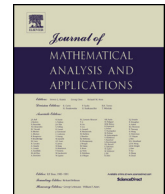




Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa



A consensus algorithm in $CAT(0)$ space and its application to distributed fusion of phylogenetic trees

Sheng Chen ^{a,*}, Peng Shi ^b, Cheng-Chew Lim ^b, Zhenyu Lu ^c^a School of Information & Control, B-DAT, CICAET, Nanjing University of Information Science & Technology, Nanjing 210044, China^b School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia^c School of Electronics & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China

ARTICLE INFO

Article history:

Received 5 December 2016

Available online xxxx

Submitted by P.L. Combettes

Keywords:

Geometric space

Consensus algorithm

Data fusion

Tree-type data

ABSTRACT

Based on convex analysis, a novel consensus algorithm of dynamical points in a $CAT(0)$ space is developed in this paper, in which the associated communication graph uniformly contains a directed spanning tree. The proposed algorithm provides an efficient method of solving consensus problems in a general $CAT(0)$ space, while having certain robustness against weak communication. The application of the new algorithm to the distributed fusion of phylogenetic trees is shown with demonstrative-case simulations, together with a study on the algorithm's robustness and efficiency.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

In practice, there exists a class of distributed computing problems of how to guarantee networked dynamical points achieving consensus in a space. The problems are found in domains ranging from multi-sensor data fusion to multi-body motion coordination, distributed synchronization control, group decision-making and parallel computing. In fact, after the average consensus algorithms [20] were proposed, consensus problems have attracted considerable attention in past years, leading to fruitful results [10,14,31,32], such as nonlinear consensus and finite-time consensus. See [6,15] and the references therein for some recent developments.

In past years, several consensus algorithms [11,25,26,29] on manifold spaces were proposed. One potential application is to fuse rotation attitudes for synchronizing rigid entities. Some of our recent works [7–9] were also devoted to this aspect. Recall that the tree structure is an efficient way for organizing data in different contexts, resulting in phylogenetic trees, decision trees, knowledge trees and language trees. See [17] for some

* Corresponding author.

E-mail address: chenshengy@163.com (S. Chen).

latent tree models and their applications. An engineering-related problem is to introduce certain efficient consensus algorithms for fusing tree-type data, such as fusing estimated phylogenetic trees for the real one in biology.

As a candidate space model for tree-type data, recall that the space of phylogenetic m -trees is a globally-nonlinear $\text{CAT}(0)$ space [4]. To our knowledge, there are few consensus algorithms proposed for $\text{CAT}(0)$ space or Hadamard space. Following the idea of set contraction, Tuna and Spulchre [30] proposed one algorithm by projecting the time-varying position convex hulls of dynamical points and their neighbors onto a common geodesic. Working from the classical work [13], Grohs [12] put forth another algorithm based on the Fréchet means of the weighted positions of dynamical points. In a nonlinear Hadamard space, however, the verification of the ω -contraction condition in [30] is not an easy task, and each running step in Grohs's consensus algorithm involves the approximation iteration problem in searching for Fréchet means.

Based on the analysis of convex hulls, this paper proposes a novel consensus algorithm of dynamical points in a $\text{CAT}(0)$ space, in which the associated communication graph uniformly contains a directed spanning tree. Unlike Grohs's consensus algorithm, the algorithm can avoid the approximation iteration problem in searching for Fréchet means in a nonlinear Hadamard space, since it does not involve Fréchet means. The algorithm provides an efficient method of solving consensus problems in a general $\text{CAT}(0)$ space while having certain robustness against weak communication. As an illustrative example for fusing tree-type data, the application of the algorithm to the distributed fusion of phylogenetic trees is shown with demonstrative-case simulations, as well as a study on the algorithm's robustness and efficiency.

The rest of this paper is organized as follows. In the next section, the analysis of convex hulls is presented with some preliminaries on $\text{CAT}(0)$ spaces. In Section 3, the consensus algorithm is shown, and the related theorem is proven. Before we conclude the paper, the algorithm is applied to the fusion problem on phylogenetic trees, together with a study on its robustness and efficiency.

2. Convex analysis

Let (X, d) be a metric space. A geodesic in (X, d) is a continuous curve $\gamma : [0, 1] \rightarrow X$ such that its arclength is minimizing locally and its speed is constant globally. As usual, a geodesic is said to be nonconstant if its speed is nonzero; otherwise, constant. The space (X, d) is called a geodesic metric space if any two points in X have a geodesic between them. For any $x, y, z \in X$, the geodesic triangle $\Delta(x, y, z)$ in (X, d) is a triangle of vertices x, y, z in (X, d) , whose three edges, denoted by $[x, y]$, $[y, z]$, $[z, x]$ respectively, are the geodesics between its vertices. A triangle $\Delta(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ in the Euclidean space $(\mathbb{R}^2, \|\cdot\|)$ is called a comparison triangle of a given geodesic triangle $\Delta(x_1, x_2, x_3)$ in (X, d) , if $d(x_i, x_j) = \|\bar{x}_i - \bar{x}_j\|$ for any $i, j \in \{1, 2, 3\}$.

Consider that a geodesic metric space (X, d) satisfies the $\text{CAT}(0)$ inequality [5], if any geodesic triangle $\Delta(x, y, z)$ in (X, d) and its comparison triangle $\Delta(\bar{x}, \bar{y}, \bar{z})$ satisfy $d(z, e) \leq \|\bar{z} - \bar{e}\|$, where e, \bar{e} are arbitrary points on the edges $[x, y]$ and $[\bar{x}, \bar{y}]$, respectively, subject to $d(x, e) = \|\bar{x} - \bar{e}\|$. A $\text{CAT}(0)$ space is a geodesic metric space satisfying the $\text{CAT}(0)$ inequality [5]. In view of the geodesic uniqueness in $\text{CAT}(0)$ spaces, thereafter, the tuple (X, d) denotes a $\text{CAT}(0)$ space, and the notation $\Gamma_{x,y}$ represents the geodesic $\Gamma_{x,y} : [0, 1] \rightarrow X$ from point x to point y in (X, d) , unless specified otherwise.

Some typical examples of $\text{CAT}(0)$ spaces include real inner product spaces, \mathbb{R} -tree spaces and the spaces of phylogenetic trees. A complete $\text{CAT}(0)$ space is called a Hadamard space [5]. Hadamard space provides more common spaces such as real Hilbert spaces and Hadamard manifolds. Some classical examples of Hadamard manifolds include Euclidean hyperbolic surfaces and the spaces of symmetric positive-definite matrices. See [2,5,27] for more examples of $\text{CAT}(0)$ spaces and Hadamard spaces.

Following the CN inequality of Bruhat and Tits, an equivalent definition of Hadamard spaces was shown in [27], where a Hadamard space is defined as a complete metric space satisfying the NPC (nonpositive curvature) inequality. Following Proposition 2.3 in [27], Lemma 1 below holds when the space (X, d) is the

Euclidean space $(\mathbb{R}^2, \|\cdot\|)$. Then, the CAT(0) inequality implies the lemma. Note that inequality (1) in the lemma guarantees that $d(z, \Gamma_{x,y}(s)) < \max(d(z, x), d(z, y))$ for any $z \in X$, provided that the point $\Gamma_{x,y}(s)$ differs from both x and y .

Lemma 1. *Let (X, d) be a CAT(0) space, and $x, y \in X$ be two points. For any $z \in X$ and $s \in [0, 1]$,*

$$d^2(z, \Gamma_{x,y}(s)) \leq (1-s)d^2(z, x) + sd^2(z, y) - s(1-s)d^2(x, y). \quad (1)$$

For later use, recall some concepts related to the sets in (X, d) . Let $C \subset X$ be a nonempty set. Consider that the set C is singular, called a singleton, if $C = \{p\}$ for a certain point $p \in X$. Let $\text{diam}(C) := \sup_{x,y \in C} d(x, y)$ denote its diameter, and define the set $\text{ext}^*(C) := \{x, y \in C : d(x, y) = \text{diam}(C)\}$. Note that the set $\text{ext}^*(C)$ only concerns the paired points in C supporting diameter $\text{diam}(C)$, hence, $\text{ext}^*(C)$ may be empty. As usual, say that the set C is convex if and only if the geodesic $\Gamma_{p,q}$ lies in the set C whenever the points $p, q \in C$. Assume that the set C is convex. A point $p \in C$ is called an extreme point of C if and only if the point p cannot be an interior point of any nonconstant geodesic lying in C . Let $\text{ext}(C)$ denote the set of all extreme points of the convex set C . Note that the set $\text{ext}(C)$ may also be empty, for example, C is an open interval in the real line space \mathbb{R} .

Lemma 2. *Let (X, d) be a CAT(0) space, and $C \subset X$ be a convex set. The inclusion relationship $\text{ext}^*(C) \subseteq \text{ext}(C)$ holds.*

Proof. It is trivial when the set $\text{ext}^*(C)$ is empty or singular. We consider the case when $\text{ext}^*(C)$ is nonempty and nonsingular, as follows.

Take any $p, q \in \text{ext}^*(C)$ such that $d(p, q) = \text{diam}(C)$. In this case, note that $\text{diam}(C) > 0$, followed by $d(p, q) > 0$. Without loss of generality, assume that $p \notin \text{ext}(C)$. It follows that there exist two distinct points $r_1, r_2 \in C$ and a number $s \in (0, 1)$ such that $\Gamma_{r_1, r_2}(s) = p$. Then, inequality (1) implies $d(p, q) < \max(d(r_1, q), d(r_2, q))$, since p is distinct from both r_1 and r_2 . Note that $d(p, q) = \text{diam}(C)$, and $d(r_1, q), d(r_2, q) \leq \text{diam}(C)$. Then, we have the contradiction $\text{diam}(C) = d(p, q) < \text{diam}(C)$, implying that the assumption does not hold. Hence, $p \in \text{ext}(C)$, followed by $\text{ext}^*(C) \subseteq \text{ext}(C)$. \square

For any nonempty subset C in (X, d) , its convex hull $\text{co}(C)$ in (X, d) is the minimal convex subset of X containing C . It follows that $\text{co}(C_1) \subseteq \text{co}(C_2)$ whenever $C_1 \subseteq C_2 \subset X$. According to [24], any convex hull $\text{co}(C)$ in (X, d) can be constructed by

$$\text{co}(C) = \bigcup_{k=0}^{\infty} A_k(C) \quad \text{with } A_0(C) = C, \quad (2)$$

where each set $A_k(C)$ is made up of all points in the geodesics of endpoints in $A_{k-1}(C)$, $k \in \mathbb{N}$. For $A_k(C)$ in construction (2), note that $A_{k-1}(C) \subseteq A_k(C)$ holds for any $k \in \mathbb{N}$, followed by $\text{co}(C) = \lim_{k \rightarrow \infty} A_k(C)$.

Following the Krein–Milman theorem in [18], any compact convex set in a Hadamard space has extreme points, which implies that some noncompact convex sets in a Hadamard space may have no extreme point. Note that the convex hull of a (nonempty) finite set in a Hadamard space is not necessarily compact, since it is not necessarily closed [19]. It follows the possibility that there exists a finite set in a certain Hadamard space such that its convex hull has no extreme point. Beyond this intuition, the next lemma implies that the convex hull of any finite set C in a CAT(0) spaces has extreme points, because the set $\text{ext}^*(C)$ is nonempty for any finite set C .

Lemma 3. *Let (X, d) be a CAT(0) space, and $C \subset X$ be a finite set. The set C satisfies that*

$$\text{ext}^*(C) = \text{ext}^*(\text{co}(C)) \subseteq \text{ext}(\text{co}(C)) \subseteq C. \quad (3)$$

Proof. It is trivial when the set C is a singleton. Based on construction (2) of the finite set C , we consider other cases as follows.

We prove that $\text{ext}^*(C) = \text{ext}^*(\text{co}(C))$ as follows. Take any integer $k \in \mathbb{N}$, and any point $x \in A_k(C) \setminus A_{k-1}(C)$, where the sets $A_{k-1}(C)$ and $A_k(C)$ are given in (2). Then, there are a number $s \in (0, 1)$ and two distinct points $p, q \in A_{k-1}(C)$ such that $\gamma_{p,q}(s) = x$. Now, consider any point $x' \in A_k(C)$. If the point $x' \in A_{k-1}(C)$, then by (1) we have

$$d(x, x') < \max(d(p, x'), d(q, x')) \leq \text{diam}(A_{k-1}(C)).$$

If the point $x' \in A_k(C) \setminus A_{k-1}(C)$, then there exist a number $s' \in (0, 1)$ and two distinct points $p', q' \in A_{k-1}(C)$ such that $\gamma_{p',q'}(s') = x'$. It follows that

$$\begin{cases} d(x, x') < \max(d(x, p'), d(x, q')), \\ d(x, p') \leq \max(d(p, p'), d(q, p')), \\ d(x, q') \leq \max(d(p, q'), d(q, q')), \end{cases}$$

according to (1). Then, we also have $d(x, x') < \text{diam}(A_{k-1}(C))$ when the point $x' \in A_k(C) \setminus A_{k-1}(C)$. Thus, for each $k \in \mathbb{N}$, $d(x, x') < \text{diam}(A_{k-1}(C))$ holds for any $x' \in A_k(C)$, when the point $x \in A_k(C) \setminus A_{k-1}(C)$. It implies that $\text{diam}(C) = \text{diam}(A_{k-1}(C))$ for each $k \in \mathbb{N}$, and $d(x, x') < \text{diam}(C)$ whenever $x \in \text{co}(C) \setminus C$ and $x' \in \text{co}(C)$. Hence, $\text{ext}^*(C) = \text{ext}^*(\text{co}(C))$.

According to (2), any $p \in \text{co}(C) \setminus C$ must be an interior point of a certain nonconstant geodesic in $\text{co}(C)$. It follows that $p \notin \text{ext}(\text{co}(C))$. In other words, $\text{ext}(\text{co}(C)) \subseteq C$. In addition, Lemma 2 implies that $\text{ext}^*(\text{co}(C)) \subseteq \text{ext}(\text{co}(C))$. Hence, this proposition holds. \square

3. Consensus algorithm

For describing the communication topology between dynamical points, we recall the following concepts from graph theory. Let $\mathbb{G}^n = \{G_1, G_2, \dots\}$ denote the set of all simple digraphs of the vertex set $\bar{n} := \{1, 2, \dots, n\}$, and S denote the subscript set of all elements in \mathbb{G}^n . Define $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For any given map $\sigma : \mathbb{N}_0 \rightarrow S$, the time-varying digraph $G_{\sigma(t)}$ is called a communication graph. Let $E_{\sigma(t)}$ denote the edge set of $G_{\sigma(t)}$. For each $i \in \bar{n}$, define its neighbor set as $N_i(G_{\sigma(t)}) := \{j \in \bar{n} : \langle j, i \rangle \in E_{\sigma(t)}\}$, and set $N_i^+(G_{\sigma(t)}) := N_i(G_{\sigma(t)}) \cup \{i\}$. Recall that a digraph $G \in \mathbb{G}^n$ contains a directed spanning tree, if there exists a vertex $k \in \bar{n}$ such that any distinct vertex $j \in \bar{n}$ has a directed path from k to j in G . We say that a communication graph $G_{\sigma(t)} \in \mathbb{G}^n$ uniformly contains a directed spanning tree, if any $\tau \in \mathbb{N}_0$ has a finite length $l_\tau \in \mathbb{N}_0$ such that the digraph $\mathcal{G}(\tau) \in \mathbb{G}^n$ having its edge set $\bigcup_{t \in [\tau, \tau + l_\tau] \cap \mathbb{N}_0} E_{\sigma(t)}$ contains a directed spanning tree. As usual, the digraph $\mathcal{G}(\tau)$ is called the union graph of $G_{\sigma(t)}$ across $[\tau, \tau + l_\tau] \cap \mathbb{N}_0$.

For n dynamical points $x_1, x_2, \dots, x_n : \mathbb{N}_0 \rightarrow X$ in (X, d) , say that they achieve consensus, if there exists a point $c \in X$ such that $\lim_{t \rightarrow \infty} x_i(t) = c$ for each $i \in \bar{n}$, where the initial positions $x_1(0), x_2(0), \dots, x_n(0)$ are given arbitrarily. The point c is called their common position, if the dynamical points achieve consensus. Motivated by the distributed fusion of phylogenetic trees, we introduce a consensus algorithm in a $\text{CAT}(0)$ space, encapsulated in the following theorem.

Theorem 1. Let $G_{\sigma(t)} \in \mathbb{G}^n$ be a communication graph, and (X, d) be a $\text{CAT}(0)$ space. Consider the following discrete-time dynamical system

$$x_i(t+1) = \Gamma_{x_i(t), x_{i_i^*}(t)}(s_i(t)), \quad i \in \bar{n}, \quad (4)$$

in (X, d) , where each time-varying subscript $i_i^* \in N_i^+(G_{\sigma(t)})$ satisfies

$$d(x_i(t), x_{i_t^*}(t)) \geq \max_{j \in N_i^+(G_{\sigma(t)})} d(x_i(t), x_j(t)), \quad (5)$$

and each time-varying parameter $s_i(t) \in [s_{\min}, s_{\max}] \subset (0, 1)$ with constants s_{\min} and s_{\max} . All dynamical points x_i in system (4) achieve consensus if and only if $G_{\sigma(t)}$ uniformly contains a directed spanning tree.

At any t -th step of the proposed algorithm, $t \in \mathbb{N}_0$, by Theorem 1, each dynamical point x_i chooses a dynamical point $x_{i_t^*}$ from the available set $\{x_j : j \in N_i^+(G_{\sigma(t)})\}$ subject to (5), i.e., the dynamical point $x_{i_t^*}$ should have the largest distance to x_i at time t ; meanwhile, each dynamical point x_i finds the point $\Gamma_{x_i(t), x_{i_t^*}(t)}(s_i(t))$ on the geodesic $\Gamma_{x_i(t), x_{i_t^*}(t)}$, and updates its position as $\Gamma_{x_i(t), x_{i_t^*}(t)}(s_i(t))$ for the next step. Following Theorem 1, it is clear that each step of the proposed algorithm calls the geodesic algorithm (covering the corresponding distance computation) in the space (X, d) at most n^2 times. In the case when (X, d) is a nonlinear Hadamard space, the next paragraph demonstrates that Grohs's consensus algorithm usually involves an approximation iteration problem in finding Fréchet means, which implies that the usage count of the geodesic algorithm is usually sufficiently large at each step of the algorithm in theory. In other words, the proposed algorithm theoretically has better efficiency than Grohs's consensus algorithm in a nonlinear Hadamard space.

Recall that any m points p_1, p_2, \dots, p_m in a Hadamard space (H, ρ) have their unique Fréchet mean, as $\operatorname{argmin}_{z \in H} \sum_{j \in \overline{m}} \omega_j \rho^2(z, p_j)$ in (H, ρ) , where all $\omega_j > 0$. If the Hadamard space (H, ρ) is nonlinear, it is accepted that there usually has no explicit formula of Fréchet means in the space. In past years, some geodesic-based approximation algorithms and methods have been proposed for finding Fréchet means, such as Sturm's algorithm in Hadamard spaces and descent gradient methods on Hadamard manifolds. For Grohs's consensus algorithm in a nonlinear Hadamard space, its approximation algorithm is required in practice for running on a computer. Note that at each step of the approximation algorithm, the approximation criterion about Fréchet means should be set sufficiently small, otherwise, it becomes an approximation algorithm of another consensus algorithm. In addition, small approximation criterion implies large iteration count, followed by large usage count of the geodesic algorithm. Hence, for Grohs's consensus algorithm in a nonlinear Hadamard space, the usage count of the geodesic algorithm at each step is usually large.

For system (4), we define the time-varying convex hull $\operatorname{co}(t) := \operatorname{co}(\{x_i(t) : i \in \overline{n}\})$ for any $t \in \mathbb{N}_0$. It is clear that $\operatorname{co}_i(t) \subseteq \operatorname{co}(t)$ for any $i \in \overline{n}$, where $\operatorname{co}_i(t) := \operatorname{co}(\{x_j(t) : j \in N_i^+(G_{\sigma(t)})\})$ is the convex hull of dynamical point x_i at time t . Lemma 4 below shows that $\operatorname{co}(t)$ is non-expansive with respect to $t \in \mathbb{N}_0$. In other words, given any $\tau \in \mathbb{N}_0$, the convex hull $\operatorname{co}(\tau)$ is an invariant set for all dynamical points x_i in system (4) after time τ . Note that the result in Lemma 4 is independent of the communication graph.

Lemma 4. Consider the time-varying convex hull $\operatorname{co}(t)$ for system (4). The inclusion relationship $\operatorname{co}(t_1) \subseteq \operatorname{co}(t_2)$ holds whenever $t_1 \geq t_2$, $t_1, t_2 \in \mathbb{N}_0$.

Proof. Take any time $t_1 \in \mathbb{N}_0$. Following (4), any updated state $x_i(t_1 + 1) \in \operatorname{co}_i(t_1)$, $i \in \overline{n}$. Hence, all $x_i(t_1 + 1) \in \operatorname{co}(t_1)$ since each $\operatorname{co}_i(t_1) \subseteq \operatorname{co}(t_1)$. It follows that $\operatorname{co}(t_1 + 1) \subseteq \operatorname{co}(t_1)$, according to the definition of convex hulls. Then, the induction method guarantees this proposition. \square

In view of the monotonicity of $\operatorname{co}(t)$ shown above, there exists the limit set $\lim_{t \rightarrow \infty} \operatorname{co}(t)$, denoted by $\operatorname{co}(\infty)$, and the limit set $\operatorname{co}(\infty)$ is convex. In addition, the monotonicity guarantees that the diameter $\operatorname{diam}(\operatorname{co}(t))$ of the convex hull $\operatorname{co}(t)$ is non-increasing with respect to $t \in \mathbb{N}_0$. It follows that the limit $\operatorname{diam}(\operatorname{co}(\infty)) := \lim_{t \rightarrow \infty} \operatorname{diam}(\operatorname{co}(t))$ exists by the Supreme Axiom of \mathbb{R} . It is clear that all dynamical points x_i in (4) achieve consensus if and only if $\operatorname{diam}(\operatorname{co}(\infty)) = 0$. For any x_i and $t \in \mathbb{N}_0 \cup \{\infty\}$, we treat any possible value $x_i(t)$ as a possible position of x_i at time t . Then, all dynamical points x_i in (4) achieve consensus if and only if the possible positions of all x_i is a common point when $t = \infty$.

Proof of Theorem 1. (Necessity Part.) For the limit set $\text{co}(\infty)$, assume that its diameter $\text{diam}(\text{co}(\infty)) > 0$. Under the preassumption that the communication graph $G_{\sigma(t)}$ uniformly contains a directed spanning tree, we now prove that the assumption does not hold as follows.

Following Lemma 3, $\text{ext}^*(\text{co}(t)) = \text{ext}^*(\{x_i(t) : i \in \bar{n}\})$ for each $t \in \mathbb{N}_0$. It follows that $\text{diam}(\text{co}(t)) = \max_{i,j \in \bar{n}} d(x_i(t), x_j(t))$. Note that \bar{n} is a finite set and $\text{co}(\infty)$ is a limit set. Therefore, there exist two subscripts $i, j \in \bar{n}$ and two corresponding divergent sequences $\{\tau_\kappa\}_{\kappa \in \mathbb{N}}, \{\bar{\tau}_\kappa\}_{\kappa \in \mathbb{N}}$ in \mathbb{N} such that $\lim_{\kappa \rightarrow \infty} x_i(\tau_\kappa) = e$ and $\lim_{\kappa \rightarrow \infty} x_j(\bar{\tau}_\kappa) = \bar{e}$, where $e, \bar{e} \in \text{co}(\infty)$ are certain fixed points such that $d(e, \bar{e}) = \text{diam}(\text{co}(\infty))$. Note that $\text{co}(\infty)$ is a convex set. It follows that the points $e, \bar{e} \in \text{ext}(\text{co}(\infty))$ by Lemma 2. In addition, the points e and \bar{e} are distinct by the assumption.

Consider dynamical point x_i with the sequence $\{\tau_\kappa\}_{\kappa \in \mathbb{N}}$. Since the fixed point $e \in \text{ext}(\text{co}(\infty))$, we have $\lim_{t \rightarrow \infty} \text{co}_i(t) = \{e\}$ by $\lim_{\kappa \rightarrow \infty} x_i(\tau_\kappa) = e$. Indeed, for any $t \in \mathbb{N}$, system (4) implies the rule that $x_i(t) \in \text{ext}(\text{co}_i(t-1))$ if and only if $\text{co}_i(t-1) = \{x_i(t)\}$; in particular, the rule still holds in the case $t = \infty$, since the parameter $s_i(t)$ in (4) belongs to the constant interval $[s_{\min}, s_{\max}] \subset (0, 1)$, where $\text{co}_i(t-1)$ in this case is understood as an arbitrary possible convex hull of dynamical point x_i under a possible neighbor set and a possible neighbor position set when $t = \infty$. For the case $t = \infty$, note that $\lim_{\kappa \rightarrow \infty} x_i(\tau_\kappa) = e$ implies that e is a possible position of x_i . In addition, we have $e \in \text{ext}(\text{co}_i^p(e))$ since $e \in \text{ext}(\text{co}(\infty))$ and $\text{co}_i^p(e) \subseteq \text{co}(\infty)$, where $\text{co}_i^p(e)$ represents an arbitrary possible convex hull $\text{co}_i^p(e)$ of dynamical point x_i at the possible position e . Hence, in the case $t = \infty$, if there exists a point in the set $\text{co}(\infty) \setminus \{e\}$ being a possible position of x_i , then the rule implies that point e cannot be a possible position of x_i , in view of $e \in \text{ext}(\text{co}_i^p(e))$. In other words, we have $\lim_{t \rightarrow \infty} x_i(t) = e$. It follows that $\lim_{t \rightarrow \infty} \text{co}_i(t) = \{e\}$, according to the rule and $e \in \text{ext}(\text{co}_i^p(e))$.

For dynamical point x_i with the sequence $\{\tau_\kappa\}_{\kappa \in \mathbb{N}}$, the above analysis shows that $\lim_{\kappa \rightarrow \infty} x_i(\tau_\kappa) = e$ implies $\lim_{t \rightarrow \infty} \text{co}_i(t) = \{e\}$. Therefore, for any subscript $k \in \bar{n}$, if there exists a divergent sequence $\{\tau'_\kappa\}_{\kappa \in \mathbb{N}}$ in \mathbb{N} such that $\lim_{\kappa \rightarrow \infty} x_k(\tau'_\kappa) = e$, we have $\lim_{t \rightarrow \infty} \text{co}_k(t) = \{e\}$. In other words, the point e has a maximal subset $\Omega_e \subset \bar{n}$ such that $\lim_{t \rightarrow \infty} \text{co}_k(t) = \{e\}$ for each $k \in \Omega_e$. Similarly, the point \bar{e} also has a maximal subset $\Omega_{\bar{e}} \subset \bar{n}$ such that $\lim_{t \rightarrow \infty} \text{co}_k(t) = \{\bar{e}\}$ for any $k \in \Omega_{\bar{e}}$.

Consider the case $t = \infty$. Since each $k \in \Omega_e$ satisfies $\lim_{t \rightarrow \infty} \text{co}_k(t) = \{e\}$, there is no directed edge from a vertex $v' \in \bar{n} \setminus \Omega_e$ to a vertex $v \in \Omega_e$ in the union graph of all possible consequences of $G_{\sigma(t)}$. For the set $\Omega_{\bar{e}}$, we have similar result. Note that the intersection $\Omega_e \cap \Omega_{\bar{e}}$ is empty, since the points e, \bar{e} are distinct. Then, the above results about Ω_e and $\Omega_{\bar{e}}$ yield a contraction with the preassumption. Indeed, the preassumption ensures that the union graph contains a directed spanning tree, whereas the results imply that the union graph contains none. Following the contradiction, the assumption does not hold, hence, all dynamical points in (4) achieve consensus.

(Sufficiency Part.) Note that achieving consensus is required for arbitrary initial positions $x_1(0), x_2(0), \dots, x_n(0)$. Without loss of generality, assume that the union graph of $G_{\sigma(t)}$ across $[0, \infty) \cap \mathbb{N}_0$ contains no directed spanning tree. It follows that there exists a pair of complementary sets $V_1, V_2 \in \bar{n}$ such that there is no directed edge between V_1 and V_2 in $G_{\sigma(t)}$ for any $t \in \mathbb{N}_0$. Let $\text{co}^k(t) := \text{co}(\{x_i(t) : i \in V_k\})$ for each $k \in \{1, 2\}$ and $t \in \mathbb{N}_0$, and consider the case where $\text{co}^1(0) \cap \text{co}^2(0) = \emptyset$. Following Lemma 4, we have $\text{co}^k(t) \subseteq \text{co}^k(0)$ for any $k \in \{1, 2\}$ and $t \in \mathbb{N}_0$. Then, it is clear that the dynamical points cannot achieve consensus in this case. \square

4. Algorithm application

Using DNA data to infer evolutionary relationship between species is a central problem in biology. It is known that gene duplication and loss, incomplete lineage sorting, lateral gene transfer and other factors may lead to the problem that a group of species possesses multiple estimated phylogenetic trees. Working towards obtaining the real phylogenetic tree, many methods were proposed around these factors. See [28] for the history and some recent developments. Considering the factor complexity, there is an upward trend of

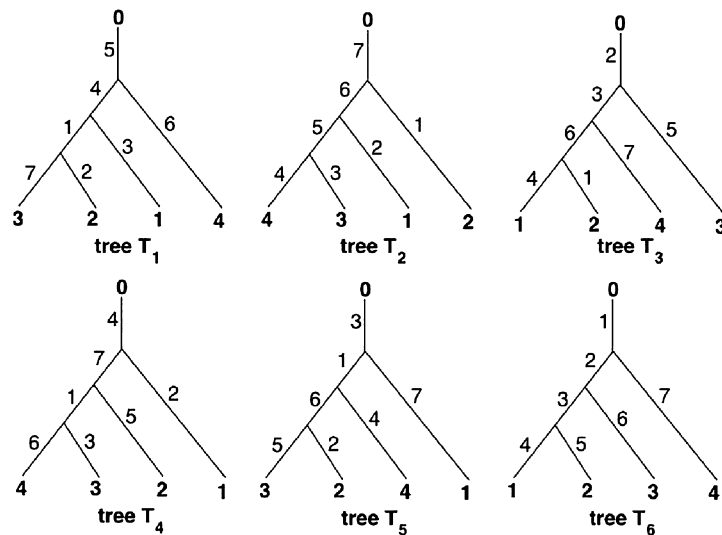


Fig. 1. Trees to be fused.

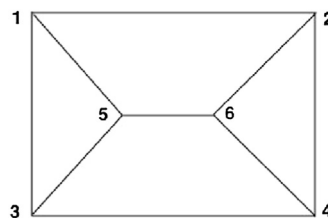
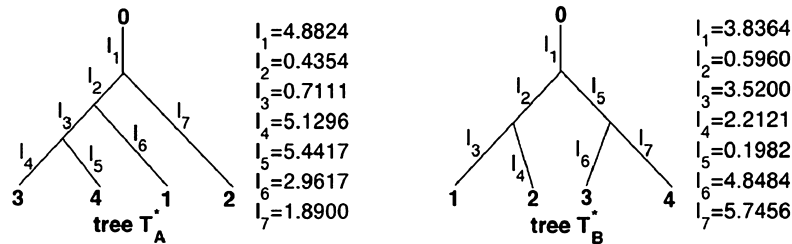
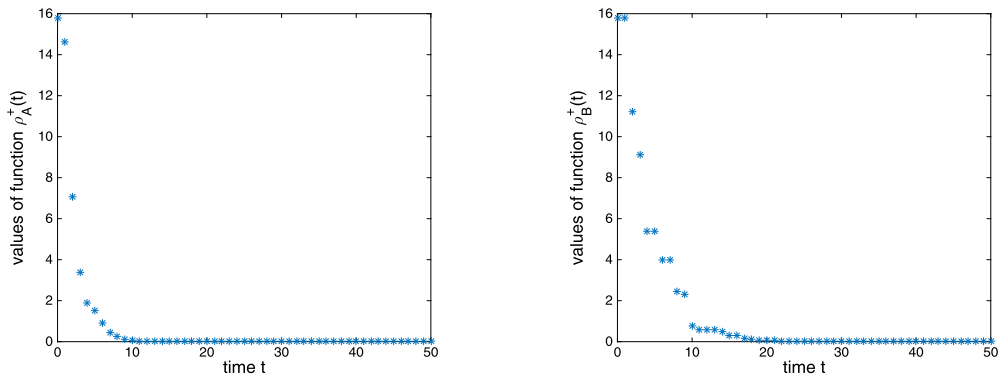


Fig. 2. Graph G.

viewing the problem from the angle of data fusion, such as [3,16]. In view of the distributed construction of phylogenetic trees [23], the viewpoint of distributed fusion seems more efficient in practice. In this section, we apply the proposed algorithm to the distributed fusion of phylogenetic trees with demonstrative-case simulations, and study its robustness and efficiency based on the simulation results.

Consider a usual case of the distributed fusion problem of weighted phylogenetic m -trees T_1, T_2, \dots, T_n , associated with a communication graph $G_{\sigma(t)} \in \mathbb{G}^n$ representing the communication topology of tree holders, where each tree holder as a computer has exactly one tree. The $n = 6$ demonstrative trees T_1, T_2, \dots, T_6 in Fig. 1, with all weights $\omega(T_i) = 1$, $i \in \bar{n}$, are the 4-trees to be fused, where the bold numbers are the respective labels of tree roots and leaves identifying the interested species, and other numbers are edge lengths representing the expected mutation numbers per DNA character site [21]. The associated communication graph could be the undirected graph G in Fig. 2, in which any undirected edge (i, j) represents two directed edges $\langle i, j \rangle$ and $\langle j, i \rangle$.

To demonstrate the applicability of our proposed algorithm to the fusion problem, recall the space of phylogenetic m -trees as follows. Following the seminal work [4], there exists a one-to-one map between the set \mathcal{T}_m of m -trees and the pasted space of some nonnegative Euclidean orthants along their common faces, leading to the induced geodesic metric d_g for the set \mathcal{T}_m . It turns out that the space (\mathcal{T}_m, d_g) is a CAT(0) space [4]. Note that the distance function d_g is defined on the inner edges of m -trees, and the vector of ordered root-edge and leaf-edge lengths in any m -tree identifies a point in the positive Euclidean orthant \mathcal{O} of dimension $m + 1$. Hence, the direct product space (\mathcal{T}_m, d_g^+) of the spaces (\mathcal{T}_m, d_g) and $(\mathcal{O}, \|\cdot\|)$ is also a CAT(0) space, since both of its subspaces are, where $\|\cdot\|$ is the usual Euclidean norm. In view of the global nonlinearity of these spaces, Owen and Provan [22] proposed the GTP (geodesic treepath problem)

Fig. 3. Fused trees T_A^* and T_B^* .Fig. 4. Functions $\rho_A^+(t)$ and $\rho_B^+(t)$.

algorithm for computing the geodesics and distances in (\mathcal{T}_m, d_g^+) . In this section, any referred geodesic or distance algorithm in (\mathcal{T}_m, d_g^+) always means the GTP algorithm.

Since the space (\mathcal{T}_m, d_g^+) is a CAT(0) space, one can regard any tree T_i as the initial position of its corresponding dynamical points x_i in (4), $i \in \bar{n}$, and apply the algorithm to searching for the fused tree of the trees T_1, T_2, \dots, T_n with weights $\omega(T_1), \omega(T_2), \dots, \omega(T_n)$. In the application, each time-varying parameter $s_i(t)$ in (4) can be chosen as $\omega(T_{i_t^*})/(\omega(T_i) + \omega(T_{i_t^*}))$. An alternative choice is $s_i(t) = \omega_i(t)/(\omega(T_i) + \omega_i(t))$, where $\omega_i(t) = \sum_{j \in N_i^+(G_{\sigma(t)})} \omega(T_j)$. The communication graph in the algorithm is just the one in association with the tree holders. Following Lemma 4, note that the algorithm is a consensus algorithm of asymptotic convergence, hence, the dynamical points usually cannot achieve consensus in finite. Let $e_f > 0$ be a small number, called the fusion criterion. A point $T^* \in \mathcal{T}_m$ is called the fused tree of the weighted trees under the fusion criterion e_f , if and only if $d_g^+(T^*, x_c) < e_f$ holds, where $x_c \in \mathcal{T}_m$ is the common position of all dynamical points x_i after they achieve consensus, $i \in \bar{n}$. In practice, the condition $\max_{i,j \in \bar{n}} d_g^+(x_i(t), x_j(t)) < e_f$ is enough to obtain the fused tree T^* at time t , according to Lemma 4.

To demonstrate the above application, consider the usual case of the distributed fusion problem discussed above. For the trees T_1, T_2, \dots, T_6 with the graph G , after running the proposed algorithm with the fusion criterion $e_f = 0.0001$, we have their fused tree T_A^* , which is shown in Fig. 3. During the simulation, the function $\rho_A^+(t) := \max_{i,j \in \bar{n}} d_g^+(x_i(t), x_j(t))$ is set up to supervise the fusion performance. Its value-variation shown in Fig. 4 implies the tendency that $\rho_A^+(t) \rightarrow 0$ when $t \rightarrow \infty$. In fact, we have $\rho_A^+(t) < e_f$ for any $t \geq 22$ by the simulation data. In other words, under the fusion criterion e_f , the fused tree T_A^* can be achieved at time $t = 22$.

The algorithm applicability is further verified by an extreme case of the distributed fusion of the trees T_1, T_2, \dots, T_6 . Let $G_{\sigma(t)} \in \mathbb{G}^n$ be the communication graph such that, for each $t \in \mathbb{N}_0$, any directed edge in the digraph $G_{\sigma(t)}$ is generated randomly with probability 0.2. This means that the communication between the tree holders is directed and dynamically switching. Note that random graph theory implies that $G_{\sigma(t)}$ uniformly contains a directed spanning tree in the sense of large probability event. All time-varying parameters $s_i(t)$ in (4) have their random values in $[0.1, 0.9]$, covering the case where the weights need

Table 1
Running data of Sturm's algorithm.

	$k = 996$	$k = 997$	$k = 998$	$k = 999$	$k = 1000$
$e_1(k)$	0.0070	0.0070	0.0089	0.0048	0.0073
$e_2(k)$	0.0057	0.0089	0.0057	0.0089	0.0075
$e_3(k)$	0.0079	0.0051	0.0075	0.0078	0.0075
$e_4(k)$	0.0056	0.0089	0.0075	0.0075	0.0075
$e_5(k)$	0.0049	0.0049	0.0049	0.0069	0.0069
$e_6(k)$	0.0091	0.0072	0.0072	0.0055	0.0076

to be dynamically adjusted. Then, after the new simulation is performed with the fusion criterion $e_f = 0.0001$, the fused tree T_B^* of T_1, T_2, \dots, T_6 , and the corresponding value-variation of function $\rho_B^+(t) := \max_{i,j \in \bar{n}} d_g^+(x_i(t), x_j(t))$, are also obtained as expected. See Fig. 3 and Fig. 4. These new simulation results validate the algorithm applicability in certain extreme cases, and demonstrate that the algorithm has certain robustness against weak communication in the application.

In the rest of the section, we study the efficiency of the proposed algorithm using the simulation-based comparison analysis method. Let (H, d) denote the completion of the tree space (\mathcal{T}_m, d_g^+) , and $x_1, x_2, \dots, x_n : \mathbb{N}_0 \rightarrow H$ be n dynamical points in the Hadamard space (H, d) such that

$$x_i(t+1) = \operatorname{argmin}_{z \in H} \sum_{j \in N_i^+(G_{\sigma(t)})} d^2(z, x_j(t)), \quad i \in \bar{n}, \quad (6)$$

where $G_{\sigma(t)} \in \mathbb{G}^n$ is a communication graph uniformly containing a directed spanning tree. Note that Corollary 1.9 in [12] guarantees that all dynamical points x_i in (6) achieve consensus for arbitrary initial positions, i.e., the algorithm encapsulated in (6) is a consensus algorithm. Following system (6), the algorithm is the simplest version of Grohs's consensus algorithm, and is called Grohs's algorithm for convenience.

For searching Fréchet means in system (6), recall that there are some approximation algorithms and methods, such as the algorithm of random order [1], Sturm's algorithm [27] and the descent method [16]. Note that the descent method is designed on Sturm's algorithm, and the algorithm of random order is similar to Sturm's algorithm. In this paper, we use Sturm's algorithm. A brief version of Sturm's algorithm is given as follows. Let p_1, p_2, \dots, p_m be m points in a Hadamard space (H, d) . Take any point $z_0 \in \{p_1, p_2, \dots, p_m\}$. For each $k \in \mathbb{N}_0$, choose a number $r(k) \in \bar{m}$ randomly according to the uniform distribution, and set

$$z_{k+1} = \Gamma_{z_k, p_{r(k)}}(1/(k+1)). \quad (7)$$

Then, the sequence $\{z_k\}_{k \in \mathbb{N}_0}$ converges to $\operatorname{argmin}_{z \in H} \sum_{j \in \bar{m}} d^2(z, p_j)$ in the sense of large probability event. See [16] for the detailed version.

Consider the application of Grohs's algorithm to the distributed fusion of the trees T_1, T_2, \dots, T_6 . For system (6), set the number $n = 6$ and define the initial position $x_i(0) = T_i$ for any $i \in \bar{n}$. For efficiency comparison, the communication graph $G_{\sigma(t)}$ in (6) is also chosen as the graph G . Then, for the first step of Grohs's algorithm, we have the running data of Sturm's algorithm in Table 1, where any $e_i(k) = d_g^+(z_k, z_{k+1})$ is the distance difference for dynamical point x_i , $i \in \bar{n}$. For Sturm's algorithm, it is reasonable to set its approximation criterion $e_a < e_f$ in the sense of large probability event, where $e_f = 0.0001$ is the fusion criterion in the application. Then, to get the positions $x_i(1)$ of all dynamical point x_i in (6), the running data in Table 1 shows that the geodesic algorithm will be called at least 6000 according to formula (7) and the criterion e_a , let alone the computation on the sample investigation about the random sequences $\{e_i(k)\}_{k \in \mathbb{N}_0}$, $i \in \bar{n}$.

For the task of fusing the trees T_1, T_2, \dots, T_6 with the graph G , the above simulation-based analysis implies that Grohs's algorithm will use the geodesic algorithm at least 6000 times. However, for the same fusion task and under the same fusion criterion, our proposed algorithm in the previous first simulation uses

the geodesic algorithm at most 1000 times, where the usage of the distance algorithm is treated as that of the geodesic algorithm since both of them are the GTP algorithm. Therefore, the above comparison analysis demonstrates that our proposed algorithm offers better efficiency than Grohs's algorithm in the fusion case that we considered.

5. Conclusions

This paper presents a consensus algorithm of dynamical points in a $CAT(0)$ space, and demonstrates its application to the distributed fusion problem of phylogenetic trees, together with a study on its robustness and efficiency. All the work is helpful for the fusion of tree-type data in practice and the motion control in a nonlinear space. In the future, we will explore the consensus of dynamical points in a metric space of curvature bounded above.

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China (61503191, 61773131), the Natural Science Foundation of Chinese Jiangsu Province (BK20150933), the Joint Key Grant of National Natural Science Foundation of China and Zhejiang Province (U1509217), and the Australian Research Council (DP170102644).

References

- [1] M. Bačák, Computing medians and means in Hadamard spaces, *SIAM J. Optim.* 24 (3) (2014) 1542–1566.
- [2] W. Ballmann, *Lecture on Spaces of Nonpositive Curvature*, Birkhäuser, Boston, 1995.
- [3] D. Barden, H. Le, M. Owen, Central limit theorems for Fréchet means in the space of phylogenetic trees, *Electron. J. Probab.* 18 (25) (2013) 1–25.
- [4] L.J. Billera, S.P. Holmes, K. Vogtmann, Geometry of the space of phylogenetic trees, *Adv. in Appl. Math.* 27 (4) (2001) 733–767.
- [5] M.R. Bridson, A. Haefliger, *Metric Spaces of Non-Positive Curvature*, Springer-Verlag, New York, 1999.
- [6] Y. Cao, W. Ren, D.W. Casbeer, C. Schumacher, Finite-time connectivity-preserving consensus of networked nonlinear agents with unknown Lipschitz terms, *IEEE Trans. Automat. Control* 61 (6) (2016) 1700–1705.
- [7] S. Chen, P. Shi, W. Zhang, L. Zhao, Consensus on complete Riemannian manifolds in finite time, *J. Math. Anal. Appl.* 400 (2) (2013) 497–504.
- [8] S. Chen, P. Shi, W. Zhang, L. Zhao, Finite-time consensus on strongly convex balls of Riemannian manifolds with switching directed communication topologies, *J. Math. Anal. Appl.* 409 (2) (2014) 663–675.
- [9] S. Chen, L. Zhao, W. Zhang, P. Shi, Consensus on compact Riemannian manifolds, *Inform. Sci.* 268 (2014) 220–230.
- [10] J. Cortés, Distributed algorithms for reaching consensus on general functions, *Automatica* 44 (3) (2008) 726–737.
- [11] R. Dong, Z. Geng, Consensus based formation control laws for systems on Lie groups, *Systems Control Lett.* 62 (2) (2013) 104–111.
- [12] P. Grohs, Wolfowitz's theorem and consensus algorithms in Hadamard spaces, *Proc. Amer. Math. Soc.* 143 (10) (2015) 4403–4414.
- [13] A. Jadbabaie, J. Lin, A.S. Morse, Coordination of groups of mobile autonomous agents using nearest neighbor rules, *IEEE Trans. Automat. Control* 48 (6) (2003) 988–1001.
- [14] Z. Lin, B. Francis, M. Maggiore, State agreement for continuous-time coupled nonlinear systems, *SIAM J. Control Optim.* 64 (1) (2007) 288–307.
- [15] Z. Miao, Y. Wang, R. Fierro, Collision-free consensus in multi-agent networks: a monotone systems perspective, *Automatica* 46 (2016) 217–225.
- [16] E. Miller, M. Owen, J.S. Provan, Polyhedral computational geometry for averaging metric phylogenetic trees, *Adv. in Appl. Math.* 68 (C) (2015) 51–91.
- [17] R. Mourad, C. Sinoquet, N.L. Zhang, T. Liu, P. Leray, A survey on latent tree model and applications, *J. Artificial Intelligence Res.* 47 (1) (2013) 157–203.
- [18] C.P. Niculescu, The Krein–Milman theorem in global NPC spaces, *Bull. Math. Soc. Sci. Math. Roumanie Tome* 50 (98) (2007) 343–346.
- [19] C.P. Niculescu, I. Roventă, Schauder fixed point theorem in spaces with global nonpositive curvature, *Fixed Point Theory Appl.* 2009 (2009) 906727.
- [20] R. Olfati-Saber, R.M. Murray, Consensus problems in networks of agents with switching topology and time-delays, *IEEE Trans. Automat. Control* 49 (9) (2004) 1520–1533.
- [21] M. Owen, Computing geodesic distances in tree space, *SIAM J. Discrete Math.* 25 (4) (2011) 1506–1529.

- [22] M. Owen, J.S. Proven, A fast algorithm for computing geodesic distances in tree space, *IEEE/ACM Trans. Comput. Biol. Bioinform.* 8 (1) (2011) 2–13.
- [23] E. Ruzgar, K. Erciyes, Clustering based distributed phylogenetic tree construction, *Expert Syst. Appl.* 39 (1) (2012) 89–98.
- [24] J. Sahattchiev, On convex hull and the quasiconvex subgroups of $F_m \times \mathbb{Z}^m$, *Groups Complex. Cryptol.* 7 (1) (2015) 69–80.
- [25] A. Sarlette, C. Lageman, Synchronization with partial state coupling on $SO(n)$, *SIAM J. Control Optim.* 50 (6) (2012) 3242–3268.
- [26] A. Sarlette, R. Sepulchre, Consensus optimization on manifolds, *SIAM J. Control Optim.* 48 (1) (2009) 56–76.
- [27] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, *Contemp. Math.* 338 (2003) 357–390.
- [28] G.J. Szöllősi, E. Tannier, V. Daubin, B. Boussau, The inference of gene trees with species trees, *Syst. Biol.* 64 (1) (2015) e42–e62.
- [29] R. Tron, B. Afsari, R. Vidal, Riemannian consensus for manifolds with bounded curvature, *IEEE Trans. Automat. Control* 58 (4) (2013) 921–934.
- [30] S.E. Tuna, R. Sepulchre, Consensus under general convexity, in: *Proc. 46th IEEE Conf. Decision and Control*, 2007, pp. 294–299.
- [31] F. Xiao, L. Wang, A. Wang, Consensus problems in discrete-time multiagent systems with fixed topology, *J. Math. Anal. Appl.* 322 (2006) 587–598.
- [32] D. Xie, S. Wang, Consensus of second-order discrete-time multi-agent systems with fixed topology, *J. Math. Anal. Appl.* 387 (2012) 8–16.