^q$ the measurement vector and the corresponding measurement noise vector, respectively.

In particular, for each agent i the measurement model y_i can be written as

$$y_i = \sum_{j=1}^{n} W_{ij} x_j + \nu_i = \sum_{j \in \bar{\mathcal{N}}_i^1} W_{ij} x_j + \nu_i \tag{25}$$

where the second equality follows from the fact that $W_{ij} = O_{q \times d}$ for any $j \notin \bar{\mathcal{N}}_i^1$. At this point, by introducing the stacked vectors $y = [y_1^T, \ldots, y_n^T]^T \in \mathbb{R}^{qn}$ and $\nu = [\nu_1^T, \ldots, \nu_n^T]^T \in \mathbb{R}^{qn}$, eq. (25) can be equivalently expressed in vector form as

$$\boldsymbol{y} = W\boldsymbol{x} + \nu \tag{26}$$

where the matrix W is defined as $W = [W_1^T, \ldots, W_n^T]^T \in \mathbb{R}^{qn \times nd}$ and the matrix W_i defined as $W_i = [W_{i1}, \ldots, W_{in}] \in \mathbb{R}^{q \times nd}$.

It follows that the global objective function to be minimized encoding the WLS problem formulation is

$$J(\boldsymbol{x}) = \sum_{i=1}^{n} J_i\left(\{x_j\}_{j \in \bar{\mathcal{N}}_i^1}\right) = \frac{1}{2} \parallel \boldsymbol{y} - W\boldsymbol{x} \parallel^2 \tag{27}$$

where each local objective $J_i\left(\{x_j\}_{j \in \bar{\mathcal{N}}_i^1}\right)$ is

$$J_i\left(\{x_j\}_{j \in \bar{\mathcal{N}}_i^1}\right) = \frac{1}{2} \parallel y_i - W_i\, \boldsymbol{x} \parallel^2 \tag{28}$$

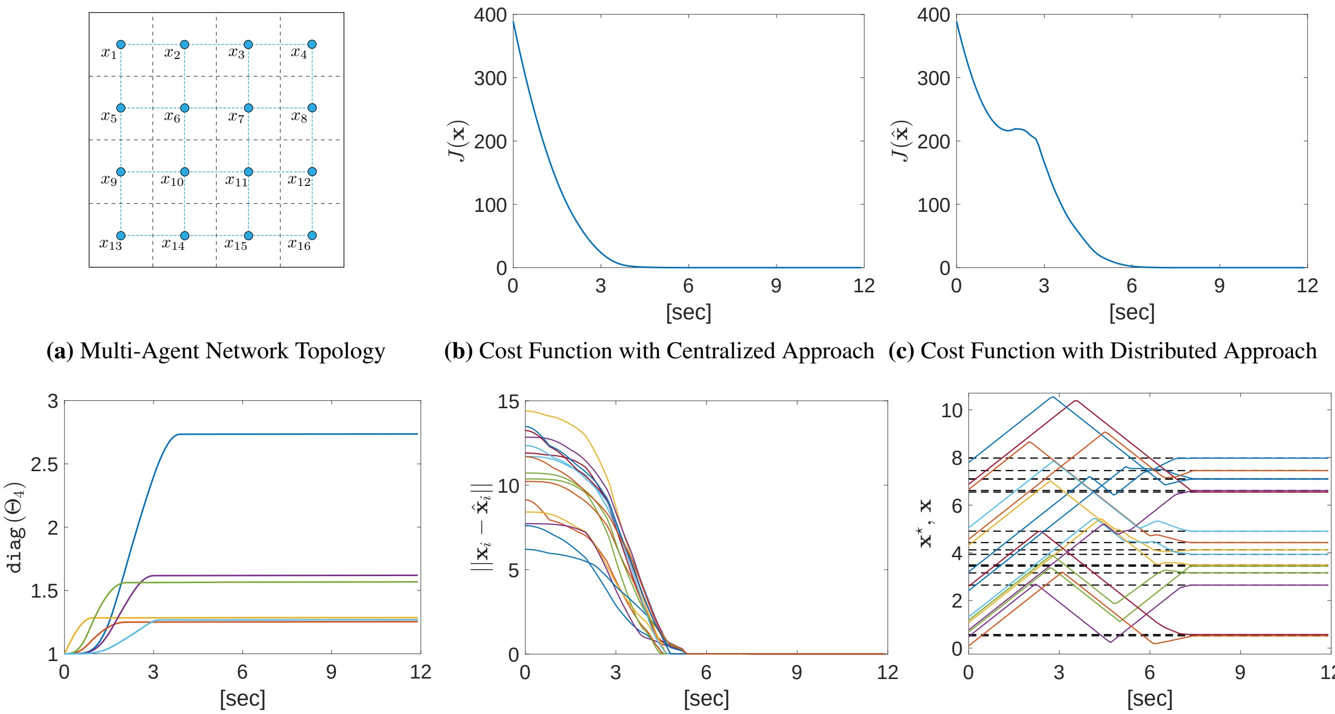
We reiterate that the minimization of the local objective function J_i in eq. (28) requires 2-hop information concerning the augmented neighborhood $\bar{\mathcal{N}}_i^1$ of each agent i to compute the components of the gradient as pointed out in [5]. Thus, it cannot be naively implemented in a fully distributed manner as it is.

For the numerical simulations, we consider a multi-agent system composed of 16 agents with continuous-time single integrator dynamics as in eq. (1) where $d = 1$ and measurement model as in eq. (25) with $q = 2$. In particular, within the context of the precision farming setting described above, an agent i represents the i-th element of an Internet of Thing (IoT) monitoring network and it is located on the i-th parcel of the field. In particular, Figure (1a) depicts the undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ encoding the pairwise interconnections among the agents, which according to the model given in [22] represent possible interactions between neighboring parcels. Both the centralized signed gradient descent given in (22) where $c_i = 1, \forall\, i$, and the proposed distributed algorithm, which relies on local observers to compute the descent direction as per (7), (8), and (13), are considered.

Figure (1b) shows the optimization carried out over time by the centralized signed gradient descent based on the knowledge of the 2-hop information, which reaches the global minimum in finite-time according to Proposition 1. Figure (1c) shows the behavior of the proposed approach where it can be noticed an increasing trend of the cost function during the first seconds of the execution due to the erroneous knowledge of the 2-hop neighborhood for each agent i. Figure (1d) shows the adaptive tuning of the gains $\Theta_i = \operatorname{diag}\{\theta_{i,f_i(1)}, \ldots, \theta_{i,f_i(n_i)}\}$ for the agent 4, where it should be noticed that the 2-hop neighborhood is composed of 5 agents, namely $\mathcal{N}_4^2 = \{2, 3, 7, 8, 12\}$, while Figure (1e) shows the finite-time convergence of the local observers according to Theorem 2. Finally, Figure (1f) numerically demonstrates how the state $\boldsymbol{x}$ computed by the proposed distributed approach reaches the (dashed) optimal value $\boldsymbol{x}^\star = (W^T\, W)^{-1}\, W^T\, \boldsymbol{y}$ in finite-time according to Theorem 2.

V. Conclusions

A distributed optimization problem for continuous time multi-agent systems where the global objective is defined as the sum of locally coupled strictly convex cost functions has been addressed. A novel alternative approach based on a distributed signed gradient descent algorithm which relies on local observers to compute the descent direction has been proposed. The proposed approach does not require 2-rounds of communications but instead exploits an estimate of the 2-hop neighborhood provided by the local observers along with the (one-time) exchange of the structure of the local components of



(**a**) Multi-Agent Network Topology (**b**) Cost Function with Centralized Approach (**c**) Cost Function with Distributed Approach

(**d**) Agent 4 $\mathrm{diag}(\Theta_4)$ dynamics (**e**) Observers finite-time convergence (**f**) Optimization finite-time convergence

Fig. 1: Simulation results involving a multi-agent system composed of 16 agents: centralized vs. distributed approach.

the gradients for pairs of neighboring agents. A weighted least squared problem formulation, which can be used in the context of precision-farming to identify the soil thermal properties in a large-scale hazelnut orchard to design an intelligent water management system, has been proposed as an optimization problem case study to numerically corroborate the theoretical findings. Future work will mainly focus on generalizing this optimization framework for time-varying global cost functions.

References

[1] M. Zhu and S. Martnez, *Distributed Optimization-Based Control of Multi-Agent Networks in Complex Environments*. Springer Publishing Company, Incorporated, 2015.

[2] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, April 2010.

[3] I. Lobel and A. Ozdaglar, "Distributed subgradient methods for convex optimization over random networks," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1291–1306, June 2011.

[4] D. E. Marelli and M. Fu, "Distributed weighted least-squares estimation with fast convergence for large-scale systems," *Automatica*, vol. 51, pp. 27 – 39, 2015.

[5] M. Todescato, N. Bof, G. Cavraro, R. Carli, and L. Schenato, "Generalized gradient optimization over lossy networks for partition-based estimation," *ArXiv e-prints*, Oct. 2017.

[6] X. Wu and J. Lu, "Fenchel dual gradient methods for distributed convex optimization over time-varying networks," in *IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec 2017, pp. 2894–2899.

[7] G. Qu and N. Li, "Accelerated distributed nesterov gradient descent for convex and smooth functions," in *IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec 2017, pp. 2260–2267.

[8] X. Zeng, P. Yi, and Y. Hong, "Distributed continuous-time algorithm for constrained convex optimizations via nonsmooth analysis approach," *IEEE Trans. on Aut. Contr.*, vol. 62, no. 10, pp. 5227–5233, Oct 2017.

[9] J. Lu and C. Y. Tang, "Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case," *IEEE Transactions on Automatic Control*, vol. 57, no. 9, pp. 2348–2354, Sept 2012.

[10] B. Gharesifard and J. Cortes, "Distributed continuous-time convex optimization on weight-balanced digraphs," *IEEE Trans. on Aut. Contr.*, vol. 59, no. 3, pp. 781–786, March 2014.

[11] S. S. Kia, J. Cortes, and S. Martinez, "Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication," *Automatica*, vol. 55, pp. 254 – 264, 2015.

[12] P. Lin, W. Ren, and J. A. Farrell, "Distributed continuous-time optimization: Nonuniform gradient gains, finite-time convergence, and convex constraint set," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2239–2253, May 2017.

[13] S. Rahili and W. Ren, "Distributed continuous-time convex optimization with time-varying cost functions," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, April 2017.

[14] Z. Feng and G. Hu, "Finite-time distributed optimization with quadratic objective functions under uncertain information," in *IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec 2017, pp. 208–213.

[15] X. Wang, J. Zhou, S. Mou, and M. J. Corless, "A Distributed Algorithm for Least Square Solutions of Linear Equations," *ArXiv e-prints*, 2017.

[16] A. Gasparri and A. Marino, "A k-hop graph-based observer for large-scale networked systems," in *56th IEEE Annual Conference on Decision and Control, CDC 2017, Melbourne, Australia, December 12-15, 2017*, 2017, pp. 4747–4752.

[17] S. Boyd and L. Vandenbergher, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[18] J. Cortes, "Finite-time convergent gradient flows with applications to network consensus," *Automatica*, vol. 42, no. 11, pp. 1993 – 2000, 2006.

[19] S. Mansour, É. Canot, and M. Muhieddine, "Identification of the thermophysical properties of the soil by inverse problem," *Journal of Heat Transfer*, vol. 138, no. 9, p. 091010, 2016.

[20] K. Oyeyemi, O. Sanuade, M. Oladunjoye, A. Aizebeokhai, A. Olaojo, J. Fatoba, O. Olofinnade, W. Ayara, and O. Oladapo, "Data on the thermal properties of soil and its moisture content," *Data in Brief*, vol. 17, pp. 900 – 906, 2018.

[21] M. Oladunjoye, O. Sanuade, and A. Olaojo, "Variability of soil thermal properties of a seasonally cultivated agricultural teaching and research farm, university of ibadan, south-western nigeria," *Global Journal of Science Frontier Research Agriculture and Veterinary*, vol. 13, no. 8, pp. 41–64, 2013.

[22] N. B. Rosselló, A. Gasparri, and E. Garone, "D5.1 Water Management Control – PANTHEON Project," http://www.project-pantheon.eu/images/PANTHEOND51submitted.pdf, 2018.