2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Hilbert Metric for Nonlinear Consensus with Varying Topology

Dongjun Wu, Anders Rantzer
Department of Automatic Control, Lund University, 22100 Lund, Sweden
Email: {dongjun.wu, anders.rantzer}@control.lth.se

***Abstract*— New results on continuous time nonlinear consensus under varying topology are presented. The results are proved utilizing non Lyapunov based methods, i.e., the Hilbert metric, showing the possibility of further investigation of Hilbert metric for consensus and synchronization problems.**

I. Introduction and preliminaries

Consensus under varying topology is ubiquitous in multi-agent and network dynamical systems. One of the major contributions on this topic was made by Moreau in 2005 [1] for discrete time systems, who showed that consensus is tightly related to the graphical properties of the multi-agent system. Based on [1] and an earlier work [2], Lin *et al.* extended the result in [1] to continuous time nonlinear systems [3]. However, the proof there was much more challenging. In particular, it relied on non-smooth analysis techniques, even though the system was possibly smooth. It is worth mentioning that in the proof of both results, [1] and [3], Lyapunov functions were used to analyze consensus, though in somewhat different ways. In [1], set valued Lyapunov theory was used while [3] relies on a technical result obtained in [4]. In [5], Angeli and Bliman extended Moreau's result to allow arbitrary time delays and relax convexity of the allowed regions for the state transition map of each agent.

The most relevant result to ours in this paper is from the paper [3]. For clarity, we briefly recall the main result obtained in [3]. The system model considered therein is

$$\begin{cases} \dot{x}_1 = f^1_{\sigma(t)}(x_1, \cdots, x_n) \\ \vdots \\ \dot{x}_n = f^n_{\sigma(t)}(x_1, \cdots, x_n) \end{cases} \tag{1}$$

where $x_i \in \mathbb{R}^m$ and σ is a piece-wise constant switching signal, taking values in a finite set $\{1, \cdots, N\}$. For each $p \in \{1, \cdots, N\}$, there is a directed graph $\mathcal{G}_p$ associated with the vector fields $\{f^i_p\}^n_{i=1}$: $\mathcal{G}_p$ has n vertices denoted as $\{1, \cdots, n\}$ and a link (i, j) is in $\mathcal{G}_p$ if f^i_p depends explicitly on x_j. The graph is called *quasi strongly connected (QSC)* if there exists an agent k such that for each agent j, there is a path from k to j. Such an agent is called a *center*. For a switching graph corresponding to system (1), we say the graph is *uniformly quasi strongly connected (UQSC)* if there exists a constant $T > 0$ such that the union of the graph on the interval $[t, t+T]$ for any t is QSC. Let $\mathcal{C}_i$ be the polytope formed by x_i and its neighboring agents and $T_{x_i}\mathcal{C}_i$ the tangent cone of $\mathcal{C}_i$ at x_i (in $\mathbb{R}^m$). The following is the key assumption made in [3]:

Assumption A0: for each $p \in \{1, \cdots, n\}$,
1) f^i_p is locally Lipschitz and f^i_p is in the relative interior of the cone $T_{x_i}\mathcal{C}_i$, or $f^i_p \in \mathrm{ri}(T_{x_i}\mathcal{C}_i)$;
2) the switching signal is piece-wise constant with minimum dwell time τ_D;

Definition 1. *Given a forward invariant[1] set $D \subseteq \mathbb{R}^m$, the system (1) is said to achieve*
1) asymptotic synchronization *on D if $|x_i(t) - x_j(t)| \leq \beta(t, |x_i(0) - x_j(0)|)$ for all $x_i(0), x_j(0) \in D$ and $i, j \in \{1, \cdots, n\}$ for some class $\mathcal{K}$ function β;*
2) exponential synchronization *on D if $|x_i(t) - x_j(t)| \leq ke^{-\lambda t}|x_i(0) - x_j(0)|$ for all $x_i(0), x_j(0) \in D$ and some constants $k, \lambda > 0$.*

We restate one of the main results in [3] below:

Theorem 1 (Lin *et al.* [3])**.** *Under the assumption **A0**, the system achieves global asymptotic synchronization* if and only if *the system (1) is UQSC.*

Note that once the first assumption of **A0** is met, the vector field f^i_p can be written as $f^p_i = \sum a^p_{ij}(x)(x_j - x_i)$ for some non-negative scalar functions a^p_{ij}. Equivalently, this means that the system can be written as $\dot{x} = A_p(x)x$ with $(A_p(x))_{ij} = a^p_{ij}(x)$, in which the matrix $A_p(x)$ is Metzler[2] and row sum zero for all x. System of the form $\dot{x} = A(x)x$ with $A(x)$ being Metzler has been recently considered by Kawano and Cao [6], where they call such systems "virtually positive". Due to the special structure $\dot{x} = A_p(x)x$, consensus of such systems shares a lot in common with linear time varying multi-agent systems. A remarkable result concerning the latter was obtained by Moreau in [7]:

Theorem 2 (Moreau [7])**.** *Consider the LTV system $\dot{x} = A(t)x$. Assume that $A(\cdot)$ is uniformly bounded and piecewise continuous. Assume that, for every time t, $A(t) = (a_{ij}(t)) \in \mathbb{R}^{n\times n}$ is Metzler with zero row sums. If there exists an index $k \in \{1, \cdots, n\}$, a threshold value $\delta > 0$ and an interval length $T > 0$ such that for all $t \geq 0$,*

$$\int_t^{T+t} a_{ik}(s)ds \geq \delta, \quad \forall i \in \{1, \cdots, n\}\backslash\{k\}, \tag{2}$$

then the system synchronizes exponentially.

The proof of Theorem 2 provided in [7] was based Lyapunov analysis, which was quite delicate, see also [8]

[1]The set D is called forward invariant if $x_i(0) \in D$ for all $i \in \{1, \cdots, n\}$ implies $x_i(t) \in D$ for all i and $t \geq 0$.
[2]A matrix A is said to be Metzler if $A_{ij} \geq 0$ for all $i \neq j$.

for a discrete time version. In particular, separable Lyapunov functions were used. This is a technique commonly used in monotone systems, see for example [9], [10], [11], [12].

As we will see in the next section, the result that we are going to present for nonlinear consensus with varying topology is analogous to Theorem 2. But we underscore that, due to the presence of nonlinearity, the same proof technique for Theorem 2 no longer goes through. Instead, we employ another approach – non Lyapunov based – to tackle this problem, i.e., through the analysis of the Hilbert metric. This approach is not new, e.g., in [13] Hilbert metric is used to obtain further relaxation and extension of the results obtained in [1]; in [14], it was used for consensus in non-commutative spaces. The novelty of the results in this paper lies in the following aspects:

1) The system to be studied is more general than (1): instead of having switching topology, we consider general time varying topology, requiring the switching to be only measurable.
2) Our results extend Theorem 1, but the proof strategy is quite different. It does not rely on Lyapunov analysis, instead on contraction properties w.r.t. the Hilbert metric. As a byproduct, we sometimes obtain stronger results: in [3], only asymptotic consensus is concerned, but we can obtain exponential consensus sometimes.
3) We report comparison results in the following sense: consensus can be characterized by comparing the accumulated graph of the system (to be defined) with respect to a static QSC graph.

Notations: $|\cdot|$ stands for Euclidean 2-norm. For a dynamical system, use $\phi(t, t_0, x_0)$ to represent the solution at t from initial state (t_0, x_0). The interior of a set S is denoted $\mathrm{Int}S$. Being X, Y some topological spaces, denote $C(X, Y)$ the space of continuous maps. Given a set $S \subseteq X$, the indicator function $1_S : X \to \{0, 1\}$ is defined to be $1_S(x) = 1$ if $x \in S$ and 0 otherwise. $E(\mathcal{G})$ stands for the edge set of graph $\mathcal{G}$. $\mathbb{1}_n$ a column vector of dimension n with all ones.

II. Main result

A. Problem setting

We consider consensus of continuous time nonlinear systems in the following form

$$\dot{x}_i = h(x_i, t) + \sum_{j=1}^{n} a_{ij}(t, x)(x_j - x_i), \quad i = 1, \cdots, n \quad (3)$$

where the state of each agent is in $\mathbb{R}$. Associated with the system there is a weighted varying directed graph $\mathcal{G}(t, x) = (a_{ij}(t, x))$. We assume

$$a_{ij}(t, x) \begin{cases} > 0 & \text{if } (i, j) \in E(\mathcal{G}(t, x)) \\ = 0 & \text{otherwise} \end{cases}$$

$a_{ij}(t, x) : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}_+$ are continuous in x for all (i, j) and $h : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is continuously differentiable. Note that we require each agent i to have the same first order internal scalar dynamics governed by h. The non-negativity assumption on the coefficients a_{ij} for $(i, j) \in E(\mathcal{G})$ is essential in our setting, which allows us to apply positive mapping theory, or more precisely, the Hilbert metric.

The model (3) is quite general, at least for scalar agent dynamics. For example, this includes the model $\dot{x} = A(t)x$ in Theorem 2 and the model in [3] when h is set to zero.

B. Accumulated graph

We have noted that in both [1] and [3], consensus is related to the accumulation of the graphs over time, either the union of the switching graph in [3], or the integration of the system matrix in [1]. This motivates us to define the accumulated graph for a time-varying graph $\mathcal{G}(t)$, $\forall t \geq 0$.

Definition 2 (Accumulated graph)**.** *Let $\mathcal{G}(\cdot) = (a_{ij}(\cdot))$ be a measurable time-varying graph, i.e., $t \mapsto a_{ij}(t)$ is a measurable function for all $a_{ij}(t) \in E(\mathcal{G}(t))$. The accumulated graph of $\mathcal{G}(\cdot)$ over the interval $[t_1, t_2]$ is the graph $\mathcal{G}|_{t_1}^{t_2}$ defined by the Lebesgue integral*

$$\mathcal{G}|_{t_1}^{t_2} = \int_{t_1}^{t_2} \mathcal{G}(t)\mathrm{d}t.$$

in the sense that $(\mathcal{G}|_{t_1}^{t_2})_{ij} = \int_{t_1}^{t_2} a_{ij}(t)\mathrm{d}t$.

Example 1. *For a system (1) with switching topology, the graph associated with it can be written as $\mathcal{G}(t) = \sum_{i=1}^{N} 1_{A_i}(t)\mathcal{G}_i$ where 1_{A_i} is the indicator function of some measurable sets A_i and $\mathcal{G}_i$ the graph corresponding to $p = i$. Now the union graph on $[t, t + T]$ used in [3] was nothing but the accumulated graph $\mathcal{G}|_t^{t+T}$. This is quite similar to the construction of the Lebesgue integration – define first for simple functions and then allow larger class of functions.*

In the sequel, we establish consensus results by studying the accumulated graph. Our proof strategy has a more geometric flavor based on Hilbert metric. It is this device that makes our extensions to more general graphs possible.

C. The Hilbert metric

Hilbert metric is a metric defined on cones. More precisely, in our setting, a cone is some closed subset $K \subseteq \mathbb{R}^n$ satisfying the following four properties

1) The interior of K is non-empty.
2) For $v, w \in K$, $v + w$ is also in K.
3) For all $\lambda \geq 0$, and $v \in K$, λv is also in K.
4) $K \cap -K = \{0\}$, where $-K := \{-x : x \in K\}$.

For example, the positive orthant $\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x_i \geq 0, \forall i = 1, \cdots, n\}$ is a cone. Given a cone as above, we can define a partial ordering as $x \leq y$ if $y - x \in K$ and $x < y$ if $y - x \in \mathrm{Int}\ K$. For $x, y \in \mathrm{Int}K$, define two numbers $M(x/y) = \inf\{\lambda : x \leq \lambda y\}$ or ∞ if the set is empty, and $m(x/y) = \sup\{\mu : \mu y \leq x\}$. Then the Hilbert metric between x, y is defined as $d(x, y) = \ln \frac{M(x,y)}{m(x,y)}$. Define the diameter of a set $S \subseteq K$ as $\mathrm{diam}(S) = \sup_{x,y \in S} d(x, y)$. We allow $\mathrm{diam}S = +\infty$, e.g., $S = \mathbb{R}_+^n$. On the other hand, for any cone $S \subseteq \mathrm{Int}\mathbb{R}_+^n \cup \{0\}$, we have $\mathrm{diam}S < +\infty$. In fact, $d(x, y) = \ln \frac{\max_i(x_i/y_i)}{\min_i(x_i/y_i)}$, where $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n)$, which is bounded on S. A mapping $A : K \to$

K on a cone is called non-negative. If in addition, A maps the interior of K into its interior, we call A a positive mapping. For example, when $K = \mathbb{R}^n_+$, then a non-negative matrix A represents a non-negative mapping while a positive matrix represents a positive mapping. One of the most important facts about positive mappings is Birkhoff's theorem in which, the requirement of homogeneity is essential, see [15], [16].

However, for general nonlinear systems, we can no longer apply Birkhoff's theorem. New techniques must be introduced. The idea is that although the system does not contract the cone, it may contract some small cones if we sample the system with certain frequency. Fig 1 shows a linear system which does not contracts the positive orthant, but it does contract any subcones containted in the interior of the positive orthant. The same phenomenon appears in nonlinear consensus. Once the sampled system is contractive with respect to those cones under the Hilbert metric, we can conclude that the system contracts to a fixed direction, which is in our case, the span of $\mathbb{1}_n$.

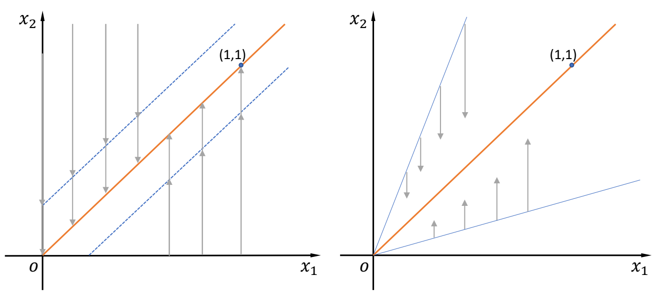

Fig. 1. The gray arrows represent the vector fields of a system. The x_2-axis is invariant and hence cannot be mapped into the interior of the positive orthant. However, the system contracts a smaller cone, see the right.

We propose to study the following type of small cones. For $\gamma \in [0, \frac{1}{\sqrt{n}})$, define a family of cones

$$\mathcal{K}(\gamma) := \left\{ x \in \mathbb{R}^n : \frac{x_i}{|x|} \geq \frac{1}{\sqrt{n}} - \gamma, \ \forall i = 1, \cdots n \right\}$$

see Fig 2. The following simple comparison lemma will be useful to us.

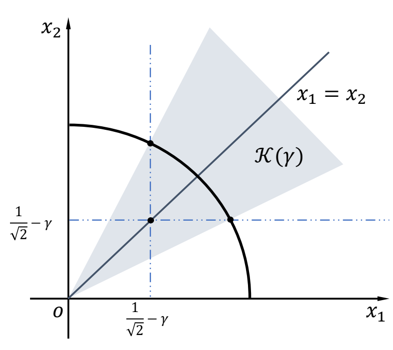

Fig. 2. Cone $\mathcal{K}(\gamma)$ when $n = 2$.

Lemma 1. *For any given $\epsilon_0 \in (0, \frac{1}{\sqrt{n}})$, there exist two positive constants, c_1, c_2 (depending on ϵ_0) such that for all $\epsilon \in [0, \frac{1}{\sqrt{n}} - \epsilon_0)$, we have the following estimate of the diameter of the cone $\mathcal{K}(\epsilon)$:*

$$c_1 \epsilon \leq \mathrm{diam}(\mathcal{K}(\epsilon)) \leq c_2 \epsilon, \quad \forall \epsilon \in [0, \frac{1}{\sqrt{n}} - \epsilon_0)$$

Remark 1. *By Lemma 1, if $0 < C < 1$, we have*

$$\mathrm{diam}\mathcal{K}(C^m \epsilon) \leq \frac{c_2}{c_1} C^m \mathrm{diam}(\mathcal{K}(\epsilon))$$

thus the cone $\mathcal{K}(C^m \epsilon)$ contracts to $\mathrm{span}\{\mathbb{1}_n\}$ exponentially as $m \to \infty$.

D. Main results

In this paper, we assume the internal dynamics for each agent is linear time varying, namely, the function $h(x_i)$ in (3) has the form

$$h(x_i, t) = a(t)x_i + b(t) \tag{4}$$

where $a(\cdot)$ and $b(\cdot)$ are bounded on $\mathbb{R}_+$. For simplicity, we assume $a(t) \leq 0$ for all t.

Remark 2. *The non-positivity assumption of $a(\cdot)$ can be relaxed. For example, we may assume that there exists some $T > 0$ such that*

$$\int_{t_0}^{t_0+T} a(t)\mathrm{d}t \leq 0 \tag{5}$$

for all $t_0 \geq 0$. e.g., $a(t)$ is periodic with period T and that $\int_0^T a(t)\mathrm{d}t \leq 0$.

Define a new state called $y = (y_1, \cdots, y_n)^\top$ through

$$y_i(t) = x_i(t) - \int_0^t e^{a(\tau)(t-\tau)} b(\tau)\mathrm{d}\tau.$$

We can readily check that

$$\dot{y}_i = a(t)y_i + \sum_{i=1}^n a_{ij}(t, y)(y_j - y_i).$$

Obviously, synchronization of y_i is equivalent to that of x_i. Therefore, without loss of generality, we assume $b(\cdot) \equiv 0$ in (4). The system (3) can now be written in matrix form as

$$\dot{x} = A(t, x)x \tag{6}$$

where $A_{ij}(t, x) = a_{ij}(t, x)$ for $i \neq j$ and $A_{ii}(t, x) = a(t) - \sum_{j \neq i}^n a_{ij}(t, x)$. The matrix $A(t, x)$ has an eigenvector $\mathbb{1}_n$, indeed, $A(t, x)\mathbb{1} = a(t)\mathbb{1}$.

Assumption A1 (Regularity): For all i, j, the mappings $x \mapsto a_{ij}(t, x)$ are

1) locally Lipschitz continuous, uniformly in t, i.e., for any compact set D, there exists a positive constant L_D, such that

$$|a_{ij}(t, x) - a_{ij}(t, y)| \leq L_D |x - y|, \quad \forall x, y \in D, \forall t \geq 0.$$

2) locally bounded in the sense that for any compact set D, there exists a constant $C_D > 0$, such that

$$|a_{ij}(t, x)| \leq C_D, \quad \forall x, y \in D, \forall t \geq 0,$$

3) continuous, uniformly in t, i.e., for any x and $\epsilon > 0$, there exists $\delta_x > 0$ such that

$$|a_{ij}(t,x) - a_{ij}(t,y)| < \epsilon, \quad \forall |y-x| < \delta_x, \forall t \geq 0,$$

Assumption A2 (Connectivity): there exists a constant $T > 0$ and a continuous hollow matrix (with zero diagonal elements) $B(x)$, which represents a quasi-strongly connected graph, such that

$$\text{off-diag} \int_{t_0}^{t_0+T} A(t, \phi(t; t_0, x))\mathrm{d}t \geq B(x), \quad \forall t_0 \geq 0, \forall x.$$

Our main result is the following theorem.

Theorem 3. *Consider the multi-agent system (6) under regularity assumption* ***A1*** *and connectivity assumption* ***A2****. The system achieves exponential synchronization on any compact convex set $D \subseteq \mathbb{R}^n$.*

Proof. For the moment assume $t \mapsto A(t,x)$ is piece-wise continuous. Consider the Euler approximation scheme

$$x_{i+1} = x_i + hA(t_0 + ih, x_i)x_i, \quad i \geq 0$$

with $h = \frac{T}{N}$, $x_0 = x$. The following estimate is well known due to the Lipschitz continuity of the mapping $x \mapsto A(t,x)x$ on any compact set D:

$$|x_i - \phi(t_0 + ih, t_0, x)| \leq \frac{Ci}{N^2}, \quad i \in \{0, \cdots, N\}.$$

In particular, $|x_N - \phi(t_0 + T, t_0, x)| \leq \frac{C}{N}$ where C depends only on the set D and the constant T. Choose $\lambda > 0$ large enough such that $\bar{A}(t,x) := A(t,x) + \lambda I \geq 0$ for all $t \geq 0$ and $x \in D$. The following calculation is in order

$$\begin{aligned}
&x_0 = x \\
&x_1 = (1-h\lambda)x + h\bar{A}(t_0,x)x \\
&x_2 = [(1-h\lambda)^2 I + h(1-h\lambda)(\bar{A}(t_0,x) + \bar{A}(t_0+h,x_1)) + *]x \\
&\quad \vdots \\
&x_N = [(1-h\lambda)^N I + (1-h\lambda)^{N-1}\sum_{i=0}^{N-1} h\bar{A}(t_0+ih,x_i) + *]x
\end{aligned} \tag{7}$$

where $*$ stands for non-negative terms. Due to the compactness of $[t_0, t_0+T]$, for any $\epsilon > 0$, there exists $\omega > 0$ such that $|A(t,x) - A(t,y)| < \epsilon, \quad \forall |x-y| < \omega, \forall t \geq 0$. Choose N large enough such that $\frac{C}{N} < \omega$, then

$$|\bar{A}(t_0+ih, \phi(t_0+ih,t_0,x)) - \bar{A}(t_0+ih,x_i)| < \epsilon$$

from which it follows that

$$\begin{aligned}
&\left|\sum_{i=0}^{N-1} h\bar{A}(t_0+ih,x_i) - \int_{t_0}^{t_0+T} \bar{A}(t,\phi(t,t_0,x))\mathrm{d}t\right| \\
\leq &\left|\sum_{i=0}^{N-1} h\bar{A}(t_0+ih,x_i) - \sum_{i=0}^{N-1} h\bar{A}(t_0+ih,\phi(t_0+ih,t_0,x))\right| \\
+ &\left|\sum_{i=0}^{N-1} h\bar{A}(t_0+ih,\phi(t_0+ih,t_0,x)) - \int_{t_0}^{t_0+T} \bar{A}(t,\phi(t,t_0,x))\mathrm{d}t\right| \\
\leq &\epsilon + \left|\sum_{i=0}^{N-1} h\bar{A}(t_0+ih,\phi(t_0+ih,t_0,x)) - \int_{t_0}^{t_0+T} \bar{A}(t,\phi(t,t_0,x))\mathrm{d}t\right|
\end{aligned} \tag{8}$$

Since ϵ is arbitrary, we get

$$\lim_{N\to\infty} \sum_{i=0}^{N-1} h\bar{A}(t_0+ih,x_i) = \int_{t_0}^{t_0+T} \bar{A}(t,\phi(t,t_0,x))dt$$

because the Riemann sum of the piece-wise continuous function $t \mapsto \bar{A}(t,\phi(t;t_0,x))$ converges to the integral. Now let $P(t_0+T,t_0,x)$ be the matrix in the bracket on the right hand side of (7), then

$$\phi(t_0+T,t_0,x) = P(t_0+T,t_0,x)x$$

for some matrix $P(t_0+T,t_0,x)x$. The matrix $P(t_0+T,t_0,x)$ has $\mathbb{1}_n$ as an eigenvector. Indeed, if we write the Euler scheme as $x_{i+1} = P_i x_i$, then obviously $P_i\mathbb{1} = (1 + a(t_0 + ih))\mathbb{1}$. Consequently,

$$\begin{aligned}
P(t_0+T,t_0,x)\mathbb{1} &= \lim_{N\to\infty} P_N \cdots P_1 \mathbb{1} \\
&= \lim_{N\to\infty} \prod_{i=1}^{N}(1 + a(t_0+ih))\mathbb{1} = \left(\exp\int_{t_0}^{t_0+T} a(t)\mathrm{d}t\right)\mathbb{1}.
\end{aligned}$$

To summarize, $P(t_0+T,t_0,x)\mathbb{1}_n = \varphi(t_0)\mathbb{1}_n$ for some $\varphi(t_0) = \exp\int_{t_0}^{t_0+T} a(t)\mathrm{d}t \leq 1$. [3]

On the other hand, let $N \to \infty$, from (7) we know that

$$P(t_0+T,t_0,x) \geq e^{-\lambda}I + e^{-\lambda}\int_{t_0}^{t_0+T} \bar{A}(t,\phi(t,t_0,x))\mathrm{d}t.$$

Assume w.l.o.g. that $B(\cdot)$ is δ-connected, i.e., there exists k such that $B_{ik}(x) > \delta$ for all i and x. (Otherwise we consider $\phi(t_0 + nT, t_0, x)$). Then since off-diag$(\int_{t_0}^{t_0+T} \bar{A}(t,\phi(t,t_0,x))) \geq B(x)$, there exists some positive constant, still denoting as δ, such that $P(t_0 + T, t_0, x)_{ik} > \delta$ for all $i = 1, \cdots, n$ (notice that P_{kk} is lower bounded by $e^{-\lambda}$) and that δ does not depend on t_0. Assume w.l.o.g., $D \subseteq \mathcal{K}(\epsilon_0)$, [4] where ϵ_0 is arbitrary, $x \in D$ can be written as $x = \frac{\mathbb{1}}{\sqrt{n}} - e$ where $e = (e_1, \cdots, e_n)^\top$, with $e_i \in [0, \epsilon_0]$, $\epsilon_0 < \frac{1}{\sqrt{n}}$, for all i. Then $\phi(t_0 + T, t_0, x) = \frac{\varphi(t_0)}{\sqrt{n}}\mathbb{1} - P(t_0 + T, t_0, x)e$. Let $q_i(x)$ be the i-th component of $P(t_0 + T, t_0, x)e$, then obviously $\delta e_k \leq q_i(x) \leq \delta e_k + (1-\delta)\epsilon_0$. Hence for all $i = 1, \cdots, n$, there holds

$$\frac{\phi(t_0+T,t_0,x)_i}{|\phi(t_0+T,t_0,x)|} \geq \frac{1 - \sqrt{n}(\varphi(t_0) - \delta)\epsilon_0 - \sqrt{n}\delta e_k}{\sqrt{n} - n\delta e_k} =: \theta(e_k) \tag{9}$$

One can easily show that θ is monotone decreasing and thus

$$\begin{aligned}
\frac{\phi(t_0+T,t_0,x)_i}{|\phi(t_0+T,t_0,x)|} &\geq \theta(\epsilon_0) = \frac{1}{1-\sqrt{n}\delta\epsilon_0}\left(\frac{1}{\sqrt{n}} - \varphi(t_0)\epsilon_0\right) \\
&\geq \frac{1}{1-\sqrt{n}\delta\epsilon_0}\left(\frac{1}{\sqrt{n}} - \epsilon_0\right) \\
&:= \frac{1}{\sqrt{n}} - C_0\epsilon_0, \quad \forall x \in D
\end{aligned}$$

where $C_0 = \frac{1-\delta}{1-\sqrt{n}\delta\epsilon_0} \in (0,1)$ which depends only on ϵ_0 and δ and hence only on D. This implies that $\phi(t_0+T,t_0,D) \subseteq \mathcal{K}(C_0\epsilon_0)$, as shown in Figure 3.

[3] At this point, we see why the relaxation in Remark 2 is sufficient. Indeed, it is the quantity $\int_{t_0}^{t_0+T} a(t)dt$ that is critical.

[4] For D not in the positive orthant, define a coordinate change $y = x + \gamma\mathbb{1}_n$. Then D can be mapped into the interior of the positive orthant for large γ, and the system in the new coordinate reads $\dot{y} = A(t, y + \gamma\mathbb{1}_n)y$.

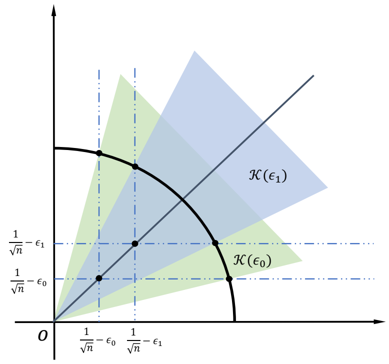

Fig. 3. Cone contraction

Similarly,

$$\frac{\phi(t_0+2T,t_0,x)_i}{|\phi(t_0+2T,t_0,x)|}=\frac{\phi(t_0+2T,t_0+T,\phi(t_0+T,t_0,x))_i}{|\phi(t_0+2T,t_0+T,\phi(t_0+T,t_0,x))|}$$
$$\geq\frac{1}{\sqrt{n}}-C_1(C_0\epsilon_0),\quad\forall x\in D$$

where $C_1=\frac{1-\delta}{1-\sqrt{n}\delta C_0\epsilon_0}<C_0$, since $P(t_0+T,t_0,x)\in\mathcal{K}(C_0\epsilon_0)$ and δ is independent of t_0. Repeating the above procedure yields

$$\frac{\phi(t_0+mT,t_0,x)_i}{|\phi(t_0+mT,t_0,x)|}\geq\frac{1}{\sqrt{n}}-C_0^{m+1}\epsilon_0,\quad\forall x\in D$$

Invoking Lemma 1, we conclude that the system synchronizes exponentially on D in Hilbert metric.

To remove the assumption on the piece-wise continuity of $t\mapsto A(t,x)$, it suffices to replace Riemann integration by Lebesgue integration. View $t\mapsto A(t,\cdot)$ as a mapping from $\mathbb{R}$ to $C(D;\mathbb{R}^{n\times n})$ equipped with norm $||g||_D=\sup_{x\in D}||g(x)||$. Since $\int_{t_0}^{t_0+T}||A(t,\cdot)||_D dt\leq TC_D$, the function $t\mapsto A(t,\cdot)$ is summable and hence can be approximated by simple functions. Let $\eta>0$ be an arbitrarily small constant, let $\mathscr{A}(t,x)=\sum_{i=0}^{N-1}1_{[t_i,t_{i+1})}(t)A_i(x)$ be the simple function such that

$$\int_{t_0}^{t_0+T}\sup_{x\in D}||\mathscr{A}(t,x)-A(t,x)||\mathrm{d}t<\eta.$$

Assume w.l.o.g. that the partition is uniform, i.e., $t_{i+1}-t_i=\frac{T}{N}$ for all i. Let $\bar{A}_i(x)=A_i(x)+\lambda I\geq 0$. As in previous section, the Euler approximation scheme gives

$$x_N=[(1-h\lambda)^N I+(1-h\lambda)^{N-1}\sum_{i=0}^{N-1}h\bar{A}(x_i)+*]x.$$

Now

$$\sum_{i=0}^{N-1}h\bar{A}(x_i)=\int_{t_0}^{t_0+T}(\mathscr{A}(t,\phi(t,t_0,x))+\lambda I)\mathrm{d}t+\epsilon\mathbb{1}_{n\times n}$$
$$=\int_{t_0}^{t_0+T}(A(t,\phi(t,t_0,x))+\lambda I)\mathrm{d}t$$
$$+\int_{t_0}^{t_0+T}(\mathscr{A}(t,\phi(t,t_0,x))-A(t,\phi(t,t_0,x)))\mathrm{d}t+\epsilon\mathbb{1}_{n\times n}$$
$$>\int_{t_0}^{t_0+T}(A(t,\phi(t,t_0,x))+\lambda I)\mathrm{d}t-(\epsilon+\eta)\mathbb{1}_{n\times n}$$

where the constant $\epsilon>0$ represents an infinitesimal error caused by the error between $\phi(t+ih,t_0,x)$ and x_i (recall A4). Let $P_N(x)=[(1-h\lambda)^N I+(1-h\lambda)^{N-1}\sum_{i=0}^{N-1}h\bar{A}(x_i)+*]$. Since $(1-h\lambda)^{N-1}$ is bounded away from 0 for all N, we can choose η sufficiently small such that $(P_N(x))_{ik}>\delta>0$ for all i. The rest of the proof is the same as in the piece-wise continuous case. Thus the theorem holds. □

III. Discussions

A. Comparison result

Corollary 1. *Consider the system (3) and assume **A1**. If there exists a QSC graph $B(x)$ such $A(t,x)\geq B(x)$ for all t, then the system synchronizes exponentially on any compact set.*

Proof. Obvious since $\int_{t_0}^{t_0+T}A(t,x)dt>TB(x)$, $\forall T$. Note that it is not clear how to obtain this result from [3]. □

B. Switching system

Now going back to the system (1), under assumption **A0**, the system can be written as

$$\dot{x}=\left[\sum_{k=1}^{p}1_k(\sigma(t))A_k(x)\right]x$$

Fix an interval $[t_0,t_0+T]$, then on $[t_0-\tau_D,t_0+T+\tau_D]$, we integrate

$$\int_{t_0-\tau_D}^{t_0+T+\tau_D}\sum_{k=1}^{p}1_k(\sigma(t))A_k(x)\mathrm{d}t\geq\tau_D\sum_{k=1}^{p}A_k(x).$$

By assumption, $\sum_{k=1}^{p}A_k(x)$ is QSC and hence the necessary part of Theorem 1 follows.

Example 2. *Consider the Kuramoto model with identical frequency $\dot{x}_i=k\sum_{j\in\mathcal{N}_i(t)}\sin(x_j-x_i)$, $i=1,\cdots n$. where $\mathcal{N}_i(t)$ is obtained by piece-wise constant switching with positive dwell time such that the graph is QSC. Let a,b be real numbers such that $0\leq b-a<\pi$, and define $\mathcal{S}=[a,b]$. Rewrite the system as*

$$\dot{x}_i=k\sum_{j\in\mathcal{N}_i(t)}\frac{\sin(x_j-x_i)}{x_j-x_i}(x_j-x_i)$$

*then obviously assumptions **A1**-**A2** are met. Thus* exponential *synchronization is achieved on compact set. This extends [3], where only asymptotic synchronization is obtained. However, this result can be proven using other methods, see e.g. [17].*

C. No continuity assumption

We relax the assumption on continuity of the functions $x\mapsto a_{ij}(t,x)$. x_N computed using Euler scheme still converges to $\phi(t_0+T,t_0,x)$ since $a_{ij}(t,x)$ are locally Lipschitz. But instead of having a uniform lower bound on $P(t_0+T,t_0,x)_{ik}$, we only know that $P(t_0+T,t_0,x)_{ik}>0$, $\forall i=1,\cdots,n$ for all $t_0\in\mathbb{R}$ and the fixed $T>0$. Thus $P(t_0+T,t_0,\cdot)$ contracts every cone $\mathcal{K}(\epsilon)$. Let $D=\mathrm{co}\{x_1,\cdots,x_n\}\subseteq\mathcal{K}(\epsilon)$, by **A2**, there exists $\epsilon_1<\epsilon$ such that

$$\min_{x\in D}\frac{\phi(t_0+T,t_0,x)_i}{|\phi(t_0+T,t_0,x)|}=\frac{1}{\sqrt{n}}-\epsilon_1$$

(the minimum is attained since ϕ is continuous). By induction, we can define a strictly decreasing sequence $\{\epsilon_m\}_{m=1}^{\infty}$, for example,

$$\begin{aligned}\frac{1}{\sqrt{n}} - \epsilon_2 &= \min_{x\in D} \frac{\phi(t_0+2T,t_0,x)_i}{|\phi(t_0+2T,t_0,x)|} \\ &= \min_{x\in\phi(t_0+T,t_0,D)} \frac{\phi(t_0+2T,t_0+T,x)_i}{|\phi(t_0+2T,t_0+T,x)|}\end{aligned}$$

with $\epsilon_2 < \epsilon_1$ since $\phi(t_0+T,t_0,D) \subseteq \mathcal{K}(\epsilon_1)$. Then $\epsilon_m \to c \geq 0$ as $m \to \infty$. Suppose $c > 0$, and choose ϵ_m

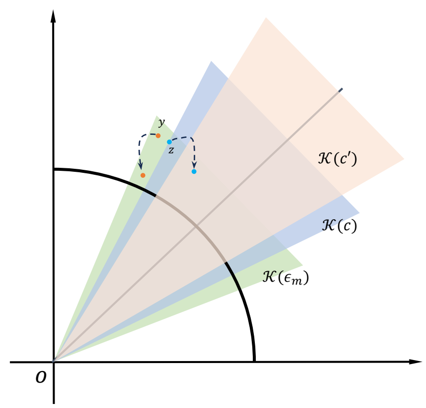

Fig. 4. Asymptotic contraction.

sufficiently close to c. There must exist $y \in \mathcal{K}(\epsilon_m)\backslash\mathcal{K}(c)$, such that $\phi(t_0+(m+1)T,t_0+mT,y) \in \mathcal{K}(\epsilon_m)\backslash\mathcal{K}(c)$. Choose $z \in \mathcal{K}(c)$ close to x, then due to the continuity of $\frac{\phi(t,t_0,x)}{|\phi(t,t_0,x)|}$, the error between $\frac{\phi(t_0+(m+1)T,t_0+mT,y)}{|\phi(t_0+(m+1)T,t_0+mT,y)|}$ and $\frac{\phi(t_0+(m+1)T,t_0+mT,z)}{|\phi(t_0+(m+1)T,t_0+mT,z)|}$ is of order $O(|\epsilon_m - c|)$ which can be made sufficiently small by choosing m large enough. But

$$\min_{w\in\mathcal{K}(c)\cap D} \frac{\phi(t_0+(m+1)T,t_0+mT,w)_i}{|\phi(t_0+(m+1)T,t_0+mT,w)|} = \frac{1}{\sqrt{n}} - c' > \frac{1}{\sqrt{n}} - c$$

implying that $\frac{\phi(t_0+(m+1)T,t_0+mT,z)_i}{|\phi(t_0+(m+1)T,t_0+mT,z)|}$ is away from $\frac{1}{\sqrt{n}} - c$ for any m. This is a contradiction since $\frac{\phi(t_0+(m+1)T,t_0+mT,y)_i}{|\phi(t_0+(m+1)T,t_0+mT,y)|} \in \mathcal{K}(\epsilon_m)\backslash\mathcal{K}(c)$. Thus we conclude that $\epsilon_m \to 0$ as $m \to \infty$ and asymptotic synchronization is guaranteed. This recovers the result in [3].

D. Higher order agent dynamics

The result can be easily extended to non scalar agent dynamics as in [3]. There, $A(t,x)$ has structure $A(t,x) = \tilde{A}(t,x) \otimes I_m$ where $\tilde{A}(t,x) \in \mathbb{R}^{n\times n}$. From previous calculation, we see that $\phi_A(t,t_0,x)$ can be decomposed as $P(t,t_0,x) \otimes I_m$ where $P(t,t_0,x)$ converges to some one rank matrix $\mathbb{1}_n\varphi(x)^\top$, where $\varphi(x) \in \mathbb{R}^n$. Thus the system states converge toward $\mathbb{1} \otimes \alpha$ exponentially for some $\alpha \in \mathbb{R}^m$. More generally, one can even consider consensus on manifolds. Let (M,g) be a Riemannian manifold. Denote $\log_x : M \to T_xM$ as the inverse mapping of the exponential map $\exp_x : T_xM \to M$ and consider the following multi-agent system:

$$\dot{x}_i = f(x_i) + \sum_{j=1}^{n} a_{ij}(t,x)\log_{x_i} x_j, \quad i = 1,\cdots,n$$

where $a_{ij}(t,x) : \mathbb{R}_+ \times M \to \mathbb{R}_+$ are non-negative continuous functions on M. Although the agent is no longer one dimensional scalar dynamics, its consensus properties is exactly the same as those in Euclidean space.

IV. Future direction

The current framework is quite flexible. It can be further studied in several different directions: 1) The internal dynamics can be nonlinear. 2) Stochastic noise can be introduced, e.g., randomness in the interaction forces. In that case, probability estimations of consensus can be derived. 3) One can consider more general accumulated graphs rather than QSC ones. Then we will be able to analyze more general phenomena than consensus, e.g., multi-synchronization.

References

[1] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Transactions on automatic control*, vol. 50, no. 2, pp. 169–182, 2005.

[2] Z. Lin, B. Francis, and M. Maggiore, "On the state agreement problem for multiple nonlinear dynamical systems," in *IFAC Proceedings Volumes*, vol. 38, no. 1. Elsevier, 2005, pp. 82–87.

[3] ——, "State agreement for continuous-time coupled nonlinear systems," *SIAM Journal on Control and Optimization*, vol. 46, no. 1, pp. 288–307, 2007.

[4] K. S. Narendra and A. M. Annaswamy, "Persistent excitation in adaptive systems," *International Journal of Control*, vol. 45, no. 1, pp. 127–160, 1987.

[5] D. Angeli and P.-A. Bliman, "Stability of leaderless discrete-time multi-agent systems," *Mathematics of Control, Signals and Systems*, vol. 18, pp. 293–322, 2006.

[6] Y. Kawano and M. Cao, "Contraction analysis of virtually positive systems," *Systems & Control Letters*, vol. 168, p. 105358, 2022.

[7] L. Moreau, "Stability of continuous-time distributed consensus algorithms," in *2004 43rd IEEE conference on decision and control (CDC)*, vol. 4. IEEE, 2004, pp. 3998–4003.

[8] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on automatic control*, vol. 48, no. 6, pp. 988–1001, 2003.

[9] G. Dirr, H. Ito, A. Rantzer, and B. Rüffer, "Separable Lyapunov functions for monotone systems: Constructions and limitations," *Discrete Contin. Dyn. Syst. Ser. B*, vol. 20, no. 8, pp. 2497–2526, 2015.

[10] A. Rantzer, "Scalable control of positive systems," *European Journal of Control*, vol. 24, pp. 72–80, 2015.

[11] H. R. Feyzmahdavian, B. Besselink, and M. Johansson, "Stability analysis of monotone systems via max-separable Lyapunov functions," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 643–656, 2017.

[12] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE transactions on automatic control*, vol. 31, no. 9, pp. 803–812, 1986.

[13] J. Cortés, "Distributed algorithms for reaching consensus on general functions," *Automatica*, vol. 44, no. 3, pp. 726–737, 2008.

[14] R. Sepulchre, A. Sarlette, and P. Rouchon, "Consensus in non-commutative spaces," in *49th IEEE conference on decision and control (CDC)*. IEEE, 2010, pp. 6596–6601.

[15] G. Birkhoff, "Extensions of Jentzsch's theorem," *Transactions of the American Mathematical Society*, vol. 85, no. 1, pp. 219–227, 1957.

[16] P. J. Bushell, "Hilbert's metric and positive contraction mappings in a banach space," *Archive for Rational Mechanics and Analysis*, vol. 52, pp. 330–338, 1973.

[17] F. Dörfler and F. Bullo, "Synchronization in complex networks of phase oscillators: A survey," *Automatica*, vol. 50, no. 6, pp. 1539–1564, 2014.