Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication ⋆

Solmaz S. Kia [a] Jorge Cortés [a] Sonia Martínez [a]

[a] *Department of Mechanical and Aerospace Engineering, University of California at San Diego, La Jolla, CA 92093, USA*

Abstract

This paper proposes a novel class of distributed continuous-time coordination algorithms to solve network optimization problems whose cost function is strictly convex and equal to a sum of local cost functions associated to the individual agents. We establish the exponential convergence of the proposed algorithm under (i) strongly connected and weight-balanced digraph topologies when the local costs are strongly convex with globally Lipschitz gradients, and (ii) connected graph topologies when the local costs are strongly convex with locally Lipschitz gradients. We also characterize the algorithm's privacy preservation properties and its correctness under time-varying interaction topologies. Motivated by practical considerations, we analyze the algorithm implementation with discrete-time communication. We consider three scenarios: periodic, centralized event-triggered, and distributed event-triggered communication. First, we provide an upper bound on the stepsize that guarantees exponential convergence over connected undirected graphs for implementations with periodic communication. Building on this result, we design a provably-correct centralized event-triggered communication scheme that is free of Zeno behavior. Finally, we develop a distributed, asynchronous event-triggered communication scheme that is also free of Zeno with asymptotic convergence guarantees. Several simulations illustrate our results.

Key words: cooperative control, distributed convex optimization, weight-balanced digraphs, event-triggered control.

1 Introduction

An important class of distributed convex optimization problems consists of the (un-)constrained network optimization of a sum of convex functions, each one representing a local cost only known to an individual agent. Such problems model a wide range of network scenarios where the global cost function is a performance metric consisting of a sum of local private utility functions. Examples are numerous and include distributed parameter estimation [Ram et al., 2010, Wan and Lemmon, 2009], distributed economic dispatch [Zhang et al., 2011], distributed statistical learning [Boyd et al., 2010], and distributed optimal resource allocation over networks [Madan and Lall, 2006, Preciado et al., 2013]. To find the network optimizers of such problems, this paper proposes a coordination model where each agent runs a purely local continuous-time evolution dynamics and communicates at discrete instants of time with its neighbors. In doing so, we are motivated by the desire of combining the conceptual ease associated with the analysis of continuous-time dynamical systems and the practical constraints imposed by real-time implementations. Our development is based on a novel continuous-time distributed algorithm design whose stability and convergence properties can be analyzed through standard Lyapunov functions.

Literature review: There are two areas on which this paper builds: distributed convex optimization and event-triggered control of networked systems. In distributed convex optimization, most coordination algorithms are time-varying, consensus-based dynamics [Boyd et al., 2010, Duchi et al., 2012, Johansson et al., 2009, Nedić and Ozdaglar, 2009, Tsitsiklis et al., 1986, Zhu and Martínez, 2012] implemented in discrete time. Recent work [Gharesifard and Cortés, 2014, Lu and Tang, 2012, Wang and Elia, 2010, 2011, Zanella et al., 2011] has introduced continuous-time dynamical solvers whose convergence properties can be analyzed via classical stability analysis. This has the added advantage of facilitating the characterization of additional properties such as speed of convergence, disturbance rejection, and robustness to parameter and model uncertainty. Wang and Elia [2011] establish asymptotic convergence under connected undirected graphs and Gharesifard and Cortés [2014] extend the design and analysis to the case of strongly connected,

⋆ Corresponding author: S.S. Kia

Email addresses: skia@ucsd.edu (Solmaz S. Kia), cortes@ucsd.edu (Jorge Cortés), soniamd@ucsd.edu (Sonia Martínez).

weight-balanced digraphs. Wang and Elia [2010] study the convergence properties of Euler discretizations of these continuous-time algorithms. The continuous-time algorithms in [Lu and Tang, 2012, Zanella et al., 2011] require twice-differentiable, strictly convex local cost functions to make use of the inverse of their Hessian and need a careful initialization to guarantee asymptotic convergence under undirected connected graphs. The novel class of continuous-time algorithms proposed here upon which our implementations with discrete-time communication are built do not suffer from the limitations discussed above and have exponential convergence guarantees. Regarding event-triggered control of networked systems, recent years have seen an increasing body of work that seeks to trade computation and decision making at the agent level for less communication, sensing or actuator effort while still guaranteeing a desired level of performance, see e.g. [Heemels et al., 2012, Mazo Jr. and Tabuada, 2011, Wang and Lemmon, 2011]. Closest to the problem considered here are works that study event-triggered communication laws for average consensus, see e.g., [Dimarogonas et al., 2012, Garcia et al., 2013, Nowzari and Cortés, 2014]. The strategies proposed in [Wan and Lemmon, 2009] save communication effort in discrete-time implementations by using local triggering events but are not guaranteed to avoid the possibility of Zeno behavior, i.e., an infinite number of triggered events in a finite period of time. Our goal in this paper is to combine the best of both approaches by synthesizing provably-correct continuous-time distributed dynamical systems, one per agent, which only require communication with neighbors at discrete instants of time. We are particularly interested in the opportunistic determination of this communication times via event triggering schemes. Such coordination algorithms are amenable to the analysis machinery offered by Lyapunov stability while, at the same time, are more in line with the practical limitations encountered in realistic scenarios.

Statement of contributions: We propose a novel class of continuous-time, gradient-based distributed algorithms for network optimization where the global objective function is strictly convex and equal to the sum of local cost functions, one per agent. We prove that these algorithms converge exponentially under strongly connected and weight-balanced agent interactions when the local cost functions are strongly convex and their gradients are globally Lipschitz. Under connected, undirected graphs, we establish exponential convergence when the local gradients are just locally Lipschitz and asymptotic convergence when the local cost functions are simply convex. We also study the algorithm convergence under networks with time-varying topologies and, motivated by privacy preservation considerations, we characterize the topological requirements on the communication graph, algorithm parameters, and initial conditions necessary for an agent to reconstruct the local gradients of other agents. Our technical approach builds on Lyapunov stability analysis and the identification of suitable Lyapunov functions. The availability of these functions makes possible our ensuing design of provably-correct continuous-time implementations with discrete-time communication. In particular, for networks with connected graph topologies, we obtain an upper bound on the suitable stepsizes that guarantee exponential convergence under periodic communication. Building on this result, we design a centralized, synchronous event-triggered communication scheme with the same exponential convergence guarantees and free of Zeno behavior. Finally, we develop a Zeno-free asynchronous event-triggered communication scheme whose execution only requires agents to interchange information with their neighbors. We establish the asymptotic convergence to the solution of the network optimization problem under this distributed communication scheme. Several simulations illustrate our results.

Organization: Section 2 introduces basic notation and concepts from graph theory and convex functions. Section 3 presents the problem statement. Section 4 introduces our novel continuous-time distributed convex optimization algorithm and characterizes its properties on convergence and privacy preservation. Section 5 discusses continuous-time implementations with discrete-time-communication of the proposed algorithm. Section 6 illustrates our results in simulation. Finally, Section 7 gathers our conclusions and ideas for future work.

2 Preliminaries

In this section, we introduce our notation and some basic concepts from convex functions and graph theory.

2.1 Notation

Let $\mathbb{R}$ and $\mathbb{N}$ denote, respectively, the set of real and natural numbers. We use $\Re(\cdot)$ to represent the real part of a complex number. The transpose of a matrix $\mathbf{A}$ is $\mathbf{A}^\top$. We let $\mathbf{1}_n$ (resp. $\mathbf{0}_n$) denote the vector of n ones (resp. n zeros), and denote by $\mathbf{I}_n$ the $n \times n$ identity matrix. We let $\mathbf{\Pi}_n = \mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top$. When clear from the context, we do not specify the matrix dimensions. For $\mathbf{A} \in \mathbb{R}^{n\times m}$ and $\mathbf{B} \in \mathbb{R}^{p\times q}$, we let $\mathbf{A}\otimes\mathbf{B}$ denote their Kronecker product. For $\mathbf{u} \in \mathbb{R}^d$, $\|\mathbf{u}\| = \sqrt{\mathbf{u}^\top\mathbf{u}}$ denotes the standard Euclidean norm. For vectors $\mathbf{u}_1, \cdots, \mathbf{u}_m$, we let $\mathbf{u} = (\mathbf{u}_1, \cdots, \mathbf{u}_m)$ represent the aggregated vector. In a networked system, we distinguish the local variables at each agent by a superscript, e.g., $\mathbf{x}^i$ is the local state of agent i. If $\mathbf{p}^i \in \mathbb{R}^d$ is a variable of agent i, the aggregated $\mathbf{p}^i$'s of the network of N agents is represented by $\mathbf{p} = (\mathbf{p}^1, \cdots, \mathbf{p}^N) \in (\mathbb{R}^d)^N$. A differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ is *strictly convex* over a convex set $C \subset \mathbb{R}^d$ iff

$$(\mathbf{z}-\mathbf{x})^\top(\nabla f(\mathbf{z}) - \nabla f(\mathbf{x})) > 0, \;\; \forall\, \mathbf{x}, \mathbf{z} \in C, \; \mathbf{x} \neq \mathbf{z},$$

and it is *m-strongly convex* ($m > 0$) iff

$$(\mathbf{z}-\mathbf{x})^\top(\nabla f(\mathbf{z}) - \nabla f(\mathbf{x})) \geq m\|\mathbf{z}-\mathbf{x}\|^2, \; \forall\, \mathbf{x}, \mathbf{z} \in C, \; \mathbf{x} \neq \mathbf{z}.$$

A function $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz with constant $M > 0$, or simply *M-Lipschitz*, over a set $C \subset \mathbb{R}^d$ iff

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq M \|\mathbf{x} - \mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in C.$$

2.2 Graph Theory

Here, we briefly review some basic concepts from graph theory and linear algebra following [Bullo et al., 2009]. A *directed graph*, or simply a *digraph*, is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, N\}$ is the *node set* and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the *edge set*. An edge from i to j, denoted by (i, j), means that agent j can send information to agent i. For an edge $(i, j) \in \mathcal{E}$, i is called an *in-neighbor* of j and j is called an *out-neighbor* of i. A graph is *undirected* if $(i, j) \in \mathcal{E}$ anytime $(j, i) \in \mathcal{E}$. A *directed path* is a sequence of nodes connected by edges.

A *weighted digraph* is a triplet $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathsf{A})$, where $(\mathcal{V}, \mathcal{E})$ is a digraph and $\mathsf{A} \in \mathbb{R}^{N \times N}$ is a weighted *adjacency* matrix with the property that $\mathsf{a}_{ij} > 0$ if $(i, j) \in \mathcal{E}$ and $\mathsf{a}_{ij} = 0$, otherwise. A weighted digraph is *undirected* if $\mathsf{a}_{ij} = \mathsf{a}_{ji}$ for all $i, j \in \mathcal{V}$. We refer to a strongly connected and undirected graph as a *connected graph*. The *weighted out-degree* and *weighted in-degree* of a node i, are respectively, $\mathsf{d}_{\text{in}}^i = \sum_{j=1}^N a_{ji}$ and $\mathsf{d}_{\text{out}}^i = \sum_{j=1}^N \mathsf{a}_{ij}$. A digraph is *weight-balanced* if at each node $i \in \mathcal{V}$, the weighted out-degree and weighted in-degree coincide (although they might be different across different nodes). The *(out-) Laplacian* matrix is $\mathsf{L} = \mathsf{D}^{\text{out}} - \mathsf{A}$, where $\mathsf{D}^{\text{out}} = \text{Diag}(\mathsf{d}_{\text{out}}^1, \cdots, \mathsf{d}_{\text{out}}^N) \in \mathbb{R}^{N \times N}$. Note that $\mathsf{L}\mathbf{1}_N = \mathbf{0}$. A digraph is weight-balanced if and only if $\mathbf{1}_N^T \mathsf{L} = \mathbf{0}$ if and only if $\text{Sym}(\mathsf{L})$ is positive semi-definite. Based on the structure of L, at least one of the eigenvalues of L is zero and the rest of them have nonnegative real parts. We denote the eigenvalues of L by $\lambda_1, \ldots, \lambda_N$, where $\lambda_1 = 0$ and $\Re(\lambda_i) \leq \Re(\lambda_j)$, for $i < j$, and the eigenvalues of $\text{Sym}(\mathsf{L})$ by $\hat{\lambda}_1, \ldots, \hat{\lambda}_N$. For a strongly connected and weight-balanced digraph, zero is a simple eigenvalue of both L and $\text{Sym}(\mathsf{L})$. In this case, we order the eigenvalues of $\text{Sym}(\mathsf{L})$ as $\hat{\lambda}_1 = 0 < \hat{\lambda}_2 \leq \hat{\lambda}_3 \leq \cdots \leq \hat{\lambda}_N$. For convenience, we define $\mathbf{L} = \mathsf{L} \otimes \mathsf{I}_d$ and $\mathbf{\Pi} = \mathbf{\Pi}_N \otimes \mathsf{I}_d$ to deal with variables of dimension $d \in \mathbb{N}$.

3 Problem Definition

Consider a network of N agents with interaction topology described by a strongly connected, weight-balanced digraph $\mathcal{G}$. Each agent $i \in \{1, \ldots, N\}$ is endowed with a local cost function $f^i : \mathbb{R}^d \to \mathbb{R}$ which is assumed differentiable. The global network cost function $f : \mathbb{R}^d \to \mathbb{R}$ is defined as $f(\mathsf{x}) = \Sigma_{i=1}^N f^i(\mathsf{x})$. We assume this function to be strictly convex. Our objective is to design a distributed optimization algorithm such that each agent obtains the global minimizer $-\infty < \mathsf{x}^\star < \infty$ of the feasible optimization problem

$$\mathsf{x}^\star = \arg\min_{\mathsf{x} \in \mathbb{R}^d} f(\mathsf{x})$$

using only its own local data and exchanged information with its neighbors (note that the strict convexity of f implies the uniqueness of the optimizer). We are also interested in characterizing the privacy preservation properties of the algorithmic solution to this distributed optimization problem. Specifically, we aim to identify conditions guaranteeing that no information about the local cost function of an agent is revealed to, or can be reconstructed by, any other agent in the network.

4 Distributed Continuous-Time Algorithm for Convex Optimization

In this section we provide a novel continuous-time distributed coordination algorithm to solve the problem stated in Section 3 and analyze in detail its convergence properties. For $i \in \{1, \ldots, N\}$, consider

$$\dot{\mathbf{v}}^i = \alpha\beta \sum_{j=1}^N \mathsf{a}_{ij}(\mathbf{x}^i - \mathbf{x}^j), \tag{1a}$$

$$\dot{\mathbf{x}}^i = -\alpha \nabla f^i(\mathbf{x}^i) - \beta \sum_{j=1}^N \mathsf{a}_{ij}(\mathbf{x}^i - \mathbf{x}^i) - \mathbf{v}^i, \tag{1b}$$

with $\alpha, \beta > 0$. In compact form, this algorithm can be written in the network variables $\mathbf{x}, \mathbf{v} \in (\mathbb{R}^d)^N$ as follows

$$\dot{\mathbf{v}} = \alpha\beta \mathbf{L}\mathbf{x}, \tag{2a}$$

$$\dot{\mathbf{x}} = -\alpha \nabla \tilde{f}(\mathbf{x}) - \beta \mathbf{L}\mathbf{x} - \mathbf{v}. \tag{2b}$$

Here, $\tilde{f} : (\mathbb{R}^d)^N \to \mathbb{R}$ is defined by $\tilde{f}(\mathbf{x}) = \sum_{i=1}^N f^i(\mathbf{x}^i)$. This algorithm is distributed because each agent only needs to receive information from its out-neighbors about their corresponding variables in $\mathbf{x}$. In contrast, the continuous-time coordination algorithms in [Gharesifard and Cortés, 2014, Wang and Elia, 2011] require the communication of the corresponding variables in both $\mathbf{x}$ and $\mathbf{v}$.

The synthesis of algorithm (2) is inspired by the following feedback control considerations. In (2b), each agent follows a local gradient descent while trying to agree with its neighbors on their estimate of the final value. However, as the local gradients are not the same, this dynamics by itself would never converge. Therefore, to correct this error, each agent uses an integral feedback term $\mathbf{v}^i$ whose evolution is driven by the agent disagreement according to (2a).

Our analysis of the algorithm convergence is structured in two parts, depending on the directed character of the interactions. Section 4.1 deals with strongly connected and weight-balanced digraphs and Section 4.2 deals with connected undirected graphs. In each case, we identify conditions on the agent cost functions that guarantee asymptotic convergence. Given the challenges posed by directed information flows, it is not surprising that we can establish stronger results under less restrictive assumptions for the case of undirected topologies.

4.1 Strongly Connected, Weight-Balanced Digraphs

Here, we study the convergence of the distributed optimization algorithm (1) over strongly connected and weight-balanced digraph topologies. We first consider the case where the interaction topology is fixed, and then discuss the time-varying interaction topologies. The following result identifies conditions on the local cost functions $\{f^i\}_{i=1}^N$ and the parameter β to guarantee the exponential convergence of (1) to the solution of the distributed optimization problem.

Theorem 1 *(Convergence of* (1) *over strongly connected and weight-balanced digraphs): Let $\mathcal{G}$ be a strongly connected and weight-balanced digraph. Assume the local cost function f^i, $i \in \{1,\dots,N\}$, is m^i-strongly convex, differentiable, and its gradient is M^i-Lipschitz on $\mathbb{R}^d$. For $m_T = \min\{m^1,\dots,m^N\}$ and $M_T = \max\{M^1,\dots,M^N\}$, let $\beta > 0$ be such that*

$$\alpha^2(\phi+1)m_T+9\alpha\beta\hat{\lambda}_2\phi-4\alpha^2M_T^2-4\alpha^2(\phi+1)^2>0, \quad (3)$$

is satisfied for some $\phi>0$ with $\phi+1>\frac{4M_T^2}{m_T}$. Then, for any $\alpha>0$ and each $i\in\{1,\dots,N\}$, the algorithm (1) *over $\mathcal{G}$ makes $\mathbf{x}^i(t)\to\mathbf{x}^\star$ exponentially fast as $t\to\infty$, starting from initial conditions $\mathbf{x}^i(0),\mathbf{v}^i(0)\in\mathbb{R}^d$ with $\sum_{i=1}^N \mathbf{v}^i(0)=\mathbf{0}_d$.*

PROOF. Note that since the digraph is weight-balanced, we have $\mathbf{1}_N^\top\mathbf{L}=\mathbf{0}_N$. Therefore, multiplying (2a) by $\mathbf{1}_N^\top\otimes\mathbf{I}_d$ from the left results in

$$\sum_{i=1}^N\dot{\mathbf{v}}^i=\mathbf{0}\Rightarrow\sum_{i=1}^N\mathbf{v}^i(t)=\sum_{i=1}^N\mathbf{v}^i(0)=\mathbf{0},\quad\forall t\geq 0. \quad (4)$$

Next, we obtain the equilibrium point of (1), $(\bar{\mathbf{v}},\bar{\mathbf{x}})$, by setting the derivatives to zero, i.e.,

$$\mathbf{0}=\alpha\beta\mathbf{L}\bar{\mathbf{x}}, \quad (5a)$$
$$\mathbf{0}=-\alpha\nabla\tilde{f}(\bar{\mathbf{x}})-\mathbf{L}\bar{\mathbf{x}}-\bar{\mathbf{v}}. \quad (5b)$$

Equation (5a) implies that $\bar{\mathbf{x}}$ belongs to the null-space of $\mathbf{L}$. For strongly connected digraphs we have rank$(\mathbf{L})=N-1$ and the null-space of $\mathbf{L}$ is spanned by $\mathbf{1}_N$. Thus,

$$\bar{\mathbf{x}}=\mathbf{1}_N\otimes\boldsymbol{\theta},\quad\boldsymbol{\theta}\in\mathbb{R}^d. \quad (6)$$

Multiplying (5b) by $\mathbf{1}_N^\top\otimes\mathbf{I}_d$ from the left and recalling (4), we obtain $\mathbf{0}=\sum_{i=1}^N\nabla f^i(\bar{\mathbf{x}}^i)$. Then, the optimality condition $\nabla f(\mathbf{x}^\star)=\mathbf{0}_d$ along with (6) and the fact that $\nabla f(\mathbf{x})=\sum_{i=1}^N\nabla f^i(\mathbf{x})$ imply

$$\bar{\mathbf{x}}^i=\mathbf{x}^\star,\quad i\in\{1,\dots,N\}.$$

Substituting this value in (5b), we obtain

$$\bar{\mathbf{v}}^i=-\alpha\nabla f^i(\mathbf{x}^\star),\quad i\in\{1,\dots,N\}. \quad (7)$$

Next, we study the stability of (1). First, we transfer the equilibrium point to the origin by means of

$$\mathbf{u}=\mathbf{v}-\bar{\mathbf{v}},\quad\mathbf{y}=\mathbf{x}-\bar{\mathbf{x}}. \quad (8)$$

Then, we apply the following change of variables

$$\mathbf{u}=([\mathbf{r}\quad\mathbf{R}]\otimes\mathbf{I}_d)\mathbf{w},\quad\mathbf{y}=([\mathbf{r}\quad\mathbf{R}]\otimes\mathbf{I}_d)\mathbf{z}, \quad (9)$$

where $\mathbf{r}=\frac{1}{\sqrt{N}}\mathbf{1}_N$ and $\mathbf{R}$ is such that $\mathbf{r}^\top\mathbf{R}=\mathbf{0}$ and $\mathbf{R}^\top\mathbf{R}=\mathbf{I}_{N-1}$. We partition the new variables as follows: $\mathbf{w}=(\mathbf{w}_1,\mathbf{w}_{2:N})$ and $\mathbf{z}=(\mathbf{z}_1,\mathbf{z}_{2:N})$, where $\mathbf{w}_1,\mathbf{z}_1\in\mathbb{R}^d$ and $\mathbf{w}_{2:N},\mathbf{z}_{2:N}\in\mathbb{R}^{(N-1)d}$. In these new variables, the algorithm (1) reads as

$$\begin{aligned}\dot{\mathbf{w}}_1&=\mathbf{0}_d,\\ \dot{\mathbf{w}}_{2:N}&=\alpha\beta(\mathbf{R}^\top\mathbf{L}\mathbf{R}\otimes\mathbf{I}_d)\mathbf{z}_{2:N},\\ \dot{\mathbf{z}}_1&=-\alpha(\mathbf{r}^\top\otimes\mathbf{I}_d)\mathbf{h},\\ \dot{\mathbf{z}}_{2:N}&=-\alpha(\mathbf{R}^\top\otimes\mathbf{I}_d)\mathbf{h}-\beta(\mathbf{R}^\top\mathbf{L}\mathbf{R}\otimes\mathbf{I}_d)\mathbf{z}_{2:N}-\mathbf{w}_{2:N},\end{aligned} \quad (10)$$

where

$$\mathbf{h}=\nabla\tilde{f}(\mathbf{y}+\bar{\mathbf{x}})-\nabla\tilde{f}(\bar{\mathbf{x}}). \quad (11)$$

Note that the first equation in (10) corresponds to the constant of motion (4). To study the stability in the other variables, consider the candidate Lyapunov function

$$\begin{aligned}V(\mathbf{z},\mathbf{w}_{2:N})&=\frac{1}{18}\alpha(\phi+1)\mathbf{z}_1{}^\top\mathbf{z}_1+\frac{\phi\alpha}{2}\mathbf{z}_{2:N}{}^\top\mathbf{z}_{2:N}\\&\quad+\frac{1}{2\alpha}(\alpha\mathbf{z}_{2:N}+\mathbf{w}_{2:N})^\top(\alpha\mathbf{z}_{2:N}+\mathbf{w}_{2:N}),\end{aligned} \quad (12)$$

with $\phi>0$ as in the statement. Note that $V(\mathbf{z},\mathbf{w}_{2:N})\leq\bar{\lambda}_{\mathbf{F}}\|\mathbf{p}\|^2$, where $\mathbf{p}=(\mathbf{z},\mathbf{w}_{2:N})$ and $\bar{\lambda}_{\mathbf{F}}$ is the maximum eigenvalue of

$$\mathbf{F}=\frac{1}{2}\begin{bmatrix}\frac{1}{9}\alpha(\phi+1)\mathbf{I}_d&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\alpha(\phi+1)\mathbf{I}_{(N-1)d}&\mathbf{I}_{(N-1)d}\\\mathbf{0}&\mathbf{I}_{(N-1)d}&\frac{1}{\alpha}\mathbf{I}_{(N-1)d}\end{bmatrix}. \quad (13)$$

The Lie derivative of V along (10) is given by

$$\begin{aligned}\dot{V}=&-\frac{1}{9}\alpha^2(\phi+1)\mathbf{y}^\top\mathbf{h}-\frac{7}{16}\mathbf{w}_{2:N}^\top\mathbf{w}_{2:N}\\&-\phi\alpha\beta\mathbf{z}_{2:N}{}^\top(\mathbf{R}^\top\,\mathrm{Sym}(\mathbf{L})\mathbf{R}\otimes\mathbf{I}_d)\mathbf{z}_{2:N}\\&+\frac{4}{9}\alpha^2\|(\mathbf{R}^\top\otimes\mathbf{I}_d)\mathbf{h}\|^2+\frac{4}{9}\alpha^2(1+\phi)^2\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N}\\&-\|\frac{3}{4}\mathbf{w}_{2:N}+\frac{2\alpha}{3}(\mathbf{R}^\top\otimes\mathbf{I}_d)\mathbf{h}+\frac{2\alpha}{3}(\phi+1)\mathbf{z}_{2:N})\|^2.\end{aligned}$$

Next, we show that under (3), $\dot{V}$ is negative definite. Note that $\|\mathbf{z}\|=\|\mathbf{y}\|$. Then, invoking the assumptions on the local cost functions in the statement, and using the

m_{T}-strongly convexity of $\tilde{f}$ and the M_{T}-Lipschitzness of $\nabla\tilde{f}$ along with $\|\mathbf{R}^\top \otimes \mathbf{I}_d\| = 1$, we have

$$\|\mathbf{y}^\top \mathbf{h}\| \geq m_{\mathrm{T}}\|\mathbf{y}\|^2 = m_{\mathrm{T}}\|\mathbf{z}\|^2, \tag{14a}$$
$$\|(\mathbf{R}^\top \otimes \mathbf{I}_d)\mathbf{h}\| \leq M_{\mathrm{T}}\|\mathbf{y}\| = M_{\mathrm{T}}\|\mathbf{z}\|. \tag{14b}$$

For a strongly connected digraph, we have $\hat{\lambda}_2 \mathbf{I} \leq \mathbf{R}^\top \operatorname{Sym}(\mathbf{L})\mathbf{R}$. Considering the relations above, we have

$$\begin{aligned}\dot{V} \leq &-\frac{\alpha^2(\phi+1)m_{\mathrm{T}}}{9}\mathbf{z}^\top\mathbf{z} - \frac{7}{16}\mathbf{w}_{2:N}^\top \mathbf{w}_{2:N} - \phi\alpha\beta\hat{\lambda}_2 \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} \\ &- \|\frac{3}{4}\mathbf{w}_{2:N} + \frac{2\alpha}{3}(\mathbf{R}^\top \otimes \mathbf{I}_d)\mathbf{h} + \frac{2\alpha(\phi+1)}{3}\mathbf{z}_{2:N})\|^2 \\ &+ \frac{4\alpha^2}{9}M_{\mathrm{T}}^2 \mathbf{z}^\top\mathbf{z} + \frac{4\alpha^2(1+\phi)^2}{9}\mathbf{z}_{2:N}^\top \mathbf{z}_{2:N}.\end{aligned}$$

Since $\mathbf{z}^\top\mathbf{z} = \mathbf{z}_1^\top\mathbf{z}_1 + \mathbf{z}_{2:N}^\top\mathbf{z}_{2:N}$, it follows that $\dot{V} < -\min\{\frac{7}{16}, \frac{1}{9}\gamma\}\|\mathbf{p}\|^2 < 0$, where γ is a shorthand notation for the expression in (3). Therefore, the convergence of $\mathbf{z} \to \mathbf{0}$ as $t \to \infty$, equivalently $\mathbf{x}^i \to \mathbf{x}^\star$, for all $i \in \{1, \ldots, N\}$, is exponential with rate no less than

$$\min\{\frac{7}{16}, \frac{1}{9}\gamma\}/(2\bar{\lambda}_{\mathbf{F}}). \tag{15}$$

(cf. [Khalil, 2002, Theorem 4.10]) □

In Theorem 1, note that the requirement $\sum_{i=1}^N \mathbf{v}^{\mathbf{i}}(0) = \mathbf{0}_d$ on the initial condition of the state $\mathbf{v}$ is trivially satisfied by each agent with the choice $\mathbf{v}^i(0) = \mathbf{0}_d$. This is an advantage with respect to the continuous-time coordination algorithms proposed in [Lu and Tang, 2012], which requires the nontrivial initialization $\sum_{i=1}^N \nabla f^i(\mathbf{x}^i(0)) = \mathbf{0}_d$, and in [Zanella et al., 2011], which requires the initialization on a state communicated among neighbors and is therefore subject to communication error.

Remark 2 *(Role of the design parameters in* (1)*):* We provide here several observations regarding the role of the design parameters α and β. First, note that there always exists β satisfying (3) (for example, any $\beta > 4(\phi+1)^2\alpha/(9\phi\hat{\lambda}_2)$). We have observed in simulation that (3) is only a sufficient condition, in fact, in all of our numerical examples, the algorithm (1) converges for any positive α and β. Although not evident in (15), one can expect that the larger α and β are, the higher the rate of convergence of the algorithm (1) is. A coefficient $\alpha > 1$ can be interpreted as a way of increasing the strong convexity coefficient of the local cost functions. A coefficient $\beta > 1$ can be interpreted as a means of increasing the graph connectivity. Our simulations have confirmed this conjecture. The relationship between these parameters and the rate of convergence of the algorithm (1) is more evident in the case of quadratic local cost functions $f^i(\mathbf{x}) = \frac{1}{2}(\mathbf{x}^\top\mathbf{x} + \mathbf{x}^\top\mathbf{a}^i + \mathbf{b}^i)$, $i \in \{1, \ldots, N\}$. In this case, the algorithm (1) is a linear time-invariant system where the eigenvalues of the system matrix are $-\alpha$, with multiplicity of Nd, and λ_i, $i \in \{1, \ldots, N\}$ (λ_i's are the eigenvalues of $\mathbf{L}$), with multiplicity d. Therefore, one can show that (1) converges regardless of the value of α, $\beta > 0$ with an exponential rate equal to $\min\{\alpha, \beta\Re(\lambda_2)\}$. •

Remark 3 (Semiglobal convergence of (1) *under local gradients that are locally Lipschitz):* The convergence result in Theorem 1 is semiglobal if the local gradients are only locally Lipschitz or, equivalently, Lipschitz on compact sets. In fact, one can see from the proof of the result that, for any compact set containing the initial conditions $\mathbf{x}^i(0) \in \mathbb{R}^d$ and $\mathbf{v}^i(0) = \mathbf{0}_d$, $i \in \{1, \ldots, N\}$, one can find $\phi > 0$ and $\beta > 0$ sufficiently large such that the compact set is contained in the region of attraction of the equilibrium point. •

Next, we study the convergence of (1) over dynamically changing, strongly connected, and weight-balanced digraphs with uniformly bounded and piecewise constant adjacency matrices. Since the proof of Theorem 1 relies on a Lyapunov function with no dependency on the system parameters and its derivative is upper bounded by a quadratic negative definite function, we can readily extend the convergence result to dynamically changing networks. The proof details are omitted for brevity.

Proposition 4 *(Convergence of* (1) *over dynamically changing interaction topologies): Let $\mathcal{G}$ be a time-varying digraph which is strongly connected and weight-balanced at all times and whose adjacency matrix is uniformly bounded and piecewise constant. Assume the local cost function f^i, $i \in \{1, \ldots, N\}$, is m^i-strongly convex, differentiable, and its gradient is M^i-Lipschitz on $\mathbb{R}^d$. Let $\beta > 0$ satisfy* (3) *with $\hat{\lambda}_2$ replaced by $(\hat{\lambda}_2)_{\min} = \min_{p\in\mathcal{P}}\{\hat{\lambda}_2(\mathbf{L}_p)\}$, where $\mathcal{P}$ is the index set of all possible realizations of $\mathcal{G}$. Then, for any $\alpha > 0$ and each $i \in \{1, \ldots, N\}$, the algorithm* (1) *over $\mathcal{G}$ makes $\mathbf{x}^i(t) \to \mathbf{x}^\star$ exponentially fast as $t \to \infty$, starting from initial conditions $\mathbf{x}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^d$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}_d$.*

We conclude this section by analyzing the privacy preservation properties of the coordination algorithm (1). More specifically, the next result characterizes the topological requirements on the communication graph and the knowledge about the algorithm's parameters and initial conditions that allow an agent to reconstruct the local gradients of other agents in the network.

Proposition 5 *(Privacy preservation under* (1)*): Let $\mathcal{G}$ be a strongly connected and weight balanced digraph. For $\alpha, \beta > 0$, consider any execution of the coordination algorithm* (1) *over $\mathcal{G}$ starting from $\mathbf{x}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^d$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}_d$. Then, an agent $i \in \{1, \ldots, N\}$ can reconstruct the local gradient of another agent $j \neq i$ only if j and all its out-neighbors are out-neighbors of i, and agent i knows $\mathbf{v}^j(0)$ and a_{jk}, $k \in \{1, \ldots, N\}$ (here we assume that the agent i is aware of the identity of neighbors of agent j and it has memory to save the time history of the data it receives from its out-neighbors).*

PROOF. Consider an arbitrary time $t^\star$. Let i be an in-neighbor of agent j and all of its out-neighbors. The algorithm (1) requires each agent to communicate its component of $\mathbf{x}$ to their in-neighbors. Since agent i has memory to save information it receives from its out-neighbors for all $t \leq t^\star$, it can use the time history of $\mathbf{x}^j(t)$ to numerically reconstruct $\dot{\mathbf{x}}^j(t^\star)$. Because i is the in-neighbor of j and its out-neighbors, it can use its knowledge of a_{jk}, $k \in \{1,\dots,N\}$ to reconstruct $\sum_{k=1}^{N} a_{jk}(\mathbf{x}^j(t) - \mathbf{x}^k(t))$ for all $t \leq t^\star$. Agent i can reconstruct $\mathbf{v}^j(t)$ from (1a) uniquely as it knows $\mathbf{v}^j(0)$. Then, agent i has all the elements to solve for $\nabla f^j(\mathbf{x}^j(t^\star))$ in (1b). The lack of knowledge about any of this information would prevent the agent i from reconstructing exactly the local gradient of agent j. □

The requirements of Proposition 5 are trivially satisfied when agent i is aware that it is the only out-neighbor of j and all agents know that the algorithm is initialized with $\mathbf{v}^k(0) = \mathbf{0}_d$, for all $k \in \{1,\dots,N\}$.

4.2 Connected Undirected Graphs

Here, we study the convergence of the algorithm (1) over connected undirected graph topologies. While the results of the previous section are of course valid for these topologies, here using the structural properties of the Laplacian matrix in the undirected case we establish the convergence of (1) for a larger family of local cost functions. In doing so, we are also able to analytically establish convergence for any $\alpha, \beta > 0$, as we show next.

Theorem 6 *(Exponential convergence of* (1) *over connected graphs): Let $\mathcal{G}$ be a connected graph. Assume the local cost function f^i, $i \in \{1,\dots,N\}$, is m^i-strongly convex and differentiable on $\mathbb{R}^d$, and its gradient is locally Lipschitz. Then, for any $\alpha, \beta > 0$ and each $i \in \{1,\dots,N\}$, the algorithm* (1) *over $\mathcal{G}$ satisfies $\mathbf{x}^i(t) \to \mathbf{x}^\star$ exponentially fast as $t \to \infty$, starting from initial conditions $\mathbf{x}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^d$ with $\sum_{j=1}^{N} \mathbf{v}^i(0) = \mathbf{0}$.*

PROOF. We use the equivalent representation (10) of the algorithm (1) obtained in the proof of Theorem 1. Consider the following candidate Lyapunov function

$$\begin{aligned} V(\mathbf{z}, \mathbf{w}_{2:N}) = &\frac{1}{2}\alpha(\phi+1)\mathbf{z}_1^\top \mathbf{z}_1 + \frac{\phi\alpha}{2}\mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} \\ &+ \frac{1}{2\alpha}(\alpha \mathbf{z}_{2:N} + \mathbf{w}_{2:N})^\top(\alpha \mathbf{z}_{2:N} + \mathbf{w}_{2:N}) \\ &+ \frac{1}{2\beta}(\phi+1)\mathbf{w}_{2:N}^\top\big((\mathbf{R}^\top \mathbf{L}\mathbf{R})^{-1} \otimes \mathbf{I}_d\big)\mathbf{w}_{2:N}, \end{aligned} \quad (16)$$

where $\phi \geq 1$ is defined below. For a connected graph, we have $\mathbf{R}^\top \mathbf{L}\mathbf{R} > 0$. Thus, V is radially unbounded and positive definite and its value can be upper bounded by $V(\mathbf{z}, \mathbf{w}_{2:N}) \leq \bar{\lambda}_{\mathbf{E}}\|\mathbf{p}\|^2$, where $\mathbf{p} = (\mathbf{z}, \mathbf{w}_{2:N})$ and $\bar{\lambda}_{\mathbf{E}} > 0$ is the maximum eigenvalue of

$$\mathbf{E} = \frac{1}{2}\begin{bmatrix} \frac{\alpha(\phi+1)}{9}\mathbf{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \alpha(\phi+1)\mathbf{I}_{(N-1)d} & \mathbf{I}_{(N-1)d} \\ \mathbf{0} & \mathbf{I}_{(N-1)d} & \frac{1}{\alpha}\mathbf{I} + \frac{1}{\beta}(\mathbf{R}^\top \mathbf{L}\mathbf{R})^{-1} \otimes \mathbf{I}_d \end{bmatrix}.$$

The Lie derivative of V along the dynamics (10) is

$$\begin{aligned} \dot{V} = &-\alpha^2(\phi+1)\mathbf{y}^\top \mathbf{h} - \phi\alpha\beta \mathbf{z}_{2:N}{}^\top(\mathbf{R}^\top \mathbf{L}\mathbf{R} \otimes \mathbf{I}_d)\mathbf{z}_{2:N} \\ &- \mathbf{w}_{2:N}{}^\top \mathbf{w}_{2:N} - \alpha \mathbf{w}_{2:N}{}^\top(\mathbf{R}^\top \otimes \mathbf{I}_d)\mathbf{h}. \end{aligned}$$

In the following, we show that $\dot{V}$ is upper bounded by a negative definite quadratic function. We start by identifying a compact set whose definition is independent of ϕ and contains the set $\mathcal{S}_0 = \{\mathbf{z} \in \mathbb{R}^{Nd} \mid V(\mathbf{z}, \mathbf{w}_{2:N}) \leq V(\mathbf{z}(0), \mathbf{w}_{2:N}(0))\}$. For any given initial condition, let $\rho_0 = \frac{1}{2}\alpha \mathbf{z}_1(0)^\top \mathbf{z}_1(0) + \frac{(\alpha+1)}{2}\mathbf{z}_{2:N}(0)^\top \mathbf{z}_{2:N}(0) + (\frac{1}{2\beta\lambda_2} + \frac{1}{2\alpha} + \frac{1}{2})\mathbf{w}_{2:N}(0)^\top \mathbf{w}_{2:N}(0)$ and define $\bar{\mathcal{S}}_0 = \{\mathbf{z} \in \mathbb{R}^{Nd} \mid \frac{1}{2}\alpha \mathbf{z}_1^\top \mathbf{z}_1 + \frac{1}{4}\alpha \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} \leq \rho_0\}$. Observe that this set is compact. Note that $(\mathbf{z}, \mathbf{w}_{2:N}) \in \mathcal{S}_0$ implies $\mathbf{z} \in \bar{\mathcal{S}}_0$ because $V(\mathbf{z}(0), \mathbf{w}_{2:N}(0)) \leq (\phi+1)\rho_0$ and

$$\begin{aligned} &\frac{1}{2}\alpha(\phi+1)\mathbf{z}_1^\top \mathbf{z}_1 + \frac{1}{4}\alpha(\phi+1)\mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} \leq \\ &\frac{1}{2}\alpha(\phi+1)\mathbf{z}_1^\top \mathbf{z}_1 + \frac{1}{2}\alpha\phi \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} \leq V(\mathbf{z}, \mathbf{w}_{2:N}), \end{aligned}$$

where we have used $\phi \geq 1$ in the first inequality. Since the change of variables (8) and (9) are linear, the corresponding $\mathbf{x}$ and $\mathbf{y}$ for $\mathbf{z} \in \bar{\mathcal{S}}_0$ also belong to compact sets as well. Then, the assumption on the gradients of the local cost functions implies that there exists $M_0 > 0$ such that

$$\|\mathbf{h}\| \leq M_0\|\mathbf{y}\| = M_0\|\mathbf{z}\|, \quad \forall \mathbf{z} \in \bar{\mathcal{S}}_0.$$

Consequently,

$$\|(\mathbf{R}^\top \otimes \mathbf{I}_d)\mathbf{h}\|^2 \leq M_0^2\|\mathbf{z}\|^2, \quad \forall (\mathbf{z}, \mathbf{w}_{2:N}) \in \mathcal{S}_0.$$

Then we can show $\alpha \mathbf{w}_{2:N}{}^\top(\mathbf{R}^\top \otimes \mathbf{I}_d)\mathbf{h} \leq \frac{1}{2}\mathbf{w}_{2:N}^\top \mathbf{w}_{2:N} + \frac{1}{2}\alpha^2 M_0^2 \mathbf{z}^\top \mathbf{z}$ for $(\mathbf{z}, \mathbf{w}_{2:N}) \in \mathcal{S}_0$. Using this inequality, the fact that the local cost functions are m^i-strongly convex (and hence (14a) holds) and $\mathbf{R}^\top \mathbf{L}\mathbf{R} \geq \lambda_2 \mathbf{I}_{N-1}$, we deduce

$$\begin{aligned} \dot{V} \leq &-\alpha^2(\phi+1)m_{\mathrm{T}}(\mathbf{z}_1^\top \mathbf{z}_1 + \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N}) - \phi\alpha\beta\lambda_2 \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} \\ &- \frac{1}{2}\mathbf{w}_{2:N}^\top \mathbf{w}_{2:N} + \frac{1}{2}\alpha^2 M_0^2(\mathbf{z}_1^\top \mathbf{z}_1 + \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N}). \end{aligned}$$

Let $\phi + 1 = \frac{1}{2m_{\mathrm{T}}}M_0^2 + \frac{1}{2m_{\mathrm{T}}\alpha^2}\delta_0$, where $\delta_0 > 0$ is such that $\phi \geq 1$ (since M_0 does not depend on ϕ, this choice is always feasible). Then,

$$\dot{V} \leq -\frac{1}{2}\min\{1, \delta_0\}\|\mathbf{p}\|^2. \quad (17)$$

The Lyapunov function (16) then satisfies all the conditions of [Khalil, 2002, Theorem 4.10], for the dynamics (10). Therefore, the convergence of $\mathbf{z} \to \mathbf{0}$ as $t \to \infty$, equivalently $\mathbf{x}^i \to \mathbf{x}^\star$, for all $i \in \{1, \dots, N\}$, is exponential. For any given initial condition the rate of convergence, at least, is $\frac{1}{4}(\min\{1, \delta_0\})/\bar{\lambda}_{\mathbf{E}} > 0$. □

Regarding Theorem 6, note that the requirement that ∇f^i is locally Lipschitz is trivially satisfied if f^i is twice differentiable. The Lyapunov function (16) identified in the proof of this result plays a key role later in our study of the algorithm implementation with discrete-time communication in Section 5. One can also see from the proof of Theorem 6 that the guaranteed rate of convergence is not uniform, unless the local gradients are globally Lipschitz. In this case, one recovers the result in Theorem 1 but for arbitrary α, $\beta > 0$.

Next, we study the convergence of the algorithm (1) over connected graphs when the local cost functions are only convex. Here, the lack of strong convexity makes us rely on a LaSalle function, rather than on a Lyapunov one, to establish asymptotic convergence to the optimizer.

Theorem 7 *(Asymptotic convergence of* (1) *over connected graphs): Let $\mathcal{G}$ be a connected graph. Assume the local cost function f^i, $i \in \{1, \dots, N\}$, is convex and differentiable on $\mathbb{R}^d$, and the global cost function f is strictly convex and differentiable on $\mathbb{R}^d$. Then, for any $\alpha, \beta > 0$ and each $i \in \{1, \dots, N\}$, the algorithm* (1) *over $\mathcal{G}$ satisfies $\mathbf{x}^i(t) \to \mathbf{x}^\star$ as $t \to \infty$, starting from any initial conditions $\mathbf{x}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^d$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}_d$.*

PROOF. We use again the equivalent representation (10) of the algorithm (1) obtained in the proof of Theorem 1. To study the stability of this system, consider the following candidate Lyapunov function

$$V(\mathbf{z}, \mathbf{w}_{2:N}) = \frac{1}{2}\mathbf{z}^\top \mathbf{z} + \frac{1}{2\alpha\beta}\mathbf{w}_{2:N}{}^\top \Big((\mathsf{R}^\top \mathsf{L} \mathsf{R})^{-1} \otimes \mathsf{I}_d\Big)\mathbf{w}_{2:N}.$$

For a connected graph, we have $\mathsf{R}^\top \mathsf{L}\mathsf{R} > 0$. Therefore, the function above is positive definite and radially unbounded. The Lie derivative of V along (10) is given by

$$\begin{aligned}\dot{V} &= -\alpha \mathbf{y}^\top (\nabla_{\mathrm{T}} f(\mathbf{y} + \bar{\mathbf{x}}) - \nabla_{\mathrm{T}} f(\bar{\mathbf{x}})) - \beta \mathbf{y}^\top \mathbf{L}\mathbf{y} \\ &= -\alpha \sum_{i=1}^N {\mathbf{y}^i}^\top (\nabla f^i(\mathbf{y}^i + \mathbf{x}^\star) - \nabla f^i(\mathbf{x}^\star)) - \beta \mathbf{y}^\top \mathbf{L}\mathbf{y}.\end{aligned}$$

To obtain the second summand, we have used $\mathbf{z}_{2:N} = (\mathsf{R}^\top \otimes \mathsf{I}_d)\mathbf{y}$ and $\mathsf{R}\mathsf{R}^\top = \mathsf{I}_N - \mathsf{r}\mathsf{r}^\top$. Since the local cost functions are convex, the first summand of $\dot{V}$ is non-positive for all $\mathbf{y}$. Because the topology is undirected, the second summand is non-positive as well, and because the graph is connected, the null-space of this summand is spanned by $\boldsymbol{\theta} \otimes \mathbf{1}_N$, $\boldsymbol{\theta} \in \mathbb{R}^d$. On this null-space, the first summand becomes

$$-\alpha \boldsymbol{\theta}^\top \sum_{i=1}^N (\nabla f^i(\boldsymbol{\theta} + \mathbf{x}^\star) - \nabla f^i(\mathbf{x}^\star)). \qquad (18)$$

Since $\nabla f(\mathbf{x}) = \sum_{i=1}^N \nabla f^i(\mathbf{x})$ and the global cost function is strictly convex by assumption, (18) can only be zero when $\boldsymbol{\theta} = \mathbf{0}$. Then, the two summands of $\dot{V}$ can be zero simultaneously only when $\mathbf{y}^i = \mathbf{0}$, for all $i \in \{1, \dots, N\}$, which is equivalent to $\mathbf{z} = \mathbf{0}$. Thus, $\dot{V}$ is negative semidefinite, with $\dot{V}(\mathbf{z}, \mathbf{w}_{2:N}) = 0$ happening on the set $\mathcal{S} = \{(\mathbf{z}, \mathbf{w}_{2:N}) \in \mathbb{R}^{Nd} \times \mathbb{R}^{(N-1)d} \mid \mathbf{z} = \mathbf{0}\}$. Note that (10) on $\mathcal{S}$ reduces to $\dot{\mathbf{w}}_{2:N} = \mathbf{0}$, $\dot{\mathbf{z}}_1 = \mathbf{0}$, and $\dot{\mathbf{z}}_{2:N} = -\mathbf{w}_{2:N}$. Therefore, the only trajectory of (10) that remains in $\mathcal{S}$ is the equilibrium point ($\mathbf{z}_1 = \mathbf{0}, \mathbf{z}_{2:N} = \mathbf{0}, \mathbf{w}_{2:N} = \mathbf{0}$). The LaSalle invariance principle (cf. [Khalil, 2002, Theorem 4.4 and Corollary 4.2]) now implies that the equilibrium is globally asymptotically stable or, in other words, $\mathbf{x}^i \to \mathbf{x}^\star$, $i \in \{1, \dots, N\}$ globally asymptotically. □

Remark 8 (Simplification of (1) *for strictly convex local cost functions):* Using the LaSalle function identified in the proof of Theorem 7, one can show that the algorithm

$$\begin{aligned}\dot{\mathbf{v}}^i &= \sum_{j=1}^N \mathsf{a}_{ij}(\mathbf{x}^i - \mathbf{x}^j), \\ \dot{\mathbf{x}}^i &= -\nabla f^i(\mathbf{x}^i) - \mathbf{v}^i,\end{aligned}$$

over a connected graph is also guaranteed to asymptotically converge to the optimizer starting from any initial conditions $\mathbf{x}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^d$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}_d$ if the local cost functions are strictly convex. •

5 Continuous-time Evolution with Discrete-Time Communication

Here, we investigate the design of continuous-time coordination algorithms with discrete-time communication to solve the distributed optimization problem stated in Section 3. The implementation of (1) requires continuous-time communication among the agents. While this abstraction is useful for analysis, in practical scenarios the communication service is only available at discrete instants of time. This observation motivates our study here. Throughout the section, we deal with communication topologies described by connected undirected graphs. Our results here build on the discussion of Section 4, particularly the identification of Lyapunov functions for asymptotic stability.

We start by introducing some useful conventions. At any given time $t \in \mathbb{R}_{\geq 0}$, let $\hat{\mathbf{x}}^j$ be the last known state of agent $j \in \{1, \dots, N\}$ transmitted to its in-neighbors. If $\{t_k^i\} \subset \mathbb{R}_{\geq 0}$ denotes the times at which agent i communicates with its in-neighbors, then one has $\hat{\mathbf{x}}^i = \mathbf{x}^i(t_k^i)$ for $t \in [t_k^i, t_{k+1}^i)$. Consider the next implementation of the algorithm (1) with discrete-time communication,

$$\dot{\mathbf{v}}^i = \alpha\beta \sum_{j=1}^{N} \mathsf{a}_{ij}(\hat{\mathbf{x}}^i - \hat{\mathbf{x}}^j), \tag{20a}$$

$$\dot{\mathbf{x}}^i = -\alpha\nabla f^i(\mathbf{x}^i) - \beta \sum_{j=1}^{N} \mathsf{a}_{ij}(\hat{\mathbf{x}}^i - \hat{\mathbf{x}}^j) - \mathbf{v}^i. \tag{20b}$$

Clearly, the evolution of (20) depends on the sequences of communication times for each agent. Here, we consider three scenarios. Section 5.1 studies periodic communication schemes where all agents communicate synchronously at Δ intervals of time, i.e., $t_k^i = t_k = \Delta k$ for all $i \in \{1, \dots, N\}$. We provide a characterization of the periods that guarantee the asymptotic convergence of (20) to the optimizer. In general, periodic schemes might result in a wasteful use of the communication resources because of the need to account for worst-case situations in determining appropriate periods. This motivates our study in Section 5.2 of event-triggered communication schemes that tie the communication times to the network state for greater efficiency. We discuss two event-triggered communication implementations, a centralized synchronous one and a distributed asynchronous one. In both cases, we pay special attention to ruling out the presence of Zeno behavior (the existence of an infinite number of updates in a finite interval of time).

5.1 Periodic Communication

The following result provides an upper bound on the size of admissible stepsizes for the execution of (20) over connected graphs with periodic communication schemes.

Theorem 9 *(Convergence of* (20) *with periodic communication): Let $\mathcal{G}$ be a connected graph. Assume the local cost function f^i, $i \in \{1, \dots, N\}$, is m^i-strongly convex, differentiable, and its gradient is M^i-Lipschitz on $\mathbb{R}^d$. Given $\alpha, \beta > 0$, consider an implementation of the algorithm* (20) *with agents communicating over $\mathcal{G}$ synchronously every Δ seconds starting at $t_1 = 0$, i.e., $t_k^i = t_k = \Delta k$ for all $i \in \{1, \dots, N\}$. Let $0 < \epsilon < 1$ and $\delta > 0$ such that*

$$\phi + 1 = \frac{1}{2m_T} M_T^2 + \frac{1}{2m_T\alpha^2}\delta > 1, \tag{21}$$

where M_T and m_T are given in the statement of Theorem 1, and define

$$\tau = \frac{1}{\alpha M_T + 1} \ln\Big(1 + \frac{(\alpha M_T + 1)\zeta}{\alpha M_T + 1 + \beta\lambda_N\sqrt{1+\alpha^2}(1+\zeta)}\Big), \tag{22}$$

where $\zeta^2 = \frac{2\epsilon\lambda_2 \min\{1-\epsilon,\delta\}}{\alpha\beta\lambda_N^2\phi + 2\alpha^2\lambda_2(1+\phi)^2}$. Then, if $\Delta \in (0, \tau)$, the algorithm evolution starting from initial conditions $\mathbf{x}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^d$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}_d$ makes $\mathbf{x}^i(t) \to \mathbf{x}^\star$ exponentially fast as $t \to \infty$, for all $i \in \{1, \dots, N\}$.

PROOF. We start by transferring the equilibrium point to the origin using (8) and then apply the change of variables (9) to write (20) as

$$\begin{aligned}
\dot{\mathbf{w}}_1 &= \mathbf{0}_d, & (23a)\\
\dot{\mathbf{w}}_{2:N} &= \alpha\beta(\mathbf{R}^\top\mathbf{L}\mathbf{R} \otimes \mathbf{I}_d)(\mathbf{z}_{2:N} + \tilde{\mathbf{z}}_{2:N}), & (23b)\\
\dot{\mathbf{z}}_1 &= -\alpha(\mathbf{r}^\top \otimes \mathbf{I}_d)\mathbf{h}, & (23c)\\
\dot{\mathbf{z}}_{2:N} &= -\alpha(\mathbf{R}^\top \otimes \mathbf{I}_d)\mathbf{h} & (23d)\\
&\quad - \beta(\mathbf{R}^\top\mathbf{L}\mathbf{R} \otimes \mathbf{I}_d)(\mathbf{z}_{2:N} + \tilde{\mathbf{z}}_{2:N}) - \mathbf{w}_{2:N},
\end{aligned}$$

where $\tilde{\mathbf{z}}_{2:N}(t) = \mathbf{z}_{2:N}(t_k) - \mathbf{z}_{2:N}(t)$, for $t \in [t_k, t_{k+1})$, and $\mathbf{h}$ is given by (11). To study the stability of (23b)-(23d), consider the candidate Lyapunov function (16) with ϕ satisfying (21). Its Lie derivative can be bounded by

$$\begin{aligned}
\dot{V} \leq &-\frac{1}{2}\delta(\mathbf{z}_1^\top\mathbf{z}_1 + \mathbf{z}_{2:N}^\top\mathbf{z}_{2:N}) & (24)\\
&- \frac{1}{2}(1-\epsilon)\mathbf{w}_{2:N}^\top\mathbf{w}_{2:N} - \phi\alpha\beta\lambda_2(1-\epsilon)\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N}\\
&+ \frac{1}{4\epsilon\lambda_2}\phi\alpha\beta\lambda_N^2\tilde{\mathbf{z}}_{2:N}^\top\tilde{\mathbf{z}}_{2:N} + \frac{1}{2\epsilon}\alpha^2(\phi+1)^2\tilde{\mathbf{z}}_{2:N}^\top\tilde{\mathbf{z}}_{2:N},\\
\leq &-\phi\alpha\beta\lambda_2(1-\epsilon)\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N} & (25)\\
&- \frac{1}{2}\min\{\delta, 1-\epsilon\}\big(\mathbf{p}^\top\mathbf{p} - \zeta^2\tilde{\mathbf{z}}_{2:N}^\top\tilde{\mathbf{z}}_{2:N}\big),
\end{aligned}$$

where $\mathbf{p} = (\mathbf{z}, \mathbf{w}_{2:N})$ and ϵ and ζ are given in the theorem's statement. Observe that at each communication time t_k, $\|\tilde{\mathbf{z}}_{2:N}(t_k)\| = 0$, then, it grows until next communication at time t_{k+1} when it becomes zero again. Our proof proceeds by showing that if $t_{k+1} < t_k + \tau$, where τ is given in (22), then we have the guarantee that

$$\|\tilde{\mathbf{z}}_{2:N}(t)\| < \zeta\,\|\mathbf{p}(t)\|, \quad t \in [t_k, t_{k+1}), \tag{26}$$

(note that, from (24), this guarantee ensures that $\dot{V}$ is negative definite for all $t \geq 0$). To this end, we study the dynamics of $q = \|\tilde{\mathbf{z}}_{2:N}\|/\|\mathbf{p}\|$ and find a lower bound on the time that it takes for q to evolve from zero (recall $\tilde{\mathbf{z}}_{2:N}(t_k) = \mathbf{0}$) to ζ. Notice

$$\begin{aligned}
&\dot{q} \leq (1+q)\frac{\|\dot{\mathbf{p}}\|}{\|\mathbf{p}\|} \leq (1+q)\times\\
&\frac{(\alpha M_T + \beta\lambda_N\sqrt{1+\alpha^2})\|\mathbf{p}\| + \|\mathbf{w}_{2:N}\| + \beta\lambda_N\sqrt{1+\alpha^2}\|\tilde{\mathbf{z}}_{2:N}\|}{\|\mathbf{p}\|}\\
&\leq (\alpha M_T + 1)(1+q) + \beta\lambda_N\sqrt{1+\alpha^2}(1+q)^2.
\end{aligned}$$

Here, we have used $d/dt(\tilde{\mathbf{z}}_{2:N}) = -\dot{\mathbf{z}}_{2:N}$, $\|\dot{\mathbf{z}}_{2:N}\| \leq \|\dot{\mathbf{p}}\|$, the evolution of $\dot{\mathbf{p}}$ given in (23b)-(23d), $\|\mathbf{R}^\top\mathbf{L}\mathbf{R}\| \leq \|\mathbf{L}\| = \lambda_N$ and

$$\left\|\begin{bmatrix} -\beta\mathbf{R}^\top\mathbf{L}\mathbf{R} \\ \alpha\beta\mathbf{R}^\top\mathbf{L}\mathbf{R}\end{bmatrix}\right\| = \beta\lambda_N\sqrt{1+\alpha^2}.$$

Using the Comparison Lemma (cf. [Khalil, 2002, Lemma 3.4]), we conclude that $q(t, q_0) \leq \psi(t, \psi_0)$, where $\psi(t, \psi_0)$ is the solution of $\dot{\psi} = (\alpha M_{\mathrm{T}} + 1)(1 + \psi) + \beta\lambda_N\sqrt{1+\alpha^2}(1+\psi)^2$ satisfying $\psi(0, \psi_0) = \psi_0$. Then,

$$
\begin{aligned}
q(t, 0) &\leq \psi(t, 0) \\
&= \frac{(\alpha M_{\mathrm{T}} + 1 + \beta\lambda_N\sqrt{1+\alpha^2})(\mathrm{e}^{\alpha M_{\mathrm{T}} t + t} - 1)}{-\beta\lambda_N\sqrt{1+\alpha^2}\mathrm{e}^{\alpha M_{\mathrm{T}} t + t} + \alpha M_{\mathrm{T}} + 1 + \beta\lambda_N\sqrt{1+\alpha^2}}.
\end{aligned}
$$

The time τ for $\psi(\tau, 0) = \zeta$ is given by (22). Then, for $\{t_{k+1} - t_k\}_{k\in\mathbb{N}} < \tau$, we have (26), and as a result $\dot{V} < 0$. Thus, $\mathbf{z} \to \mathbf{0}$, as $t \to \infty$, which is equivalent to $\mathbf{x}^i \to \mathbf{x}^\star$ as $t \to \infty$. The Lyapunov function (16) and the negative definite upper bound on its derivative are quadratic, therefore, the convergence is exponential. □

Remark 10 *(Dependence of the communication period on the design parameters):* It is interesting to note that the value of τ in Theorem 9 depends on the graph topology, the parameters of the local cost functions, the algorithm design parameters α and β, and the variables ϵ and δ. One can use this dependency to maximize the value of τ. Notice that the argument of ln(.) in (22) is a monotonically increasing function of $\zeta > 0$. Therefore, the smaller the value of β, the larger the value of τ. However, the dependency of τ on the rest of the parameters listed above is more complex. For given local cost functions, fixed network topology and fixed values of α, β, the maximum value of ζ is when $\phi + 1$ is at its minimum and $\epsilon\lambda_2 \min\{1-\epsilon, \delta\}$ is at its maximum. •

5.2 Event-Triggered Communication

This section studies the design of event-triggered communication schemes for the execution of the coordination algorithm (20). In contrast to periodic schemes, event-triggered implementations tie the determination of the communication times to the current network state, resulting in a more efficient use of the resources.

Interestingly, the proof of Theorem 9 reveals that the satisfaction of condition (26) guarantees the monotonic evolution of the Lyapunov function, which in turn ensures the correct asymptotic behavior of the algorithm. One could therefore specify when communication should occur by determining the times when this condition is not satisfied. There is, however, a serious drawback to this approach, that has to do with the fact that we are dealing with an optimization problem: the evaluation of the condition (26) requires the knowledge of the global minimizer $\mathbf{x}^\star$, which is of course not available. To see this, note that

$$\|\tilde{\mathbf{z}}_{2:N}\| = \|\mathbf{\Pi}(\mathbf{x}(t_k) - \mathbf{x})\|, \tag{27a}$$

$$\|\mathbf{p}\| = \sqrt{\|\mathbf{x} - \bar{\mathbf{x}}\|^2 + \|\mathbf{\Pi}(\mathbf{v} - \bar{\mathbf{v}})\|^2}, \tag{27b}$$

where we have used $\mathbf{R}^\top\mathbf{r} = \mathbf{0}$, $\mathbf{R}\mathbf{R}^\top = \mathbf{\Pi}_N = \mathbf{\Pi}_N^2$, and (4) (recall $\mathbf{\Pi} = \mathbf{\Pi}_N \otimes \mathbf{I}_d$). From (7), $\bar{\mathbf{v}}^i = -\alpha\nabla f^i(\mathbf{x}^\star)$ for $i \in \{1, \ldots, N\}$, and therefore the evaluation of the triggering condition (26) requires knowledge of the global optimizer. Our forthcoming discussion illustrates how one can get around this problem. We first consider the design of centralized triggers that require global knowledge and then move on to distributed triggering schemes that only require agent communication with neighbors.

5.2.1 Centralized Synchronous Implementation

In this section, we present a centralized event-triggered scheme to determine the sequence of synchronous communication times in (20). Our discussion builds upon the examination of the Lie derivative of the Lyapunov function used in the proof of Theorem 9 and the observations made above regarding the lack of knowledge of the solution $\mathbf{x}^\star$ of the optimization problem.

From (27), we see that an event-triggered law should not employ $\mathbf{p}$, but rather rely on $\tilde{\mathbf{z}}_{2:N}$ and $\mathbf{z}_{2:N}$, in order to be independent of $\mathbf{x}^\star$. With this in mind, the examination of the upper bound (24) on $\dot{V}$ reveals that, if

$$
\begin{aligned}
\|\tilde{\mathbf{z}}_{2:N}(t)\|^2 &= \|\mathbf{\Pi}(\mathbf{x}(t_k) - \mathbf{x}(t))\|^2 \\
&\leq \kappa\|\mathbf{z}_{2:N}(t)\|^2 = \kappa\|\mathbf{\Pi}\mathbf{x}(t)\|^2,
\end{aligned} \tag{28}
$$

where κ is shorthand notation for

$$\kappa = 2\frac{\epsilon\delta\lambda_2 + 2\phi\alpha\beta\lambda_2^2\epsilon^2(1-\epsilon)}{\alpha\beta\phi\lambda_N^2 + 2\lambda_2\alpha^2(1+\phi)^2}, \tag{29}$$

(here $0 < \epsilon < 1$ and ϕ is given by (21)), then we have

$$
\begin{aligned}
\dot{V} \leq &-\frac{1}{2}\delta\mathbf{z}_1^\top\mathbf{z}_1 - \phi\alpha\beta\lambda_2(1-\epsilon)^2\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N} \\
&-\frac{1}{2}(1-\epsilon)\mathbf{w}_{2:N}^\top\mathbf{w}_{2:N} < 0.
\end{aligned} \tag{30}
$$

Then, we can reproduce the proof of Theorem 9 and conclude the exponential convergence of the algorithm executions to the solution of the optimization problem.

According to the above discussion, the sequence of synchronous communication times $\{t_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_{\geq 0}$ for (20) should be determined by (28). However, for a truly implementable law, one should guarantee that no Zeno behavior occurs, i.e., the sequence of times does not have any finite accumulation point. However, observing (28), one can see that Zeno behavior will arise at least near the agreement surface $\mathbf{\Pi}\mathbf{x} = \mathbf{0}_{dN}$. The following result details how we address this problem to design a Zeno-free centralized event-triggered communication law.

Theorem 11 (Convergence of (20) *with Zeno-free centralized event-triggered communication): Let $\mathcal{G}$ be a connected graph. Assume the local cost function f^i, $i \in \{1, \ldots, N\}$, is m^i-strongly convex, differentiable, and its gradient is M^i-Lipschitz on $\mathbb{R}^d$. Consider an implementation of the algorithm* (20) *with agents communicating over $\mathcal{G}$ synchronously at times $\{t_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_{\geq 0}$,* *starting at $t_1 = 0$, determined by*

$$t_{k+1} = \operatorname{argmax}\{t \in [t_k + \tau, \infty) \mid \|\boldsymbol{\Pi}(\mathbf{x}(t_k) - \mathbf{x}(t))\|^2 \leq \kappa \|\boldsymbol{\Pi}\mathbf{x}(t)\|^2\}, \quad (31)$$

where τ and $\kappa < 1$ are defined in (22) *and* (29)*, respectively. Then, for any given $\alpha, \beta > 0$ and each $i \in \{1, \ldots, N\}$, the algorithm evolution starting from initial conditions $\mathbf{x}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^d$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}_d$ makes $\mathbf{x}^i(t) \to \mathbf{x}^\star$ exponentially fast as $t \to \infty$.*

PROOF. The proof is a consequence of our discussion above. We start by showing that $\kappa < 1$. This is an important property guaranteeing that, if agents start in agreement at a point other than the optimizer $\mathbf{x}^\star$, then the condition (28) is eventually violated, enforcing information updates. Notice that (a) $4\epsilon^2(1-\epsilon)\lambda_2^2 < \lambda_N^2$ and (b) $\epsilon\delta < \alpha^2(1+\phi)^2$ imply that the numerator in (29) is smaller than its denominator, and hence $\kappa < 1$. (a) follows from noting that the maximum of $4\epsilon^2(1-\epsilon)$ for $\epsilon \in (0,1)$ is $16/27 < 1$ and the fact that $\lambda_2 \leq \lambda_N$. We prove (b) reasoning by contradiction. Assume $\delta > \alpha^2(1+\phi)^2$ or equivalently $\alpha(1+\phi) - \sqrt{\delta} < 0$. Using (21) and multiplying both sides of the inequality by $2m_{\mathrm{T}}\alpha$, we obtain

$$\alpha^2 M_{\mathrm{T}}^2 + \delta - 2\alpha m_{\mathrm{T}}\sqrt{\delta} = (\sqrt{\delta} - \alpha m_{\mathrm{T}})^2 + \alpha^2(M_{\mathrm{T}}^2 - m_{\mathrm{T}}^2) < 0,$$

which, since $M_{\mathrm{T}} \geq m_{\mathrm{T}}$, is a contradiction. Having established the consistency of (28), consider now the candidate Lyapunov function V defined in (16) and let t_k be the last time at which a communication among all neighboring agents occurred. From the proof of Theorem 9, we know that the time derivative of V is negative, $\dot{V} < 0$ as long as $t < t_k + \tau$. After this time, (30) above shows that as long as (28) is satisfied, $\dot{V}$ remains negative. □

Interestingly, given that (28) does not use the full state of the network but rather relies on the computation of disagreement, one can interpret it as an output feedback event-triggered controller. Guaranteeing the existence of lower bounded inter-execution times for such controllers is in general a difficult problem, see e.g., [Donkers and Heemels, 2012]. Augmenting (28) with the condition $t_{k+1} \geq t_k + \tau$ results in Zeno-free executions by lower bounding the inter-event times by τ. The knowledge of this value also allows a designer to compute bounds on the maximum energy spent by the network on communication during any given time interval.

5.2.2 Distributed Asynchronous Implementation

Here, we present a distributed event-triggered scheme for determining the sequence of communication times in (20). At each agent, the execution of the distributed communication scheme depends only on local variables and the triggered states received from its neighbors. This naturally results in asynchronous communication. We also show that the resulting executions are free from Zeno behavior.

Theorem 12 (Convergence of (20) *with Zeno-free distributed event-triggered communication): Let $\mathcal{G}$ be a connected graph. Assume the local cost function f^i, $i \in \{1, \ldots, N\}$, is m^i-strongly convex, differentiable, and its gradient is M^i-Lipschitz on $\mathbb{R}^d$. For $\boldsymbol{\epsilon} \in \mathbb{R}^n_{>0}$, consider an implementation of the algorithm* (20) *where agent $i \in \{1, \ldots, N\}$ communicates with its neighbors in $\mathcal{G}$ at times $\{t_k^i\}_{k\in\mathbb{N}} \subset \mathbb{R}_{\geq 0}$, starting at $t_1^i = 0$, determined by*

$$t_{k+1}^i = \operatorname{argmax}\{t \in [t_k^i, \infty) \mid \quad (32)$$
$$4\mathsf{d}_{out}^i \|\hat{\mathbf{x}}^i(t) - \mathbf{x}^i(t)\|^2 \leq \sum_{j=1}^N \mathsf{a}_{ij}\|\hat{\mathbf{x}}^i(t) - \hat{\mathbf{x}}^j(t)\|^2 + (\epsilon^i)^2\}.$$

Let $\beta > 0$ be such that

$$\gamma' = \alpha^2(\phi+1)m_T + 4.5\alpha\beta\lambda_2\phi - 4\alpha^2 M_T^2 - 4\alpha^2(\phi+1)^2 > 0,$$

is satisfied for some $\phi > 0$ with $\phi + 1 > \frac{4M_T^2}{m_T}$, where M_T and m_T are given in Theorem 1. Then, for any $\alpha > 0$ and each $i \in \{1, \ldots, N\}$, the algorithm evolution starting from initial conditions $\mathbf{x}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^d$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}_d$ makes $\|\mathbf{x}^i(t) - \mathbf{x}^\star\| \leq \frac{\phi\alpha\beta\bar{\lambda}_F}{4\eta\underline{\lambda}_F}\|\boldsymbol{\epsilon}\|^2$ as $t \to \infty$. (Here, $\eta = \min\{\frac{7}{16}, \frac{1}{9}\gamma'\}$, and $\underline{\lambda}_F$ and $\bar{\lambda}_F$ are the minimum and maximum eigenvalues of $\mathbf{F}$ given in (13)*). Furthermore, given initial conditions, the inter-execution times of agent $i \in \{1, \ldots, N\}$ are lower bounded by*

$$\tau^i = \frac{1}{\alpha M^i}\ln\Big(1 + \frac{\alpha M^i \epsilon^i}{2\sqrt{\mathsf{d}_{out}^i}(\alpha M^i + 2\beta\mathsf{d}_{out}^i + 1)\theta}\Big), \quad (33)$$

where $\theta = \frac{\bar{\lambda}_F}{\underline{\lambda}_F}\sqrt{\|\mathbf{x}(0) - \bar{\mathbf{x}}\|^2 + \|\mathbf{v}(0) - \bar{\mathbf{v}}\|^2 + \frac{\phi\alpha\beta\bar{\lambda}_F}{4\eta\underline{\lambda}_F}\|\boldsymbol{\epsilon}\|^2}$.

PROOF. Given an initial condition, let $[0, T)$ be the maximal interval on which there is no accumulation point in the set of event times $\{t_k\}_{k\in\mathbb{N}} = \cup_{i=1}^N \cup_{k\in\mathbb{N}} t_k^i$. Note that $T > 0$, since the number of agents is finite and, for each $i \in \{1, \ldots, N\}$, $\epsilon^i > 0$ and $\tilde{\mathbf{x}}^i(0) = \hat{\mathbf{x}}^i(0) - \mathbf{x}^i(0) = \mathbf{0}$. The dynamics (20), under the event-triggered communication scheme (32), has a unique solution in the time interval $[0, T)$. Next, we use Lyapunov analysis to show that the trajectory stays bounded in the time interval $[0, T)$. Consider the Lyapunov function V given in (12). The Lie derivative of V along (23b)-(23d) is given by

$$\begin{aligned}\dot{V} = &-\frac{1}{9}\alpha^2(\phi+1)\mathbf{y}^\top\mathbf{h} - \frac{7}{16}\mathbf{w}_{2:N}^\top\mathbf{w}_{2:N} \\ &- \|\frac{3}{4}\mathbf{w}_{2:N} + \frac{2\alpha}{3}(\mathbf{R}^\top \otimes \mathbf{I}_d)\mathbf{h} + \frac{2\alpha}{3}(\phi+1)\mathbf{z}_{2:N})\|^2 \\ &+ \frac{4}{9}\alpha^2\|(\mathbf{R}^\top \otimes \mathbf{I}_d)\mathbf{h}\|^2 + \frac{4}{9}\alpha^2(1+\phi)^2\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N} \\ &- \phi\alpha\beta\mathbf{z}_{2:N}^\top(\mathbf{R}^\top\mathbf{L}\mathbf{R}\otimes\mathbf{I}_d)(\mathbf{z}_{2:N} + \tilde{\mathbf{z}}_{2:N}),\end{aligned}$$

where $\tilde{\mathbf{z}}_{2:N} = \hat{\mathbf{z}}_{2:N} - \mathbf{z}_{2:N}$. Using the assumptions on the local cost functions and following steps similar to those taken in the proof of Theorem 1 to lower bound $\mathbf{y}^\top \mathbf{h}$ and upper bound $\|(\mathsf{R}^\top \otimes \mathsf{I}_d)\mathbf{h}\|$, one can show that

$$\begin{aligned}
\dot{V} \leq & -\frac{\alpha^2(\phi+1)m_{\mathrm{T}}}{9}\mathbf{z}^\top\mathbf{z} - \frac{7}{16}\mathbf{w}_{2:N}^\top\mathbf{w}_{2:N} \\
& - \|\frac{3}{4}\mathbf{w}_{2:N} + \frac{2\alpha}{3}(\mathsf{R}^\top \otimes \mathsf{I}_d)\mathbf{h} + \frac{2\alpha(\phi+1)}{3}\mathbf{z}_{2:N})\|^2 \\
& + \frac{4\alpha^2}{9}M_{\mathrm{T}}^2\mathbf{z}^\top\mathbf{z} + \frac{4\alpha^2(1+\phi)^2}{9}\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N} - \frac{\phi\alpha\beta\lambda_2}{2}\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N} \\
& - \frac{\phi\alpha\beta}{2}\Big(\mathbf{z}_{2:N}^\top(\mathsf{R}^\top\mathsf{L}\mathsf{R}\otimes\mathsf{I}_d)\mathbf{z}_{2:N} + 2\mathbf{z}_{2:N}^\top(\mathsf{R}^\top\mathsf{L}\mathsf{R}\otimes\mathsf{I}_d)\tilde{\mathbf{z}}_{2:N}\Big).
\end{aligned}$$

The last term in the third line of the inequality above comes from lower bounding half of $\mathbf{z}_{2:N}^\top(\mathsf{R}^\top\mathsf{L}\mathsf{R}\otimes\mathsf{I}_d)\mathbf{z}_{2:N}$ using $\lambda_2\mathsf{I} \leq \mathsf{R}^\top\mathsf{L}\mathsf{R}\otimes\mathsf{I}_d$. Define

$$s = -\mathbf{z}_{2:N}^\top(\mathsf{R}^\top\mathsf{L}\mathsf{R}\otimes\mathsf{I}_d)\mathbf{z}_{2:N} - 2\mathbf{z}_{2:N}^\top(\mathsf{R}^\top\mathsf{L}\mathsf{R}\otimes\mathsf{I}_d)\tilde{\mathbf{z}}_{2:N},$$

corresponding to the last term in the Lie derivative above. Our next objective is to show that $s \leq \frac{1}{2}\|\boldsymbol{\epsilon}\|^2$ for all $t \in [0,T)$. Using $\mathsf{R}\mathsf{R}^\top = \mathbf{\Pi}_N$, $\mathsf{L}\mathbf{\Pi}_N = \mathbf{\Pi}_N\mathsf{L} = \mathsf{L}$, $\mathbf{z}_{2:N} = (\mathsf{R}^\top \otimes \mathsf{I}_d)\mathbf{y} = (\mathsf{R}^\top \otimes \mathsf{I}_d)\mathbf{x}$, and $\tilde{\mathbf{z}}_{2:N} = (\mathsf{R}^\top \otimes \mathsf{I}_d)\tilde{\mathbf{y}} = (\mathsf{R}^\top \otimes \mathsf{I}_d)\tilde{\mathbf{x}}$, we obtain $s = -\mathbf{x}^\top\mathbf{L}\mathbf{x} - 2\mathbf{x}^\top\mathbf{L}\tilde{\mathbf{x}}$. Since $\mathbf{x} = \hat{\mathbf{x}} - \tilde{\mathbf{x}}$, then

$$\begin{aligned}
s &= -(\hat{\mathbf{x}} - \tilde{\mathbf{x}})^\top\mathbf{L}(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) - 2(\hat{\mathbf{x}} - \tilde{\mathbf{x}})^\top\mathbf{L}\tilde{\mathbf{x}} \\
&= -\hat{\mathbf{x}}^\top\mathbf{L}\hat{\mathbf{x}} + \tilde{\mathbf{x}}^\top\mathbf{L}\tilde{\mathbf{x}}.
\end{aligned}$$

Using the fact that $\mathsf{D}_{\mathrm{out}} + \mathsf{A} \geq 0$, we have

$$\begin{aligned}
\tilde{\mathbf{x}}^\top\mathbf{L}\tilde{\mathbf{x}} &= \tilde{\mathbf{x}}^\top(\mathsf{D}_{\mathrm{out}} \otimes \mathsf{I}_d)\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^\top(\mathsf{A} \otimes \mathsf{I}_d)\tilde{\mathbf{x}} \\
&\leq 2\tilde{\mathbf{x}}^\top(\mathsf{D}_{\mathrm{out}} \otimes \mathsf{I}_d)\tilde{\mathbf{x}} = 2\sum_{i=1}^{N}\mathsf{d}_{\mathrm{out}}^i\|\tilde{\mathbf{x}}^i\|^2.
\end{aligned}$$

Therefore, we can write

$$\begin{aligned}
s &= \frac{1}{2}\sum_{i=1}^{N}\Big(4\mathsf{d}_{\mathrm{out}}^i\|\tilde{\mathbf{x}}^i\|^2 - \sum_{j=1}^{N}\mathsf{a}_{ij}\|\hat{\mathbf{x}}^i - \hat{\mathbf{x}}^j\|^2\Big) \\
&= \frac{1}{2}\sum_{i=1}^{N}\big(4\mathsf{d}_{\mathrm{out}}^i\|\hat{\mathbf{x}}^i - \mathbf{x}^i\|^2 - \sum_{j=1}^{N}\mathsf{a}_{ij}\|\hat{\mathbf{x}}^i - \hat{\mathbf{x}}^j\|^2\big),
\end{aligned}$$

which, together with (32), implies that $s \leq \frac{1}{2}\|\boldsymbol{\epsilon}\|^2$ for $t \in [0,T)$. Then,

$$\begin{aligned}
\dot{V} \leq & -\frac{\alpha^2(\phi+1)m_{\mathrm{T}}}{9}\mathbf{z}^\top\mathbf{z} - \frac{7}{16}\mathbf{w}_{2:N}^\top\mathbf{w}_{2:N} - \frac{\phi\alpha\beta\lambda_2}{2}\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N} \\
& - \|\frac{3}{4}\mathbf{w}_{2:N} + \frac{2\alpha}{3}(\mathsf{R}^\top \otimes \mathsf{I}_d)\mathbf{h} + \frac{2\alpha(\phi+1)}{3}\mathbf{z}_{2:N})\|^2 \\
& + \frac{4\alpha^2}{9}M_{\mathrm{T}}^2\mathbf{z}^\top\mathbf{z} + \frac{4\alpha^2(1+\phi)^2}{9}\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N}^2 + \frac{\phi\alpha\beta}{4}\|\boldsymbol{\epsilon}\|^2 \\
\leq & -\eta\|\mathbf{p}\|^2 + \frac{\phi\alpha\beta}{4}\|\boldsymbol{\epsilon}\|^2, \quad t \in [0,T),
\end{aligned}$$

where η is given in the statement and $\mathbf{p} = (\mathbf{z}, \mathbf{w}_{2:N})$. Recall from the proof of Theorem 1 that $\underline{\lambda}_F\|\mathbf{p}\|^2 \leq V(\mathbf{p}) \leq \bar{\lambda}_F\|\mathbf{p}\|^2$. Then, using the Comparison Lemma (cf. [Khalil, 2002, Lemma 3.4]), we deduce that

$$\begin{aligned}
\|\mathbf{p}(t)\| &\leq \frac{1}{\underline{\lambda}_F}\|V(0)\|\,\mathrm{e}^{-\frac{\eta}{\bar{\lambda}_F}t} + \frac{\phi\alpha\beta\bar{\lambda}_F\|\boldsymbol{\epsilon}\|^2}{4\eta\underline{\lambda}_F}(1 - \mathrm{e}^{-\frac{\eta}{\bar{\lambda}_F}t}) \\
&\leq \frac{\bar{\lambda}_F}{\underline{\lambda}_F}\|\mathbf{p}(0)\|\mathrm{e}^{-\frac{\eta}{\bar{\lambda}_F}t} + \frac{\phi\alpha\beta\bar{\lambda}_F\|\boldsymbol{\epsilon}\|^2}{4\eta\underline{\lambda}_F}(1 - \mathrm{e}^{-\frac{\eta}{\bar{\lambda}_F}t}),
\end{aligned} \tag{34}$$

for $t \in [0,T)$. Notice that regardless of value of T,

$$\|\mathbf{p}(t)\| \leq \frac{\bar{\lambda}_F}{\underline{\lambda}_F}\|\mathbf{p}(0)\| + \frac{\phi\alpha\beta\bar{\lambda}_F}{4\eta\underline{\lambda}_F}\|\boldsymbol{\epsilon}\|^2, \tag{35}$$

for $t \in [0,T)$. Notice that the righthand side is equal to the constant θ given in the statement. This can be seen by noting that, from (9), we have $\|\mathbf{z}\| = \|\mathbf{x} - \bar{\mathbf{x}}\|$ and $\|\mathbf{w}\| = \|\mathbf{w}_{2:N}\| = \|\mathbf{v} - \bar{\mathbf{v}}\|$ (recall that $\sum_{i=1}^{N}\mathbf{v}^i(0) = \mathbf{0}$ results in $\mathbf{w}_1 = \mathbf{0}$ for all $t \geq 0$).

Our final objective is to show that $T = \infty$. To achieve this, we start by establishing a lower bound on the inter-execution times of any agent. To do this, let us determine a lower bound on the amount of time it takes for agent $i \in \{1,\ldots,N\}$ to have $\|\hat{\mathbf{x}}^i - \mathbf{x}^i\|$ evolve from 0 to $\epsilon^i/(2\sqrt{\mathsf{d}_{\mathrm{out}}^i})$. Using (7) and (20b), we have

$$\begin{aligned}
&\frac{d}{dt}\|\hat{\mathbf{x}}^i - \mathbf{x}^i\| = -\frac{(\hat{\mathbf{x}}^i - \mathbf{x}^i)^\top\dot{\mathbf{x}}^i}{\|\hat{\mathbf{x}}^i - \mathbf{x}^i\|} \leq \|\dot{\mathbf{x}}^i\| \\
&= \| - \alpha(\nabla f^i(\mathbf{x}^i) - \nabla f^i(\mathbf{x}^\star)) - \beta\sum_{j=1}^{N}\mathsf{a}_{ij}(\hat{\mathbf{x}}^i - \hat{\mathbf{x}}^j) \\
&\quad -(\mathbf{v}^i + \alpha\nabla f^i(\mathbf{x}^\star)\| \\
&\leq \alpha M^i\|\mathbf{x}^i - \mathbf{x}^\star\| + \beta\sum_{j=1}^{N}\mathsf{a}_{ij}\|\hat{\mathbf{x}}^i - \hat{\mathbf{x}}^j\| + \|\mathbf{v}^i - \bar{\mathbf{v}}^i\|,
\end{aligned}$$

which leads to

$$\frac{d}{dt}\|\hat{\mathbf{x}}^i - \mathbf{x}^i\| \leq \alpha M^i \|\hat{\mathbf{x}}^i - \mathbf{x}^i\| + \alpha M^i \|\hat{\mathbf{x}}^i - \mathbf{x}^\star\| + \beta \sum_{j=1}^{N} \mathsf{a}_{ij}(\|\hat{\mathbf{x}}^i - \mathbf{x}^\star\| + \|\hat{\mathbf{x}}^j - \mathbf{x}^\star\|) + \|\mathbf{v}^i - \bar{\mathbf{v}}^i\|.$$

From (35), we have $\|\mathbf{x}(t) - \bar{\mathbf{x}}\| \leq \theta$ and $\|\mathbf{v}(t) - \bar{\mathbf{v}}\| \leq \theta$. This implies $\|\mathbf{v}^i(t) - \bar{\mathbf{v}}^i\| \leq \theta$, $\|\hat{\mathbf{x}}^i - \mathbf{x}^\star\| \leq \theta$, and $\sum_{j=1}^{N} \mathsf{a}_{ij}(\|\hat{\mathbf{x}}^i - \mathbf{x}^\star\| + \|\hat{\mathbf{x}}^j - \mathbf{x}^\star\|) \leq 2\mathsf{d}^i_{\text{out}}\theta$, for all $i \in \{1,\dots,N\}$. Therefore, from the inequality above,

$$\frac{d}{dt}\|\hat{\mathbf{x}}^i - \mathbf{x}^i\| \leq \alpha M^i \|\hat{\mathbf{x}}^i - \mathbf{x}^i\| + c^i,$$

where $c^i = (\alpha M^i + 2\beta \mathsf{d}^i_{\text{out}} + 1)\theta$. Using the Comparison Lemma (cf. [Khalil, 2002, Lemma 3.4]) and the fact that $\|\hat{\mathbf{x}}^i - \mathbf{x}^i(t^i_k)\| = 0$, we deduce

$$\|\hat{\mathbf{x}}^i - \mathbf{x}^i(t)\| \leq \frac{c^i}{\alpha M^i}(\mathrm{e}^{\alpha M^i (t - t^i_k)} - 1), \quad t \geq t^i_k.$$

Then, the time it takes $\|\hat{\mathbf{x}}^i - \mathbf{x}^i\|$ to reach $\epsilon^i/(2\sqrt{\mathsf{d}^i_{\text{out}}})$ is lower bounded by $\tau^i > 0$ given by (33). Next, to show $T = \infty$, we proceed by contradiction. Our argument is similar to [Donkers and Heemels, 2012, proof of Theorem IV]. Suppose that $T < \infty$. Then, the sequence of events $\{t_k\}_{k\in\mathbb{N}}$ has an accumulation point at T. Because we have a finite number of agents, this means that there must be an agent $i \in \{1,\dots,N\}$ for which $\{t^i_k\}_{k\in\mathbb{N}}$ has an accumulation point at T, implying that agent i transmits infinitely often in the time interval $[T - \Delta, T)$ for any $\Delta \in (0, T]$. However, this is in contradiction with the fact that inter-event times are lower bounded by $\tau^i > 0$ on $[0, T)$. Having established $T = \infty$, note that this fact implies that under the event-triggered communication law (32), the algorithm (20) does not exhibit Zeno behavior. Furthermore, from (34), we deduce that, for each $i \in \{1,\dots,N\}$, one has $\|\mathbf{x}^i(t) - \mathbf{x}^\star\| \leq \|\mathbf{p}(t)\| \leq \frac{\phi\alpha\beta\bar{\lambda}_F}{4\eta\underline{\lambda}_F}\|\boldsymbol{\epsilon}\|^2$ as $t \to \infty$. □

The lower bound on the inter-event times allows a designer to compute bounds on the maximum energy spent by each agent (and hence the network) on communication during any given time interval. The combination of the facts that the total number of agents is finite and each agent's inter-event times are lower bounded implies that the total number of events in any finite time interval is finite. In general, an explicit expression lower bounding the network inter-event times is not available. It is also worth noticing that the farther away the agents starts from the final convergence point (larger θ in (33)), the smaller the guaranteed lower bound between inter-event times becomes. Notice also that the value of τ^i in (33) depends on the graph topology, the parameters of the local cost function, the algorithm design parameters α and β, and the variable ϵ^i. One can use this dependency to maximize the value of τ^i in a similar fashion as discussed in Remark 10. Finally, we should point out that for quadratic local cost functions of the form $f^i(\mathsf{x}) = \frac{1}{2}(\mathsf{x}^\top \mathsf{x} + \mathsf{x}^\top \mathsf{a}^i + \mathsf{b}^i)$, $i \in \{1,\dots,N\}$ the claim of Theorem 12 holds for any $\alpha, \beta > 0$. The technical details are omitted for brevity.

6 Simulations

In this section, we illustrate the performance of the coordination algorithm (1) and its implementation with discrete-time communication (20). In all simulations, we consider a network of 10 agents, with strongly convex local cost functions defined on $\mathbb{R}$ given by

$$\begin{aligned}
&f^1(x) = 0.5\,\mathrm{e}^{-0.5x} + 0.4\,\mathrm{e}^{0.3x}, && f^2(x) = (x-4)^2,\\
&f^3(x) = 0.5x^2\ln(1+x^2) + x^2, && f^4(x) = x^2 + \mathrm{e}^{0.1x},\\
&f^5(x) = \ln(\mathrm{e}^{-0.1x} + \mathrm{e}^{0.3x}) + 0.1x^2, && f^6(x) = x^2/\ln(2+x^2),\\
&f^7(x) = 0.2\,\mathrm{e}^{-0.2x} + 0.4\,\mathrm{e}^{0.4x}, && f^8(x) = x^4 + 2x^2 + 2,\\
&f^9(x) = x^2/\sqrt{x^2+1} + 0.1x^2, && f^{10}(x) = (x+2)^2.
\end{aligned}$$

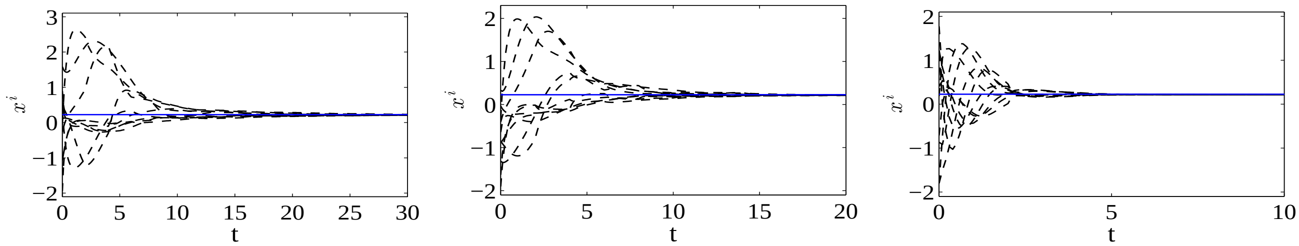
The gradient of the local cost function of agents 1, 4, 7, 8 are locally Lipschitz, while the rest are globally Lipschitz.

Figure 1 shows executions of the algorithm (1) for different values of β when the network topology alternates every 2 seconds among three strongly connected, weight-balanced digraphs (with unitary edge weights, Figure 2 shows one of these graphs). Convergence is achieved as guaranteed by Proposition 4 (see also Remark 3). The plot also shows that larger values of β result in faster convergence, cf. Remark 2. In all the simulations we ran, convergence is achieved for any $\alpha, \beta > 0$.

Figure 3(a)-(b) show executions of the algorithm (20) with periodic communication over the network depicted in Fig 2 for different β and Δ. Even though Theorem 9 is established for undirected graphs, our simulations show convergent behavior for strongly connected and weight-balanced digraphs. In case (a), it takes 33 communication events for all agents to be within 0.02 of the optimizer $\mathsf{x}^\star$. This value is 31 for case (b). This suggests a trade-off, where larger values of Δ (corresponding to smaller values of β, see Remark 10) result on savings on the energy consumed by agents for communication at the cost of slower convergence.

Figure 4(a)-(b) compares the evolution of the agents for algorithm (20) with periodic communication and for an Euler discretization of (1) over the network in Figure 2. In these simulations, we fixed Δ and varied β until the algorithm becomes close to the divergence. The results show that the algorithm (20) can use a larger β. This reveals that, for the same amount of communication effort, (20) achieves faster convergence.

Figure 5 shows the time history of the states x^i and the communication execution times of each agent



(a) $\alpha = 1,\ \beta = 0.5$ (b) $\alpha = 1,\ \beta = 1$ (c) $\alpha = 1,\ \beta = 5$

Fig. 1. Executions of (1) over a time-varying digraph that remains weight-balanced and strongly connected: the dashed lines (resp. solid blue line) show the time history of x^i's (resp. $\mathsf{x}^\star$).

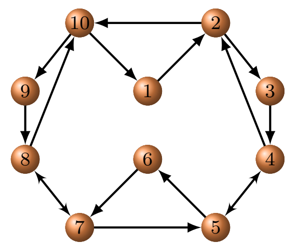
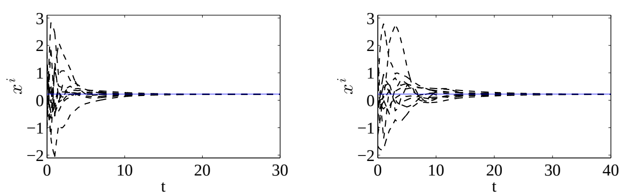


Fig. 2. A strongly connected and weight-balanced digraph with 10 agents (edge weights are 1).



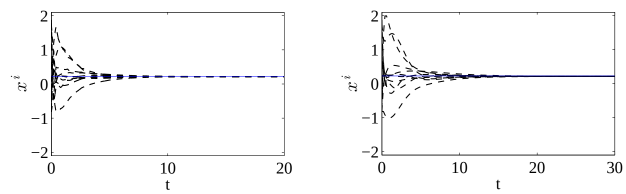
(a) Algorithm (20): $\alpha=1$, $\beta=1$, $\Delta=0.5$ sec. (b) Algorithm (20): $\alpha=1$, $\beta=0.5$, $\Delta=1$ sec.

Fig. 3. Performance evaluation of the algorithm (20) when the communication is periodic: the dashed lines (resp. solid blue line) show the time history of x^i's (resp. $\mathsf{x}^\star$).

$i \in \{1, \ldots, 10\}$ of the algorithm (20) when the distributed event-triggered communication law (32) is employed. The results illustrate the behavior guaranteed by Theorem 12: the communications times are asynchronous, the operation is Zeno-free, and the states converge to an $\|\boldsymbol{\epsilon}\|^2$-neighborhood of $\mathsf{x}^\star$.

7 Conclusions

We have presented a novel class of distributed continuous-time coordination algorithms that solve network optimization problems where the objective function is strictly convex and equal to a sum of local agent cost functions. For strongly connected and weight-balanced agent interactions, we have shown that our algorithms converge exponentially to the solution of the optimization problem when the local cost functions are strongly convex and their gradients are globally Lipschitz. This property is preserved in dynamic networks as long as the topology stays strongly connected and weight-balanced.



(a) Algorithm (20): $\alpha=1$, $\beta=2, \Delta=0.2$ sec. (b) Euler discretization of (1) with $\alpha=1, \beta=1, \Delta=0.2$ sec.

Fig. 4. Performance evaluation of the algorithm (20) when the communication is periodic vs. the Euler-discretized implementation of the algorithm (1): the dashed lines (resp. solid blue line) show the time history of x^i's (resp. $\mathsf{x}^\star$).

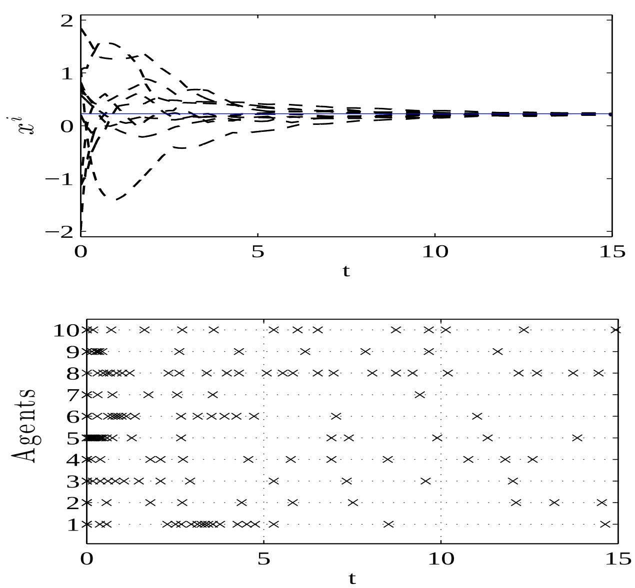


Fig. 5. Performance evaluation of the algorithm (20) with parameters $\alpha = \beta = 1$ when the distributed event-triggered communication law (32) with $\epsilon^i = 0.002$, $i \in \{1, \ldots, N\}$ is employed: in the top plot the dashed lines (resp. solid blue line) show the time history of x^i's (resp. $\mathsf{x}^\star$); in the bottom plot $\times$ shows the time an event is triggered by an agent.

For connected and undirected agent interactions, we have shown that exponential convergence still holds under the relaxed conditions of strongly convex local cost functions with locally Lipschitz gradients. In this case, asymptotic convergence also holds when the local cost functions are just convex. We have also explored the implementation of our algorithms with discrete-time communication. Specifically, we have established asymptotic convergence under periodic, centralized synchronous, and distributed asynchronous event-triggered communication schemes, paying special attention to establishing the Zeno-free nature of the algorithm executions. Future work will focus on strengthening the results presented in this paper, the exploration of abstractions about other agents' behaviors for self-triggered implementations, and the use of triggered control methods in other distributed optimization and coordination problems, including constrained, time-varying, and online scenarios, and networked games.

Acknowledgements

The second author is thankful to Cameron Nowzari for discussions on event-triggered control and distributed optimization.

References

S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2010.

F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. American Mathematical Society. Princeton University Press, 2009. ISBN 978-0-691-14195-4. Available at http://www.coordinationbook.info.

D.V. Dimarogonas, E. Frazzoli, and K.H. Johansson. Distributed event-triggered control for multi-agent systems. *IEEE Transactions on Automatic Control*, 57(5):1291–1297, 2012.

M.C.F. Donkers and W.P.M.H. Heemels. Output-based event-triggered control with guaranteed $\mathcal{L}_\infty$-gain and improved and decentralized event-triggering. *IEEE Transactions on Automatic Control*, 57(6):1362–1376, 2012.

J. Duchi, A. Agarwal, and M. Wainwright. Dual averaging for distributed optimization: Convergence analysis and network scaling. *IEEE Transactions on Automatic Control*, 57(3):592–606, 2012.

E. Garcia, Y. Caob, H. Yuc, P. Antsaklis, and D. Casbeerb. Decentralised event-triggered cooperative control with limited communication. *International Journal of Control*, 86(9):1479–1488, 2013.

B. Gharesifard and J. Cortés. Distributed continuous-time convex optimization on weight-balanced digraphs. *IEEE Transactions on Automatic Control*, 59(3), 2014. To appear.

W.P.M.H. Heemels, K.H Johansson, and P. Tabuada. An introduction to event-triggered and self-triggered controlg. In *IEEE Int. Conf. on Decision and Control*, pages 3270–3285, Maui, HI, 2012.

B. Johansson, M. Rabi, and M. Johansson. A randomized incremental subgradient method for distributed optimization in networked systems. *SIAM Journal on Optimization*, 20(3):1157–1170, 2009.

H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 3 edition, 2002. ISBN 0130673897.

J. Lu and C. Y. Tang. Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case. *IEEE Transactions on Automatic Control*, 57(9):2348–2354, 2012.

R. Madan and S. Lall. Distributed algorithms for maximum lifetime routing in wireless sensor networks. *IEEE Transactions on Wireless Communications*, 5(8):2185–2193, 2006.

M. Mazo Jr. and P. Tabuada. Decentralized event-triggered control over wireless sensor/actuator networks. *IEEE Transactions on Automatic Control*, 56(10):2456–2461, 2011.

A. Nedić and A. Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61, 2009.

C. Nowzari and J. Cortés. Zeno-free, distributed event-triggered communication and control for multi-agent average consensus. In *American Control Conference*, Portland, OR, 2014. Submitted.

V.M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. Pappas. Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks. Available in arXiv:1303.3984v1, 2013.

S. S. Ram, A. Nedić, and V.V. Veeravalli. Distributed and recursive parameter estimation in parametrized linear state-space models. *IEEE Transactions on Automatic Control*, 55(2):488–492, 2010.

J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9):803–812, 1986.

P. Wan and M.D. Lemmon. Event-triggered distributed optimization in sensor networks. In *Symposium on Information Processing of Sensor Networks*, pages 49–60, San Francisco, CA, 2009.

J. Wang and N. Elia. Control approach to distributed optimization. In *Allerton Conf. on Communications, Control and Computing*, Illinois, USA, October 2010.

J. Wang and N. Elia. A control perspective for centralized and distributed convex optimization. In *IEEE Int. Conf. on Decision and Control*, Florida, USA, December 2011.

X. Wang and M. D. Lemmon. Event-triggering in distributed networked control systems. *IEEE Transactions on Automatic Control*, 56(3):586–601, 2011.

F. Zanella, D. Varagnolo, A. Cenedese, G. Pillonetto, and L. Schenato. Newton-Raphson consensus for distributed convex optimization. In *IEEE Int. Conf. on Decision and Control*, pages 5917–5922, Florida, USA, December 2011.

Z. Zhang, X. Ying, and M. Chow. Decentralizing the economic dispatch problem using a two-level incremental cost consensus algorithm in a smart grid environment. In *North American Power Symposium*, Boston, MA, 2011. Electronic Proceedings.

M. Zhu and S. Martínez. On distributed convex optimization under inequality and equality constraints. *IEEE Transactions on Automatic Control*, 57(1):151–164, 2012.