2020 59th IEEE Conference on Decision and Control (CDC)
Jeju Island, Republic of Korea, December 14-18, 2020

Distributed Continuous-Time Optimization with Time-Varying Objective Functions and Inequality Constraints

Shan Sun, Wei Ren

Abstract— **This paper is devoted to the distributed continuous-time optimization problem with time-varying objective functions and time-varying nonlinear inequality constraints. Different from most studied distributed optimization problems with time-invariant objective functions and constraints, the optimal solution in this paper is time varying and forms a trajectory. To minimize the global time-varying objective function subject to time-varying local constraint functions using only local information and local interaction, we present a distributed control algorithm that consists of a sliding-mode part and a Hessian-based optimization part. The asymptotical convergence of the proposed algorithm to the optimal solution is studied under suitable assumptions. The effectiveness of the proposed scheme is demonstrated through a simulation example.**

I. Introduction

Distributed optimization algorithms allow for decomposing certain optimization problems into smaller, more manageable sub-problems that can be solved in parallel. Therefore, they are widely used to solve large-scale optimization problems such as optimization of network flows [1], big-data analysis [2], design of sensor networks [3], multi-robot teams [4], and resource allocation [5]. There has been significant attention on distributed convex optimization problems, where the goal is to cooperatively seek the optimal solution that minimizes the sum of private convex objective functions available to each individual agent. In this context, discrete-time distributed optimization algorithms have been studied extensively (see e.g., [6], [7] and references therein).

There exists another body of literature on distributed continuous-time optimization algorithms (see e.g., [8]–[14]). The distributed continuous-time optimization algorithms have applications in coordinated control of multi-agent teams. For example, multiple physical robots modeled by continuous-time dynamics might need to track a team optimal trajectory. Note that most studies in the literature focus on stationary optimization problems in which both the objective functions and constraints do not explicitly depend on time. However, in many applications, the local performance objectives or engineering constraints may evolve in time, reflecting the fact that the optimal solution could be changing over time and create a trajectory (see e.g., [15]–[18]), which makes the design and analysis much more complex. Moreover in practical optimization problems, constraints are always inevitable. In this paper, we are interested in the distributed continuous-time algorithms for time-varying constrained optimization problems.

The literature on the distributed continuous-time algorithms for time-varying optimization problems focuses on first-order gradient methods [19], [20] and second-order optimization methods [22]. Specifically, [19] and [20] solve the distributed time-varying optimization problems with convex set constraints using, respectively, the projected gradient method and gradient-based penalty function method. However, they are limited to solve certain kinds of optimization problems (e.g., [19]: quadratic objective functions; [20]: linear programming) and there exist tracking errors to the optimal solutions. Although the second-order optimization methods work well in centralized time-varying optimization problems (see e.g., [17], [18], [23]), their use in distributed settings has been prohibited as they require global information of the network to compute the inverse of the global Hessian matrix. Ref. [22] solves the distributed time-varying optimization problems using second-order optimization methods. However, the consensus-based algorithm in [22] (Section III.B) is limited to the unconstrained problem with local objective functions that have identical Hessians. While the estimator-based algorithm in [22] (Section III.C) can deal with certain objective functions with nonidentical Hessians, it relies on the distributed average tracking techniques [24] and hence poses restrictive assumptions that the derivatives of the Hessians and the gradients of the local objective functions are bounded. In addition, because the estimator-based algorithm has to estimate the Hessian inverse of the global objective function, it necessitates the communication of certain virtual variables between neighbors with increased computation costs. While it is possible to convert the constrained optimization problem to an unconstrained one using penalty methods, the resulting penalized objective functions would not have identical Hessians due to the involvement of the nonuniform local constraint functions (even if the original objective functions would), and they might not satisfy the restrictive assumptions mentioned above. As a result, the algorithms in [22] cannot be applied to address the distributed time-varying constrained optimization problem (see Remark 2 for a more detailed comparison). For discrete-time distributed online optimization algorithms, the readers are referred to [21] and references therein.

This paper aims to develop a distributed algorithm to solve the continuous-time optimization problem with private time-varying objective functions and time-varying nonlinear inequality constraints. In this work, the time-varying optimization problem is deformed as a consensus subproblem

This work was supported by National Science Foundation under Grant ECCS-1920798. S. Sun and W. Ren are with the Department of Electrical and Computer Engineering, University of California at Riverside, Riverside, CA, 92521 USA (e-mail: ssun029@ucr.edu, ren@ece.ucr.edu).

Authorized licensed use limited to: Univ of Calif Riverside. Downloaded on July 22,2021 at 01:15:08 UTC from IEEE Xplore. Restrictions apply.

and a minimization subproblem on the team objective function. We develop a sliding-mode method with a Hessian-dependent gain for all the agents to achieve consensus on the states. Meanwhile, a Hessian-based (second-order) optimization method coupled with the log-barrier penalty functions is proposed to track the local time-varying optimal solution. To implement the algorithm, each agent just needs its own state and the relative states between itself and its neighbors. When the agents' states are their positions, the algorithm can be implemented based on purely local sensing (e.g., absolute and relative positions) without the need for communicating virtual variables. The asymptotical convergence to the optimal solution is established based on nonsmooth analysis, Lyapunov theory and convex optimization theory. Numerical simulation results are presented to illustrate the effectiveness of the theoretical results. To the best of our knowledge, this is the first work in the literature on distributed continuous-time time-varying constrained optimization problems that guarantees zero tracking errors.

II. PRELIMINARIES

A. Notation

Let $\mathbb{R}, \mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ denote the sets of real numbers, real vectors of dimension n, and real matrices of size $n \times m$, respectively. Let $\mathbb{R}_{>0}$ represent the set of positive real numbers. The cardinality of a set S is denoted by $|S|$. Let $\mathbf{1}_n$ (resp. $\mathbf{0}_n$) denote the vector of n ones (resp. n zeros), and I_n denote the $n \times n$ identity matrix. For a matrix $A \in \mathbb{R}^{m\times n}$, $[A]_{k\bullet} \in \mathbb{R}^{1\times n}$ is the k-th row of A, and A^T (resp. A^{-1}) is the transpose (resp. inverse) of A, $\lambda_{\min}(A)$ is the smallest eigenvalue of A. For a vector $x = [x_1, \cdots, x_n]^T \in \mathbb{R}^{n\times 1}$, $\mathrm{diag}(x) \in \mathbb{R}^{n\times n}$ represents the diagonal matrix with the elements in the main diagonal being the elements of x, $\|x\|_p$ represents the p-norm of the vector x, $B(x, \delta)$ represents the open ball of radius δ centered at x, and $\mathrm{sgn}(x) = [\mathrm{sgn}(x_1), \cdots, \mathrm{sgn}(x_n)]^T$, where $\mathrm{sgn}(x_i)$ denotes the signum function defined as $\mathrm{sgn}(x_i) = -1$ if $x_i < 0$, $\mathrm{sgn}(x_i) = 0$ if $x_i = 0$, and $\mathrm{sgn}(x_i) = 1$ if $x_i > 0$. Let $\nabla f(x, t)$ and $\nabla^2 f(x, t)$ denote, respectively, the gradient and Hessian of function $f(x, t)$ with respect to the vector x. Let $\otimes$ denote the Kronecker product and $\overline{\mathrm{co}}$ the convex closure.

B. Graph Theory

An undirected graph, is denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, where $\mathcal{V} = \{1, ..., n\}$ is the node set, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set, and $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ is the weighted adjacency matrix with entries $a_{ij},\ i, j \in \mathcal{V}$. For an undirected graph, an edge (j, i) implies that node i and node j are able to share data with each other, and $a_{ij} = 1$ if $(j, i) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. Here $a_{ij} = a_{ji}$. Let $\mathcal{N}_i = \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$ denote the set of neighbors of node i. A path is a sequence of nodes connected by edges. An undirected graph is connected if for every pair of nodes there is a path connecting them. The Laplacian matrix $\mathcal{L} = [l_{ij}] \in \mathbb{R}^{n\times n}$ associated with $\mathcal{A}$ is defined as $l_{ii} = \sum_{j=1, j\neq i}^{n} a_{ij}$ and $l_{ij} = -a_{ij}$, where $i \neq j$. The incidence matrix $\mathcal{D} = [d_{ij}] \in \mathbb{R}^{n\times|\mathcal{E}|}$ associated with $\mathcal{G}$ is defined as $d_{ik} = -1$ if the kth edge leaves node i, $d_{ik} = 1$ if it enters node i, and $d_{ik} = 0$ otherwise. For the incidence matrix of an undirected graph, the orientation of the edges is assigned arbitrarily. Note that for an undirected graph, $\mathcal{L}\mathbf{1}_n = \mathbf{0}_n$, $\mathcal{L}^T = \mathcal{L}$ and $\mathcal{L} = \mathcal{D}\mathcal{D}^T$.

C. Nonsmooth Analysis

In this subsection, we recall some important definitions of the nonsmooth systems that will be exploited in our main result.

Definition 1. *(Filippov Solution) [25] Consider the vector differential equation*

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

where $f : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ is measurable and locally essentially bounded. A vector function $x(\cdot)$ is called a Filippov solution of (1) on $[t_0, t_1]$, if $x(\cdot)$ is absolutely continuous on $[t_0, t_1]$ and for almost all $t \in [t_0, t_1]$, $\dot{x}(t) \in K[f](x, t)$, where $K[f](x, t) := \bigcap_{\delta>0} \bigcap_{\mu(N)=0} \overline{\mathrm{co}} f(B(x, \delta) - N, t)$ is the Filippov set-valued map of $f(x, t)$ and $\bigcap_{\mu(N)=0}$ denotes the intersection over all sets N of Lebesgue measure zero.

Definition 2. *(Clarke's Generalized Gradient) [25] Consider a locally Lipschitz continuous function $V(x) : \mathbb{R}^d \to \mathbb{R}$, the generalized gradient of the function V at x is given by $\partial V(x) := \overline{\mathrm{co}}\{\lim \nabla V(x) | x_i \to x, x_i \notin \Omega_V\}$, where Ω_V is the set of measure zero where the gradient of V is not defined.*

Definition 3. *(Chain Rule) [25] Let $x(\cdot)$ be a Filippov solution of $\dot{x} = f(x, t)$ and $V(x) : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz continuous function. Then for almost all t,*

$$\frac{d}{dt} V[x(t)] \in \dot{\tilde{V}},$$

where $\dot{\tilde{V}}$ is the set-valued Lie derivative defined as $\dot{\tilde{V}} := \bigcap_{\xi\in\partial V} \xi^T K[f]$.

III. MAIN RESULTS

Consider a network consisting of n agents. Each agent is regarded as a node in an undirected graph, and each agent can only interact with its local neighbors in the network. Suppose that each agent satisfies the following continuous-time dynamics

$$\dot{x}_i(t) = u_i(t), \tag{2}$$

where $x_i(t) \in \mathbb{R}^m$ is the state of agent i, and $u_i(t) \in \mathbb{R}^m$ is the control input of agent i. In this work, we study the distributed time-varying optimization problem with time-varying nonlinear inequality constraints. The goal is to design $u_i(t)$ using only local information and interaction, such that all the agents work together to find the optimal trajectory $y^*(t) \in \mathbb{R}^m$ which is defined as

$$\begin{aligned} y^*(t) = \mathrm{argmin} \quad & \sum_{i=1}^{n} f_i[y(t), t], \\ \text{s.t.} \quad & g_i[y(t), t] \leq \mathbf{0}_{q_i},\ i \in \mathcal{V}, \end{aligned} \tag{3}$$

Authorized licensed use limited to: Univ of Calif Riverside. Downloaded on July 22,2021 at 01:15:08 UTC from IEEE Xplore. Restrictions apply.

where $f_i[y(t),t]:\mathbb{R}^m\times\mathbb{R}_{>0}\to\mathbb{R}$ are the local objective functions, and $g_i[y(t),t]:\mathbb{R}^m\times\mathbb{R}_{>0}\to\mathbb{R}^{q_i}$ are the local inequality constraint functions. It is assumed that $f_i[y(t),t]$ and $g_i[y(t),t]$ are known only to agent i. We assume that the local objective functions $f_i[y(t),t]$ and inequality constraint functions $g_i[y(t),t]$ are twice continuously differentiable with respect to $y(t)$ and continuously differentiable with respect t.

If the underlying network is connected, the above problem (3) is equivalent to the problem that all the agents reach consensus while optimizing the team objective function $\sum_{i=1}^{n} f_i[x_i(t),t]$ under constraints, more formally,

$$\begin{aligned} &x^*(t)\in\mathbb{R}^{m*n}=\operatorname{argmin}\quad \sum_{i=1}^{n} f_i[x_i(t),t],\\ &\text{s.t.}\quad g_i[x_i(t),t]\le \mathbf{0}_{q_i},\quad x_i(t)=x_j(t),\quad \forall i,j\in\mathcal{V}, \end{aligned}\tag{4}$$

where $x(t)\in\mathbb{R}^{m*n}$ is the stack of all the agents' states. Here, the goal is that each state $x_i(t),\ \forall i\in\mathcal{V}$, converges to the optimal solution $y^*(t)$, i.e.,

$$\lim_{t\to\infty}[x_i(t)-y^*(t)]=\mathbf{0}_m.\tag{5}$$

This architecture of the distributed time-varying constrained optimization problem with networked agents finds broad applications in distributed cooperative control problems, including multi-robot navigation [15], [16] and resource allocation of power network [18]. For notational simplicity, we will remove the time index t from the variables $x_i(t)$ and $u_i(t)$ in most remaining parts of this paper and only keep it in some places when necessary.

Lemma 1. *[28] Let $f(r):R^m\to R$ be a continuously differentiable convex function with respect to r. The function $f(r)$ is minimized at r^* if and only if $\nabla f(r^*)=0$.*

We make the following assumptions which are all standard in the literature and are used in recent works like [22], [23].

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

Assumption 2. *All the objective functions $f_i(x_i,t)$ are uniformly strongly convex in x_i, for all $t\ge 0$.*

Assumption 3. *All the constraints $g_i(x_i,t)$ are uniformly convex in x_i, for all $t\ge 0$.*

Assumption 4. *For all $t\ge 0$ and for all $i\in\mathcal{V}$, there exists at least one y such that $g_i(y,t)<\mathbf{0}_{q_i}$. Therefore, the Slater's condition holds for all time.*

The uniform strong convexity of the objective function implies that the optimal trajectory $y^*(t)$ is unique for all $t\ge 0$.

A. Distributed Algorithm Design

In this subsection, we derive our distributed control algorithm for the time-varying constrained optimization problem in (4). We design the following controller for agent i:

$$\begin{aligned} u_i&=-\beta[\nabla^2\tilde{L}_i(x_i,t)]^{-1}\sum_{j\in\mathcal{N}_i}\operatorname{sgn}(x_i-x_j)+\phi_i(t),\\ \phi_i(t)&=-[\nabla^2\tilde{L}_i(x_i,t)]^{-1}\left[\nabla\tilde{L}_i(x_i,t)+\frac{\partial}{\partial t}\nabla\tilde{L}_i(x_i,t)\right], \end{aligned}\tag{6}$$

where $\beta\in\mathbb{R}_{>0}$ is a constant control gain, and $\tilde{L}_i(x_i,t)$ is a penalized objective function of agent i, defined as,

$$\tilde{L}_i(x_i,t)=f_i(x_i,t)-\frac{1}{\rho(t)}\sum_{j=1}^{q_i}\log[1-\rho(t)g_{ij}(x_i,t)],\tag{7}$$

where $g_{ij}(x_i,t):\mathbb{R}^m\times\mathbb{R}_{>0}\to\mathbb{R}$ denotes the $j-$th component of function $g_i(x_i,t)$, and $\rho(t)\in\mathbb{R}_{>0}$ is a time-varying barrier parameter satisfying

$$\rho(t)=a_1e^{a_2t},\ a_1,a_2\in\mathbb{R}_{>0}.\tag{8}$$

Remark 1. *In this work, the time-varying optimization problem* (4) *is deformed as a consensus subproblem and a minimization subproblem on the team objective function. We develop a distributed sliding-mode control law to address the consensus part. That is, the role of term $-\beta[\nabla^2\tilde{L}_i(x_i,t)]^{-1}\sum_{j\in\mathcal{N}_i}\operatorname{sgn}(x_i-x_j)$ in* (6) *is to drive all the agents to reach a consensus on the states ($\lim_{t\to\infty}\|x_i-\sum_{j=1}^{n}x_j\|_2=0$). Here, the Hessian-dependent gain $\beta[\nabla^2\tilde{L}_i(x_i,t)]^{-1}$ is introduced to guarantee the convergence of our algorithm under nonidentical $\nabla^2\tilde{L}_i(x_i,t)$. While the second term, $\phi_i(t)\in\mathbb{R}^m$, is an auxiliary variable playing a role in minimizing the penalized objective function $\tilde{L}_i(x_i,t)$ given by* (7)*. Note that we use the log-barrier penalty functions (see the second term in* (7)*) to incorporate the inequality constraints into the penalized objective function. As shown in* (6)*, we use the second-order/Hessian information of the penalized objective function to achieve the optimization goal.*

In addition, we have

$$\nabla\tilde{L}_i(x_i,t)=\nabla f_i(x_i,t)+\sum_{j=1}^{q_i}\frac{\nabla g_{ij}(x_i,t)}{1-\rho(t)g_{ij}(x_i,t)},\tag{9}$$

$$\begin{aligned} \frac{\partial}{\partial t}\nabla\tilde{L}_i(x_i,t)=&\frac{\partial}{\partial t}\nabla f_i(x_i,t)+\sum_{j=1}^{q_i}\frac{\partial\nabla g_{ij}(x_i,t)/\partial t}{1-\rho(t)g_{ij}(x_i,t)}\\ &+\sum_{j=1}^{q_i}\frac{\dot{\rho}(t)g_{ij}(x_i,t)\nabla g_{ij}(x_i,t)}{[1-\rho(t)g_{ij}(x_i,t)]^2},\\ &+\sum_{j=1}^{q_i}\frac{\rho(t)\nabla g_{ij}(x_i,t)\partial g_{ij}(x_i,t)/\partial t}{[1-\rho(t)g_{ij}(x_i,t)]^2}, \end{aligned}\tag{10}$$

$$\begin{aligned} \nabla^2\tilde{L}_i(x_i,t)=&\nabla^2 f_i(x_i,t)+\sum_{j=1}^{q_i}\frac{\nabla^2 g_{ij}(x_i,t)}{1-\rho(t)g_{ij}(x_i,t)}\\ &+\sum_{j=1}^{q_i}\frac{\rho(t)\nabla g_{ij}(x_i,t)\nabla g_{ij}(x_i,t)^T}{[1-\rho(t)g_{ij}(x_i,t)]^2}, \end{aligned}\tag{11}$$

Authorized licensed use limited to: Univ of Calif Riverside. Downloaded on July 22,2021 at 01:15:08 UTC from IEEE Xplore. Restrictions apply.

where $\frac{\partial}{\partial t}\nabla f_i(x_i,t)$, $\frac{\partial}{\partial t}\nabla g_{ij}(x_i,t)$ and $\frac{\partial}{\partial t}g_{ij}(x_i,t)$ are, respectively, the partial derivatives of $\nabla f_i(x_i,t)$, $\nabla g_{ij}(x_i,t)$ and $g_{ij}(x_i,t)$ with respect to t. To make the algorithm (6) work, the initial states $x_i(0)$ need satisfy

$$g_i[x_i(0),0] < \mathbf{0}_{q_i}, \qquad \forall i \in \mathcal{V}. \tag{12}$$

Also, for notational simplicity, we will remove the time index t from the auxiliary variable $\phi_i(t)$ in most remaining parts of this paper and only keep it in some places when necessary.

Remark 2. *In this work, we convert the considered constrained optimization problem into an unconstrained one using the log-barrier penalty functions. It is worth noting that the proposed algorithm* (6) *is not a simple extension of the existing distributed time-varying unconstrained optimization algorithms in [22]. Especially, to apply the consensus-based algorithm in [22] (Section III.B), it is required that the Hessians of all the local objective functions be identical. In contrast, in our context with the penalized objective functions, the Hessians of them are nonuniform due to the involvement of the nonuniform local constraint functions even if the original objective functions have identical Hessians. The estimator-based algorithm in [22] (Section III.C) can deal with certain objective functions with nonidentical Hessians. However, it not only necessitates the communication of certain virtual variables between neighbors with increased computation costs, but requires that the derivatives of the Hessians and the gradients of the objective functions be bounded. Unfortunately, due to the complexity of the penalized objective functions in the considered constrained problem, such a requirement is no longer guaranteed to hold and hence the result therein is not applicable to our problem. In this paper, we introduce a novel algorithm with a Hessian-dependent gain to account for the complexity caused by the penalized objective functions. The novel algorithm design in turn introduces new challenges in theoretical analysis, which will be addressed in the following.*

Remark 3. *In algorithm* (6)*, each agent just needs its own information and the relative states between itself and its neighbors. In some robotic applications, the agents' states are their spatial positions. As a result, the relative positions can be obtained by local sensing and the communication necessity might be eliminated.*

B. Convergence Analysis

In this subsection, the asymptotical convergence of system (2) under controller (6) to the optimal solution is established. To establish our results, we require the following assumptions.

Assumption 5. *If all the local states x_i are bounded, then there exists a constant $\bar{\alpha}$ such that $\sup_{t\in[0,\infty)} \|\frac{\partial}{\partial t}\nabla f_i(x_i,t)\|_2 \leq \bar{\alpha}$, for all $i \in \mathcal{V}$ and $t \geq 0$.*

Assumption 6. *If all the local states x_i are bounded, then there exist constants $\bar{\beta}$ and $\bar{\gamma}$ such that $\sup_{t\in[0,\infty)} \|\frac{\partial}{\partial t}\nabla g_{ij}(x_i,t)\|_2 \leq \bar{\beta}$ and $\sup_{t\in[0,\infty)} \|\frac{\partial}{\partial t}g_{ij}(x_i,t)\|_2 \leq \bar{\gamma}$, for all $i \in \mathcal{V}$, $j \in [1,\cdots,q_i]$ and $t \geq 0$.*

Remark 4. *In Assumption 5, we assume that all $\|\frac{\partial}{\partial t}\nabla f_i(x_i,t)\|_2$ are bounded under bounded x_i. The assumption holds for an important class of situations. For example, consider the normal quadratic objective functions $\|c_i x_i + h_i(x_i,t)\|_2^2$. As long as $\frac{\partial}{\partial t}h_i(x_i,t)$ (e.g. $\sin(t),t$) are bounded under bounded x_i, $\|\frac{\partial}{\partial t}\nabla f_i(x_i,t)\|_2$ will be bounded. In Assumption 6, we assume that all $\|\frac{\partial}{\partial t}\nabla g_{ij}(x_i,t)\|_2$ and $\|\frac{\partial}{\partial t}g_{ij}(x_i,t)\|_2$ are bounded under bounded x_i. The assumption holds for an important class of situations. The boundedness of $\|\frac{\partial}{\partial t}\nabla g_{ij}(x_i,t)\|_2$ and $\|\frac{\partial}{\partial t}g_{ij}(x_i,t)\|_2$ holds for most commonly used boundary constraint functions, e.g., $x_i \leq b(t)$ or $x_i^2 \leq b(t)$ under bounded $\dot{b}(t)$.*

Remark 5. *With the piecewise-differentiable signum function involved in algorithm* (6)*, the solution should be investigated in the sense of Filippov [26]. However, since the signum function is measurable and locally essentially bounded, the Filippov solutions of the proposed system dynamics always exist. Hence if the Lyapunov function candidates are continuously differentiable, the set-valued Lie derivative of them is a singleton at the discontinuous points and the proof still holds without employing the nonsmooth analysis.*

In this work, we convert the considered constrained optimization problem into an unconstrained optimization problem using the log-barrier penalty functions. It is important that the log-barrier penalty function involved in (7) is always well defined under our proposed algorithm. This is described in the next lemma.

Lemma 2. *Suppose that Assumption 4 and the initial condition* (12) *hold. For the system* (2) *under controller* (6)*, each $x_i(t)$ belongs to set $D_i = \{x_i \in \mathbb{R}^m \mid g_i(x_i,t) < \frac{1}{\rho(t)}\mathbf{1}_{q_i}\}$ for all $t \geq 0$. That is,* (7) *is always well defined.*

Due to space limitation, we omit the proofs to Lemmas 2 to 5 in this paper. The proofs can be found in [27]. In the following, in Lemma 3, we prove that the eventual states of the agents satisfy the optimal requirement shown in Lemma 1, i.e., $\lim_{t\to\infty} \sum_{i=1}^{n} \nabla \tilde{L}_i[x_i(t),t] = 0$. The goal of problem (4) is that all the agents' states reach consensus on the optimal trajectory, and thus in Lemma 4, we prove that consensus can be achieved eventually if all ϕ_i in the controller (6) are bounded, i.e., $\lim_{t\to\infty} \|x_i(t) - \sum_{j=1}^{n} x_j(t)\| = 0$ can be achieved under bounded ϕ_i. Then in Lemma 5, we prove that all ϕ_i associated with the system (2) under controller (6) are indeed bounded. Finally, in Theorem 1, we present that the original constrained optimization problem (4) can be achieved, i.e., $\lim_{t\to\infty} \|x_i(t) - y^*(t)\|_2 = 0$ for all $i \in \mathcal{V}$.

Lemma 3. *Suppose that Assumptions 1 to 4, the gain condition* (8) *and initial condition* (12) *hold. For the system* (2) *under controller* (6)*, the summation of all $\nabla \tilde{L}_i(x_i,t)$*

Authorized licensed use limited to: Univ of Calif Riverside. Downloaded on July 22,2021 at 01:15:08 UTC from IEEE Xplore. Restrictions apply.

satisfies $\lim_{t\to\infty} \sum_{i=1}^{n} \nabla \tilde{L}_i(x_i, t) = \mathbf{0}_m$.

Lemma 4. *Suppose that Assumptions 1 to 4, the gain condition* (8) *and initial condition* (12) *hold. For the system* (2) *under controller* (6)*, if there exists a constant* $\bar{\phi}$ *such that* $\sup_{t\in[0,\infty)} \|\phi_i(t)\|_2 \leq \bar{\phi}$, $\forall i \in \mathcal{V}$ *and* β *satisfies that,*

$$\beta > \frac{2\bar{\phi} m n^2 |\mathcal{E}|}{\min\{\lambda_{\min}[(\nabla^2 \tilde{L}_i)^{-1}]\}}, \tag{13}$$

all the states x_i *will achieve consensus eventually, i.e.,* $\lim_{t\to\infty} \|x_i(t) - x_j(t)\|_2 = 0$, $\forall i, j \in \mathcal{V}$.

Lemma 5. *Suppose that Assumptions 1 to 6, the gain condition* (8) *and initial condition* (12) *hold. For the system* (2) *under controller* (6)*, all* ϕ_i *remain bounded. That is, there exists a constant* $\bar{\phi}$ *such that* $\sup_{t\in[0,\infty)} \|\phi_i(t)\|_2 \leq \bar{\phi}$, $\forall i \in \mathcal{V}$.

Theorem 1. *Suppose that Assumptions 1 to 6, the initial condition* (12) *and gain conditions* (8) *and* (13) *hold. For the system* (2) *under controller* (6)*, all the states* x_i *will converge to the optimal solution* $y^*(t)$ *in* (3) *eventually.*

Proof: Define

$$\tilde{y}(t)^* \in \mathbb{R}^m = \text{argmin} \sum_{i=1}^{n} \tilde{L}_i[y(t), t], \tag{14}$$

where $\tilde{L}_i[y(t), t]$ is each agent's penalized objective function defined by (7). Note that Assumptions 1 to 6, the initial condition (12) and gain condition (8) hold. It follows from Lemma 5 that all ϕ_i associated with the system (2) under controller (6) are bounded for all $t \geq 0$, which in turn implies that $x_i(t) = x_j(t)$, $\forall i, j \in \mathcal{V}$ eventually according to Lemma 4. Moreover, based on Lemma 3 we know that $\lim_{t\to\infty} \sum_{i=1}^{n} \nabla \tilde{L}_i(x_i, t) = \mathbf{0}_m$. Using a similar analysis to that in Lemma 3, we have each $\tilde{L}_i(x_i, t)$ is continuously differentiable and strongly convex in x_i. Then it follows from Lemma 1 that all x_i will converge to the optimal solution $\tilde{y}^*(t)$ in (14), i.e., $\lim_{t\to\infty} x_i(t) = \tilde{y}^*(t)$, $\forall i \in \mathcal{V}$.

Define

$$\hat{y}^*(t) \in \mathbb{R}^m = \text{argmin} \sum_{i=1}^{n} f_i[y(t), t]$$
$$\text{s.t.} \quad g_{ij}[y(t), t] \leq \frac{1}{\rho(t)}, \ \forall i \in \mathcal{V}, j = 1, \cdots, q_i.$$

According to [28] (Sec. 11.2), we know that

$$\left| \sum_{i=1}^{n} f_i[\hat{y}^*(t), t] - \sum_{i=1}^{n} f_i[\tilde{y}^*(t), t] \right| \leq \sum_{j=1}^{n} \frac{q_j}{\rho(t)}.$$

Note that $y^*(t) \in \mathbb{R}^m$ is the optimal solution of problem (3). Under Assumption 4, the optimal solution $y^*(t)$ can be characterized using the Karush–Kuhn–Tucker (KKT) conditions for all $t \geq 0$. Then based on [28] (Sec. 5.9), we have

$$\left| \sum_{i=1}^{n} f_i[\hat{y}^*(t), t] - \sum_{i=1}^{n} f_i[y^*(t), t] \right| \leq \sum_{j=1}^{n} \sum_{k=1}^{q_j} \frac{\nu_{jk}^*(t)}{\rho(t)},$$

where $\nu_{jk}(t)$ are the Lagrangian multipliers corresponding to the inequality constraint defined in (3), and $\nu_{jk}^*(t)$ are the optimal Lagrangian multipliers. Hence, because $\lim_{t\to\infty} \rho(t) = \infty$, we have

$$\lim_{t\to\infty} \left| \sum_{i=1}^{n} f_i[y^*(t), t] - \sum_{i=1}^{n} f_i[\tilde{y}^*(t), t] \right| = 0.$$

Under Assumption 2, the optimal solution $y^*(t)$ is unique, which indicates that $\lim_{t\to\infty} x_i(t) = y^*(t)$, $\forall i \in \mathcal{V}$.

IV. NUMERICAL SIMULATION RESULTS

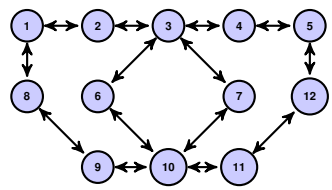

Fig. 1: An undirected graph.

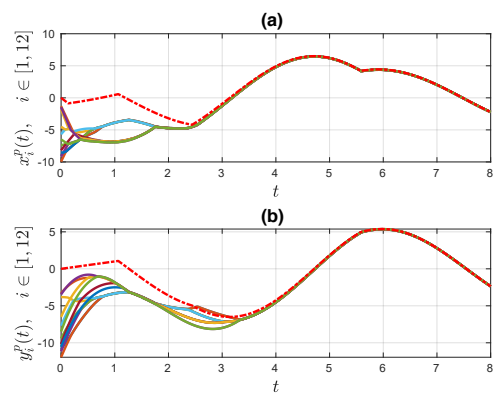

Fig. 2: State trajectories of all the agents corresponding to Theorem 1. The red dashed line is the optimal solution and the other solid lines are the trajectories of all agents' states.

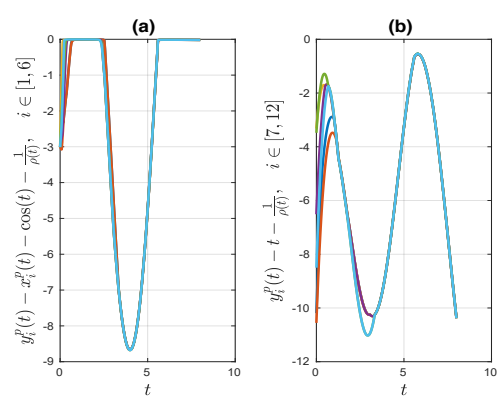

Fig. 3: (a) The constraint results of agents $1-6$.
(b) The constraint results of agents $7-12$.

In this section, we use a simulation case to illustrate Theorem 1. We consider a network with $n = 12$ and $m = 2$.

Authorized licensed use limited to: Univ of Calif Riverside. Downloaded on July 22,2021 at 01:15:08 UTC from IEEE Xplore. Restrictions apply.

The network topology is shown by the undirected graph in Figure 1. Let $x_i = [x_i^p, y_i^p]^T \in \mathbb{R}^2$ denote the states of each agent and consider the following optimization problem:

$$\begin{aligned} \min \quad & \sum_{i=1}^{12} \frac{1}{2}[x_i^p(t) + i\sin(t)]^2 + \frac{3}{2}[y_i^p(t) - i\cos(t)]^2, \\ \text{s.t.} \quad & y_j^p(t) - x_j^p(t) - \cos(t) \leq 0, \ \forall j \in [1, 2, \cdots, 6], \\ & y_k^p(t) - t \leq 0, \ \forall k \in [7, 8, \cdots, 12], \\ & x_i^p = x_j^p \text{ and } y_i^p = y_j^p, \ \forall i, j \in \mathcal{V}. \end{aligned} \tag{15}$$

The initial states $x_i^p(0), \ i \in \mathcal{V}$ are generated randomly from the range $[-10, 0]$, and $y_i^p(0) = x_i^p(0) - 2, \ i \in \mathcal{V}$. Therefore, the initial condition (12) is satisfied. We choose $\beta = 25$, $\rho(t) = 100\exp(0.1t)$. Therefore, the gain condition (8) is satisfied. The state trajectories of all the agents are shown in Figure 2. We can see that all the agents are able to track the optimal trajectory without tracking errors eventually which is consistent with Theorem 1. The constraint result is shown in Figure 3. In our simulation, agents $1-6$ are assigned to the constraint function $y_i^p(t) - x_i^p(t) - \cos(t) \leq 0, \ i \in [1, \cdots, 6]$, so $y_i^p(t) - x_i^p(t) - \cos(t) - 1/\rho(t), \ i \in [1, \cdots, 6]$ always remain negative. Agents $7-12$ are assigned to the constraint function $y_i^p(t) - t \leq 0, \ i \in [7, \cdots, 12]$, and thus $y_i^p(t) - t - 1/\rho(t), \ i \in [7, \cdots, 12]$ always remain negative.

V. CONCLUSIONS

In this paper, we have studied the distributed continuous-time constrained optimization problem with time-varying objective functions and time-varying constraints. The goal is that a set of networked agents cooperate to track the time-varying optimal solution that minimizes the summation of all the local time-varying objective functions subject to all the local time-varying constraints, where each agent can only receive information from its neighbors. We propose a distributed sliding-mode algorithm built on the Hessian-based optimization methodology. We have shown that asymptotical convergence is guaranteed under some reasonable assumptions. Numerical simulation result is given to illustrate the theoretical algorithm.

REFERENCES

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

[2] R. Zhang, and J. Kwok. "Asynchronous distributed ADMM for consensus optimization." in *Proceedings of International Conference on Machine Learning*, Beijing, China, 2014, pp. 1701-1709.

[3] M. Rabbat, and R. Nowak. "Distributed optimization in sensor networks." in *Proceedings of International Symposium on Information Processing in Sensor Networks*, Berkeley, USA, 2004, pp. 20-27.

[4] M. T. Shahab, and M. Elshafei. "Distributed optimization of multi-robot motion with time-energy criterion." *Path Planning for Autonomous Vehicles*, IntechOpen, Apr. 2019.

[5] H. Lakshmanan, and D. P. De Farias. "Decentralized resource allocation in dynamic networks of agents." *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 911-940, Sept. 2008.

[6] T. Yang, X. Yi, J. Wu, Y. Yuan, D. Wu, Z. Meng, Y. Hong, H. Wang, Z. Lin, and K. H. Johansson. "A survey of distributed optimization." *Annual Reviews in Control*, vol. 47, pp. 278–305, 2019.

[7] A. Nedic, and A. Ozdaglar. "Distributed subgradient methods for multi-agent optimization." *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48-61, Jan. 2009.

[8] 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.

[9] P. Lin, W. Ren, C. Yang, and W. Gui. "Distributed optimization with nonconvex velocity constraints, nonuniform position constraints, and nonuniform stepsizes." *IEEE Transactions on Automatic Control*, vol. 64, no. 6, pp. 2575-2582, June 2019.

[10] Z. Qiu, S. Liu, and L. Xie. "Distributed constrained optimal consensus of multi-agent systems." *Automatica*, vol. 68, pp. 209-215, June 2016.

[11] J. Lu, and C. 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.

[12] S. Liang, X. Zeng, and Y. Hong. "Distributed nonsmooth optimization with coupled inequality constraints via modified Lagrangian function." *IEEE Transactions on Automatic Control*, vol. 63, no. 6, pp. 1753-1759, June 2018.

[13] Y. Zhu, W. Yu, G. Wen, G. Chen, and W. Ren. "Continuous-time distributed subgradient algorithm for convex optimization with general constraints." *IEEE Transactions on Automatic Control*, vol. 64, no. 4, pp. 1694-1701, Apr. 2019.

[14] S. Yang, Q. Liu, and J. Wang. "A multi-agent system with a proportional-integral protocol for distributed constrained optimization." *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3461-3467, July 2017.

[15] S. G. Lee, Y. Diaz-Mercado, and M. Egerstedt. "Multirobot control using time-varying density functions." *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 489-493, Apr. 2015.

[16] D. Verscheure, B. Demeulenaere, J. Swevers, J. De Schutter, and M. Diehl. "Time-optimal path tracking for robots: A convex optimization approach." *IEEE Transactions on Automatic Control*, vol. 54, no. 10, pp. 2318-2327, Oct. 2009.

[17] W. Su."Traffic engineering and time-varying convex optimization." Ph.D dissertation, Department of Electrical Engineering, Pennsylvania State University, Philadelphia, 2009.

[18] Y. Tang. "Time-varying optimization and its application to power system operation." Ph.D dissertation, Department of Electrical Engineering, California Institute of Technology, Pasadena, 2019.

[19] C. Sun, M. Ye, and G. Hu. "Distributed time-varying quadratic optimization for multiple agents under undirected graphs." *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3687-3694, July 2017.

[20] H. Mehdi, E. Garone, and L. Schenato. "A distributed method for linear programming problems with box constraints and time-varying inequalities." *IEEE Control Systems Letters,* vol. 3, no. 2, pp. 404-409, Apr. 2019.

[21] X. Yi, X. Li, L. Xie, and K. H. Johansson. "Distributed online convex optimization with time-varying coupled inequality constraints." *IEEE Transactions on Signal Processing,* vol. 68, pp, 731-746, 2020.

[22] S. Rahili, and W. Ren. "Distributed continuous-time convex optimization with time-varying cost functions." *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1590-1605, Apr. 2017.

[23] M. Fazlyab, S. Paternain, V. M. Preciado and A. Ribeiro. "Prediction-correction interior-point method for time-varying convex optimization." *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 1973-1986, July 2018.

[24] F. Chen, Y. Cao, and W. Ren. "Distributed average tracking of multiple time-varying reference signals with bounded derivatives." *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 3169-3174, Dec. 2012.

[25] D. Shevitz, and B. Paden. "Lyapunov stability theory of nonsmooth systems." *IEEE Transactions on Automatic Control*, vol. 39, no. 9, pp. 1910-1914, Sept. 1994.

[26] J. Cortes. "Discontinuous dynamical systems." *IEEE Control Systems Magazine,* vol. 28, no. 3, pp. 36-73, June 2008.

[27] S. Sun and W. Ren. "Distributed continuous-time optimization with time-varying objective functions and inequality constraints." arXiv preprint arXiv:2009.02378, 2020.

[28] S. Boyd, and L. Vandenberghe. "*Convex optimization*." Cambridge, U.K.: Cambridge University Press, 2004.

[29] D. G. Luenberger, and Y. Ye. "Linear and nonlinear programming." Reading, U. S.: Addison-Wesley, 1984.

Authorized licensed use limited to: Univ of Calif Riverside. Downloaded on July 22,2021 at 01:15:08 UTC from IEEE Xplore. Restrictions apply.