2015 European Control Conference (ECC)
July 15-17, 2015. Linz, Austria

A Simple Proof of the Continuous Time Linear Consensus Problem with Applications in Non-Linear Flocking Networks

Christoforos Somarakis[1] and John S. Baras[2]

***Abstract*— We revisit the linear distributed consensus problem in continuous time, to provide a simple and elegant proof under very mild assumptions. Our approach is based on a novel extension of the contraction coefficient that can be adapted to the continuous time version of the model. We apply our results in non-linear second order consensus networks of Cucker-Smale type and we obtain new initial conditions for asymptotic flocking.**

I. INTRODUCTION

Self-organized dynamics lie in the core of modern complex dynamics, a most interesting branch of which is this of networked control systems. Examples of networks that illustrate a collective behavior, as a result of local interaction among nodes in the network, are ubiquitous both in nature and in human societies. The self-organized aspect of these systems is usually understood by a distributed, local exchange of information between autonomous agents who seek some form of co-operation. The core phenomenon in these examples is that through this dynamic interaction all agents' states eventually concentrate around a common value. These problems are known as consensus problems and enjoy a durable interdisciplinary interest in the applied sciences. As a result several mathematical models have been introduced to appraise this *emergence of consensus* among agents.

In its simplest version, a formal framework includes a finite number of agents N, each agent i of which possesses a value of interest denoted by the real number x_i and evolves it under the following averaging scheme:

$$\dot{x}_i = \sum_j a_{ij}(x_j - x_i), \qquad i = 1, \ldots, N. \tag{0.1}$$

The coupling weights a_{ij} are non-negative numbers that quantify the influence of agent j on i. They essentially characterize the interdependence of agents, the connectivity regime and the process of the asymptotic alignment with respect to x_i. The majority of the proposed frameworks, many of which are discussed below, is in fact concerned with different versions of the connectivity weights a_{ij}.

A. *Review of the existing literature*

Consensus algorithms have been studied under deterministic settings of linear or non-linear versions, static or switching communication topologies, time or state dependent connectivity weights: [18], [12], [6], [11], [8], [13], [7], [2], [14]. An alternative perspective is this of stochastic settings either in the communication regime or as a noisy perturbation to the nominal model. This setting is beyond the scope of this paper.

In their vast majority, the aforementioned works rely on a common and fundamental assumption: The exchange of information among any two communicating nodes occurs via an established connection with a communication weight that is uniformly bounded away from zero. This automatically ensures the applicability of an abundance of results from linear algebra, algebraic graph theory, probability theory [15], [1], [10], [4] towards proving asymptotic consensus. The importance of this underlying assumption is noted before [11] and we strenuously suggest that whichever work does not explicitly state it, should be subject to criticism. In [16] the authors work on the discrete time version of Eq. (0.1) arguing as well, that this assumption is instrumental for proving uniform convergence and that it can be relaxed with considerable care.

Distributed consensus systems that bear non-uniform positive weights have appeared in the literature [7], [2] and it is this condition that makes the corresponding stability problems particularly challenging.

This paper is an excerpt from a technical report which discusses this new approach to the continuous time consensus model for a set of important variations in the literature [17].

B. *Organization of the paper*

In section II, we state the main nomenclature to be used, and we review elements of graph and non-negative matrix theories. Additionally, we provide vital preliminary results by extending parts of the theory of non-negative matrices.

In Section III, we consider the deterministic linear consensus problem and provide a new and simple proof using the main mathematical tool implemented for consensus systems in discrete time, after turning the problem from a differential equation into an integral equation one. One of the advantages is that just like in [16] we drop the uniform lower bound on the connectivity weights and impose new conditions for consensus on the rate that the non-zero connectivity weights are allowed to vanish. These conditions heavily depend on the type of the connectivity regime.

In Section IV we apply the obtained results to an important flocking model with state dependent connectivity weights and obtain new conditions for flocking under the mildest connectivity regime within the deterministic framework.

*This work is supported by US AFOSR MURI grant FA9550-09-1-0538, NSF grant CNS-1035655, NIST grant 70NANB11H148.

[1]C. Somarakis is with the Applied Mathematics Dept. , the Institute for Systems Research, University of Maryland, College Park, MD, USA csomarak@umd.edu

[2]J. S. Baras is with the Department of Electrical Engineering, the Institute for Systems Research, University of Maryland, College Park, MD, USA baras@umd.edu

In the discussion section, V, we comment on our work and in particular we mention a number of models to which our framework is applicable.

Due to space limitations the proofs of a few technical lemmas and propositions are omitted. For complete results the reader is referred to [17].

II. Notations and Definitions

Typically, $\mathbb{N}, \mathbb{Z}, , \mathbb{R}$ denote the sets of naturals, integers and reals respectively. For $N \in \mathbb{N}$, $\mathcal{V} = \{1, \ldots, N\}$. For any set $\mathcal{B}$ we understand $\mathcal{B}^C$ as its complement. We will work in the N-dimensional Euclidean space $\mathbb{R}^N$ a vector $\mathbf{x} \in \mathbb{R}^N$ of which is considered as a column vector, unless otherwise stated. The *agreement* or *consensus* space Δ is defined as the subset of $\mathbb{R}^N$ such as

$$\Delta = \{\mathbf{x} \in \mathbb{R}^N : x_1 = x_2 = \cdots = x_N\}$$

Next, we define the *spread* of a vector $\mathbf{x} \in \mathbb{R}^N$ as

$$S(\mathbf{x}) = \max_{i,j} \ x_i - x_j.$$

This quantity will serve as a pseudo-norm for the stability analysis to follow. Indeed it is always non-negative, satisfies the triangular inequality and $S(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in \Delta$. By $\mathbf{1}$ we understand the $N-$dimensional vector with all entries equal to 1 and obviously $S(c\mathbf{1}) = 0$ for any $c \in \mathbb{R}$. By I we understand the $N \times N$ identity matrix. Finally, be $||\cdot||_p$ we denote the p-norm.

A. Graph theory

By a *topological directed graph* $\mathbb{G}$ we understand the pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V} =$ is the set of vertices, $\mathcal{E} = \{(i,j) : i,j \in \mathcal{V}\}$ is the set of edges where $(i,j) \neq (j,i)$. The *degree* N_i of a vertex $i \in \mathcal{V}$ is defined as the cardinality of the set $\{j \in \mathcal{V}, (i,j) \in \mathcal{E}\}$. The graph $\mathbb{G}$ is *routed-out branching* if there exists a vertex $i \in \mathcal{V}$ (called the route of the graph) such that for any $j \neq i \in \mathcal{V}$ there is a *path* of edges $(l_k, l_{k-1})|_{k=0}^m$ such that $l_0 = i$ and $l_m = j$. For two graphs $\mathbb{G}_1 = (\mathcal{V}, \mathcal{E}_1)$ and $\mathbb{G}_2 = (\mathcal{V}, \mathcal{E}_2)$, we say that $\mathbb{G}_1$ is a sub-graph of $\mathbb{G}_2$ if $\mathcal{E}_1 \subset \mathcal{E}_2$. The *adjacency matrix* A is a $0-1$, $N \times N$ matrix with elements $A_{ij} = 1 \Leftrightarrow (i,j) \in E$. The *degree matrix* $D := \text{Diag}\big[d_i\big]$. Finally, the *Laplacian* of $\mathbb{G}$ is the matrix $L := D - A$ with the sum of its rows being identically equal to zero.

By $\mathcal{S}$ we denote the family of graphs with fixed N vertices and self-edges on every node, and by $\mathcal{T} \subset S$ the subset of graphs each of which is routed-out branching.

B. Non-negative Matrix theory

A *non-negative* matrix $P = \{p_{ij}\}$ is such that $p_{ij} \geq 0$ for all i, j.[1] The non-negative matrix P is *generalized stochastic*, or m-stochastic, if $\sum_j p_{ij} = m$ for all i. A crucial property of an m-stochastic matrix is that m is always an eigenvalue of it. For $m = 1$ P reduces to the standard stochastic matrix. Given an m-stochastic matrix $P = [p_{ij}]$, the quantity

$$\tau(P) = \frac{1}{2} \max_{i,j} \sum_s |p_{is} - p_{js}| = m - \min_{i,j} \sum_s \min\{p_{is}, p_{js}\} \quad (1)$$

is *the coefficient of ergodicity* of P. τ measures the averaging effect of stochastic matrices. Its history dates back to one of Markov's first papers [9] and in the literature there exist an abundance of similar tools [5]. A crucial set of properties of τ is presented in the following Theorem:

Theorem 2.1: [4] For any m-stochastic matrix P and $\mathbf{z} \in \mathbb{R}^N$ it holds that:

$$S(P\mathbf{z}) \leq \tau(P) S(\mathbf{z}).$$

The coefficient of ergodicity measures the averaging effect of stochastic matrices and it is the central concept behind any convergence result in linear consensus algorithms. Its history dates back to one of Markov's first papers [4] and it applies to "dynamics" of the type:

$$\mathbf{w} = P\mathbf{z}$$

with P being m-stochastic.

C. Preliminary Results

A not so straightforward extension is illustrated in Theorem 2.4 below, the proof of which relies on the following two lemmas:

Lemma 2.2 (The first mean value theorem for integration): If a real-valued function G is continuous in a compact subset J and ϕ is integrable that does not change sign on J, then there exists $x \in J$ such that

$$G(x) \int_J \phi(t) dt = \int_J G(t) \phi(t) dt.$$

Lemma 2.3 (Lemma 1.1 of [4]): Suppose $\boldsymbol{\delta} \in \mathbb{R}^N$ such that $\boldsymbol{\delta}^T \mathbf{1} = 0$ and $\boldsymbol{\delta} \neq 0$. Then there is an index $\mathcal{I} = \mathcal{I}(\delta)$ of ordered pairs (i,j) with $i, j \in \mathcal{V}$ such that

$$\boldsymbol{\delta}^T = \sum_{(i,j) \in \mathcal{I}} \frac{T_{ij}}{2} (\mathbf{e}_i - \mathbf{e}_j)$$

where $T_{ij} > 0$ and $\mathbf{e}_i$ is the row vector $(0, \ldots, 0, 1, 0, \ldots, 0)$ with 1 at the i^{th} position.

Theorem 2.4: Let I be a compact subset of $\mathbb{R}$ and assume that for any compact $I' \subset I$, $W_{I'} = \int_{s \in I'} P(s) ds$ is m-stochastic. If $\mathbf{w} = \int_{s \in I} P(s) \mathbf{z}(s) ds$ then

$$S(\mathbf{w}) = \tau(W_I) S(\mathbf{z}^*)$$

for some $\mathbf{z}^* = (z_1(s_1), \ldots z_N(s_N))$ for some $s_i \in I$ and

$$\begin{aligned}\tau(W_I) &= \frac{1}{2} \max_{h,h'} \sum_{k=1}^N \int_{s \in I} |p_{hk}(s) - p_{h'k}(s)| ds \\ &= m - \min_{h,h'} \sum_{k=1}^N \min\left\{ \int_{s \in I} p_{hk}(s) ds, \int_{s \in I} p_{h'k}(s) ds \right\}\end{aligned}$$

Proof:

Pick $h, h^{'} \in \mathcal{V}$. Then for $\mathbf{p}_h, \mathbf{p}_{h'}$ the h^{th} and $h^{'th}$ rows of P respectively, we have

$$\int_{s \in I} \big(\mathbf{p}_h(s) - \mathbf{p}_{h'}(s)\big) \mathbf{z}(s) ds$$

[1] Unless otherwise specified each matrix is supposed to be square and of dimension $N \times N$.

Now, since $N < \infty$ there is a partition $\{I_l\}_{l=1}^m$ of I which depends on h, h' such that for any I_l, $p_{hk}(s) - p_{h'k}(s)$ does not change sign for $s \in I_l$, $k \in \mathcal{V}$ and it is not identically zero. Then for fixed I_l we apply Lemma 2.2 to obtain

$$\sum_k \int_{s\in I_l} \big(p_{hk}(s) - p_{h'k}(s)\big) z_k(s) ds =$$
$$\sum_k \int_{s\in I_l} \big(p_{hk}(s) - p_{h'k}(s)\big) ds z_k(s_k^*) = \boldsymbol{\delta}_l^T \mathbf{z}_l^*$$

for some $s_k^* = s(I_l, h, h')$, $\boldsymbol{\delta}_l^T = \int_{I_l} \big(\mathbf{p}_h(s) - \mathbf{p}'_h(s)\big) ds \neq 0$ and $\mathbf{z}_l^* = (z_1(s_1^*), \dots z_N(s_N^*))^T$. By Assumption $\int_{I_l} P(s) ds$ is m-stochastic and therefore $\boldsymbol{\delta}_l^T \mathbf{1} = 0$. Hence Lemma 2.3 is applied and together with the triangle inequality

$$|\boldsymbol{\delta}_l^T \mathbf{z}_l^*| \leq \frac{1}{2} ||\boldsymbol{\delta}_l||_1 S(\mathbf{z}_l^*)$$

(see also [4]). Then if we let $S(\mathbf{z}*) = \max_l S(\mathbf{z}_l^*)$, we obtain the bound

$$S(\mathbf{w}) = \max_{h,h'} \left| \int_{s\in I} \big(\mathbf{p}_h(s) - \mathbf{p}_{h'}(s)\big) \mathbf{z}(s) ds \right|$$
$$= \sum_l |\boldsymbol{\delta}_l^T \mathbf{z}_l^*| \leq \max_{h,h'} \frac{1}{2} \int_I ||\mathbf{p}_h(s) - \mathbf{p}_{h'}(s)||_1 ds S(\mathbf{z}^*)$$

Finally, from the identity $|x - y| = x + y - 2\min\{x, y\}$ for any $x, y \in \mathbb{R}$ and the fact that $\forall h, h' \in \mathcal{V}$ it holds $\sum_k \int_{s\in I} p_{hk}(s) ds = \sum_k \int_{s\in I} p_{h'k}(s) ds = m$ we obtain the desired result.

■

Similarly, for the expression

$$\mathbf{w} = \int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q) \mathbf{z}(q) dq ds$$

one can show, along the lines of the proof of Theorem 2.4 that if $W_I^{(2)} = \int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q) dq ds$ is stochastic, then

$$S(\mathbf{w}) \leq \tau(W_I^{(2)}) S(\mathbf{z}^*) \tag{2}$$

for some $\mathbf{z}^* = \big(z_1(s_{(ij)}^{(1)}), z_2(s_{(ij)}^{(2)}), \dots, z_N(s_{(ij)}^{(N)})\big)$ all $s_{(ij)}^{(l)}$ of which are in $I_1 \cup I_2$.

Finally, the sub-multiplicativity property for pairs of stochastic matrices of the particular form discussed in this section, applies to expressions of the type

$$\mathbf{w} = \int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q) \mathbf{z}(q) dq ds$$

so long as $\int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q) dq ds$ is stochastic.

Regardless if we are working with products of matrices within integrals or not, a crucial point in this work is to ask for p_{ij} such that $\rho < m$. It is this feature that characterizes the contractive (averaging) nature of the stochastic matrices. It can be easily verified that $\min_{i,j} \sum_s \min\{p_{is}, p_{js}\}$ (or equivalently $\min_{h,h'} \sum_{k=1}^N \min\big\{\int_{s\in I} p_{hk}(s) ds, \int_{s\in I} p_{h'k}(s) ds\big\}$) is strictly positive for any P which possesses a strictly positive column. These matrices are called *scrambling* and lie in the core of the analysis of non-homogeneous discrete Markov Chains [4].

The properties of stochastic matrices and their products play a crucial role and the standard approach is through graph theory: Any non-negative matrix P can be represented as a graph $\mathbb{G}_P$ with it's adjacency matrix A_P the elements of which satisfy the property $A_{ij} = 1 \Leftrightarrow P_{ij} \neq 0$. For two stochastic matrices P_1 and P_2, we write $P_1 \sim P_2$ if $\mathbb{G}_{P_1} = \mathbb{G}_{P_2}$ (consequently $P_1 = P_2$). This way we can study P from the point of view of graph theory. A non-negative matrix P is called *regular* if $\mathbb{G}_P$ is routed-out branching.

A classical result in the theory of products of stochastic matrices is that for a regular matrix P there is a power of it that makes it scrambling: i.e. $\exists\gamma \geq 1 : \tau(P^\gamma) < 1$ and from the sub-multiplicative property $P^t \to \mathbf{1}\mathbf{1}^T c$ for some $c \in \mathbb{R}$, as $t \to \infty$. The power of P that makes it scrambling is known as the *scrambling index* and the aforementioned statement on the asymptotic behavior of P^t is the ergodic theorem of stochastic matrices [4]. As the product of stochastic matrices is stochastic as well, the preceding notions can be extended to study the behavior of non-homogeneous products of stochastic matrices. We exclusively study *backward products* of stochastic matrices defined as

$$P_B^{(l)}(t) := \begin{cases} \int_{t-B}^t C(t, s_1) ds_1, \ l = 1 \\ \int_{t-B}^t C(t, s) P_B^{(l-1)}(s) ds, \ l > 1 \end{cases} \tag{3}$$

for some integrable matrix function $C(t, s)$, $B > 0$, $t \in \mathbb{R}$ so that $P_B^{(l)}(t)$ is stochastic for every $l \geq 1$.

We recall now the set $\mathcal{S}$ and its subset $\mathcal{T}$. Let $R = R_N$ denote the cardinality of $\mathcal{T}$. Each member $\mathbb{G}_i$ of it, has a scrambling index γ_i. In fact, $\mathcal{T}$ can be partitioned in such mutually disjoint subsets: $\mathcal{T} = \bigsqcup_v \mathcal{Y}_v$ so that for $\mathbb{G}_1 \in \mathcal{Y}_{z_1}$, $\mathbb{G}_2 \in \mathcal{Y}_{z_2}$, $z_1 \neq z_2$ if and only if $\gamma_{z_1} \neq \gamma_{z_2}$. Consequently, we can enumerate

$$1 = \gamma_0 < \gamma_1 < \cdots < \gamma_{\max} \leq \left[\frac{N}{2}\right]$$

For instance, $\mathcal{Y}_0$ is the subclass of routed-out branching graphs, each member $\mathbb{G}_{\mathcal{Y}_0}$ of which has scrambling index, $\gamma_0 = 1$, i.e. there exist i such that $[\mathbb{G}_{\mathcal{Y}_0}]_{ji} \in \mathcal{E}_{\mathbb{G}_{\mathcal{Y}_0}}$. Next we note that for any $\mathbb{G}_1, \mathbb{G}_2 \in \mathcal{T}$ with $\mathbb{G}_2$ being a subgraph of $\mathbb{G}_1$, it holds that $\gamma_1 \leq \gamma_2$ and thus we understand that by adding an edge to any graph, the scrambling index will certainly not increase. In fact, there exists a sufficient number of new edges that will decrease the scrambling index. Fix $j < i$. Then for any $\mathbb{G}_i \in \mathcal{Y}_i$ there exists a positive number $l_{i,j}$ such that the graph $\mathbb{G}_j$ formed out of $\mathbb{G}_i$ with $l_{i,j}$ additional edges will be a member of $\bigcup_{v=0}^j \mathcal{Y}_v$, in which case $\gamma_j \leq \gamma_i - 1$.

Remark 2.5: The minimum number of edges needed to be added on an arbitrary member of $\mathcal{Y}_i$ so that the resulting graph is a member of $\bigcup_{v=0}^{i-1} \mathcal{Y}_v$, is denoted by $l^* := \max_i\{l_{i,i-1}\}$.

III. A NEW PROOF OF CONVERGENCE TO CONSENSUS.

In this section we will presents a (yet another) proof of consensus convergence. For $N < \infty$ number of agents we consider the following initial value problem:

$$i \in \mathcal{V} : \begin{cases} \dot{x}_i(t) & = \sum_{j\in N_i} a_{ij}(t)\big(x_j(t) - x_i(t)\big),\ t \geq t_0 \\ x_i(t_0) & = x_i^0,\ t = t_0 \end{cases} \tag{4}$$

Modeling failing signals and an overall switching connectivity regime must consider the connectivity weights $a_{ij}(t)$ to "jump" from a positive value to zero in a discontinuous fashion.

Assuming $a_{ij}(t)$ to be right continuous, a solution $\mathbf{x}(t, t_0, \mathbf{x}^0)$ of (4) is a continuous function with a right t-derivative which satisfies the differential equation for every $t \geq t_0$.[2] This solution also satisfies any classical integral equation the derivative of which satisfies Eq. (4) in the classical sense.

Assumption 3.1: For the connectivity weights $a_{ij}(t)$, it holds that $a_{ij}(t) \neq 0 \Rightarrow a_{ij}(t) \geq f(t)$, where $f(t)$ is a positive, monotonically non-increasing function such that $f(t) \to 0$.

Assumption 3.2: The connectivity weights a_{ij} are upper bounded, right continuous, non-negative functions of time.

This, although hardly an assumption, together with $N < \infty$ implies that $m := \sup_{t\geq t_0} \max_i \sum_j a_{ij}(t) < \infty$. We recall now the matrix representation of the graph $\mathbb{G}_t$ in terms of the degree matrix $D = D(t)$ and the adjacency matrix $A = A(t)$. Then the matrix $W(t) := mI - D(t) + A(t)$ is m-stochastic. We begin as in the case of discrete time with two elementary but crucial remarks:

Lemma 3.3: [17] Under Assumption 3.2 the solution $\mathbf{x}\big(t, t_0, \mathbf{x}^0\big)$ of Eq. (4) satisfies $x^0_{min} \leq x_i(t) \leq x^0_{\max}$, $\forall\ t \geq t_0, i \in \mathcal{V}$.

Lemma 3.4: [17] If $\mathbf{x}(t, t_0, \mathbf{x}^0)$ is the solution of Eq. (4) such that $S\big(\mathbf{x}(t)\big) \to 0$ as $t \to \infty$ then the forward limit set $\omega(\mathbf{x}^0)$ is a singleton with a point in Δ.

1) Static & Switching Networks I: We begin the first round of results assuming increased connectivity among agents. This means that the overall connectivity regime may be static or switching provided that $P(t)$ is scrambling on the average:

Theorem 3.5: Let Assumption 3.2 hold. If $f(t) := \min_{i,j} \sum_s \min\{a_{is}(t), a_{js}(t)\}$ satisfies:

$$\int^{\infty} f(t)dt = \infty$$

then we have global convergence of the system (4) to a constant value.

Proof: We write Eq. (4) in vector form

$$\dot{\mathbf{x}} = -D(t)\mathbf{x} + A(t)\mathbf{x} = -m\mathbf{x} + \big(mI - D(t) + A(t)\big)\mathbf{x} =$$
$$= -m\mathbf{x} + W(t)\mathbf{x} \Leftrightarrow \frac{d}{dt}(e^{mt}\mathbf{x}) = e^{mt}P(t)\mathbf{x}$$

[2] If one is not willing to accept this premise, a discontinuous $a_{ij}(t)$ on an subset of $[t_0, \infty)$ with Lebesgue measure zero implies a solution that satisfies (4) in almost every t and the same analysis applies.

so that from Theorem 2.1 we obtain the bound

$$S\bigg(\frac{d}{dt}\big(e^{mt}\mathbf{x}(t)\big)\bigg) \leq e^{mt}(m - f(t))S\big(\mathbf{x}(t)\big)$$

then

$$\begin{aligned}\frac{d}{dt}S\big(\mathbf{x}(t)\big) &= -me^{-mt}S\big(e^{mt}\mathbf{x}(t)\big) + e^{-mt}\frac{d}{dt}S\big(e^{mt}\mathbf{x}(t)\big) \\ &\leq -mS\big(\mathbf{x}(t)\big) + e^{-mt}S\bigg(\frac{d}{dt}\big(e^{mt}\mathbf{x}(t)\big)\bigg) \\ &\leq -mS\big(\mathbf{x}(t)\big) + (m - f(t))S\big(\mathbf{x}(t)\big) \\ &\leq -f(t)S\big(\mathbf{x}(t)\big)\end{aligned}$$

which implies

$$S\big(\mathbf{x}(t)\big) \leq e^{-\int_{t_0}^{t} f(s)ds}S\big(\mathbf{x}^0\big)$$

and the result follows in view of the imposed condition on f and Lemma 3.4. ■

Remark 3.6: This is a generalization of the results obtained in [14] concerning continuous time consensus algorithms. In addition, further improved results on non-linear continuous time models are to be obtained in the following. On condition that there is always an agent $i = i(t) \in \mathcal{V}$ that affects every other agent j in the group it then suffices for $\int^{\infty} f(s)ds = \infty$.

2) Static & Switching Networks II: We will escalate the analysis with the study of the dynamics of Eq. (4) under a recurrent connectivity condition. Define for $B \geq 0$, $s \in [t - B, t]$

$$C(t, s) = e^{-mB}\delta(s - (t - B))I + e^{-m(t-s)}W(s) \tag{5}$$

with $\delta(\cdot)$ being the delta function and $W(s) = mI - D(s) + A(s)$, as before.

Proposition 3.7: [17] Let Assumption 3.2 hold and $C(t, s)$ defined in eq. (5). For any $B > 0$, $l \geq 1$, the matrix $P_B^{(l)}(t)$ as defined in (3) is stochastic.

Assumption 3.8: There exist $B > 0$ and $M > t_0$ so that for any $t \geq M$ the graph $\mathbb{G}_{P_B(t)}$ that corresponds to $P_B(t)$ is routed-out branching.

Due to the upper boundedness of $a_{ij}(t)$, the weights are also assumed to satisfy the dwelling time condition [6].

Assumption 3.9: For any $t \geq t_0$ there exists $\epsilon > 0$ independent of t such that $a_{ij}(t) \neq 0 \Rightarrow a_{ij}(s) \geq f(s)$ for $s \in I_\epsilon(t^*) = [t^* - \epsilon, t^* + \epsilon]$ for some $t^* \in \mathbb{R}$ and $t \in I_\epsilon(t^*)$.

Theorem 3.10: Let Assumptions 3.1, 3.2, 3.8 and 3.9 hold. Unconditional asymptotic consensus for the solution of system (4) is achieved under one of the following conditions:

1. $\mathbb{G}_{P(t)}$ is independent of time (static connectivity) and there exists a sequence $t_i \geq M$ with $t_{i+1} - t_i \geq \gamma B$, such that
$$\sum_i f^{\gamma}(t_i) = \infty.$$
2. $\mathbb{G}_{P(t)}$ depends on time (switching connectivity) and there exists a sequence $t_i \geq t_0$ with $t_{i+1} - t_i \geq \sigma B$, such that

$$\sum_i f^{\sigma}(t_i) = \infty.$$

where $\sigma = l^*([N/2]+1)$ and l^* with the meaning of Remark 2.5.

Proof: We begin with the static case. The solution $\mathbf{x}$ of (4) satisfies

$$\dot{\mathbf{x}} = -m\mathbf{x} + \big(mI - D(t) + A(t)\big)\mathbf{x} \Rightarrow$$
$$\frac{d}{dt}\big(e^{mt}\mathbf{x}\big) = e^{mt}\big(mI - D(t) + A(t)\big)\mathbf{x} \Rightarrow$$
$$e^{mt}\mathbf{x}(t) - e^{m(t-B)}\mathbf{x}(t-B) = \int_{t-B}^{t} e^{ms}W(s)\mathbf{x}(s)ds$$
$$\Rightarrow$$

$$\mathbf{x}(t) = \int_{t-B}^{t} C(t,s_1)\mathbf{x}(s_1)ds_1$$
$$= \int_{t-B}^{t}\int_{s_1-B}^{s_1}\cdots\int_{s_{\gamma-1}-B}^{s_{\gamma-1}}\prod_{k=1}^{\gamma} C(s_{k-1},s_k)\mathbf{x}(s_\gamma)ds_\gamma\cdots ds_1$$

with $s_0 = t$. Consequently from Theorem 2.4, Lemma 3.3 and Proposition 3.7 we have

$$S\big(\mathbf{x}(t)\big) \le \tau\big(P_B^{(\gamma)}(t)\big)S\big(\mathbf{x}(t-\gamma B)\big)$$

a condition that illustrates the contractive dynamics exactly because $\tau\big(P_B^{(\gamma)}(t)\big) < 1$ on the assumption of static connectivity. Equivalently, there exists $\gamma \ge 1$ so that $P_B^{\gamma}(t)$ is scrambling, i.e. for some $j^* \in \mathcal{V}$, $[P_B^{\gamma}(t)]_{j^*i} > 0$ for all $i \in \mathcal{V}$. Then

$$[P_B^{(\gamma)}(t)]_{j^*j^*} \ge$$
$$\int_{s_0-B}^{s_0}\int_{s_1-B}^{s_1}\cdots\int_{s_{\gamma-1}-B}^{s_{\gamma-1}}\prod_{k=1}^{\gamma}\Big(e^{-mB}\delta\big(s_k-(s_{k-1}-B)\big)+$$
$$+\,e^{-m(s_{k-1}-s_k)}\big(m-d_i(s_k)\big)\Big)ds_\gamma\cdots ds_1$$
$$> e^{-\gamma mB}$$

with $s_0 = t$ and for $i \ne j^*$ set $(k_{-1} = i, k_0 = l_0, \ldots, k_\gamma = l_{\gamma-1}, k_\gamma = j^*$ to get

$$[P_B^{(\gamma)}(t)]_{j^*i} > \int_{s_0-B}^{s_0}\cdots\int_{s_{\gamma-1}-B}^{s_{\gamma-1}} e^{-m(s_0-s_\gamma)}ds_\gamma\cdots ds_1 f^{\gamma}(t)$$
$$= \frac{(1-e^{-mB})^{\gamma}}{m^{\gamma}}f^{\gamma}(t)$$

For $t' \le M$ large enough so that $f(t) \le \frac{me^{-mB}}{1-e^{-mB}}$ whenever $t \ge t'$ we obtain the estimate:

$$\tau\big(P_B^{\gamma}(t)\big) \le 1 - c_1 f^{\gamma}(t) \tag{6}$$

where $c_1 = \frac{(1-e^{-mB})^{\gamma}}{m^{\gamma}} > 0$. Finally, for the aforementioned sequence $\{t_i\}$, for any $t \ge t'$, there exists i such that $t \in [t_i, t_{i+1}]$. Then

$$S\big(\mathbf{x}(t)\big) \le S\big(\mathbf{x}(t_i)\big) \le \big(1 - c_1 f^{\gamma}(t_i)\big)S\big(\mathbf{x}(t_i-\gamma B)\big)$$
$$\le \big(1 - c_1 f^{\gamma}(t_i)\big)S\big(\mathbf{x}(t_{i-1})\big)$$

For $\varepsilon > 0$, pick i_1 and i_2 large enough so that $t_{i_1} \ge t'$ and $\sum_{j=i_1}^{i_2} f(t_j) \ge c_1^{-1}\log(\frac{\varepsilon}{S(\mathbf{x}^0)})$ and then for $t \ge t_i$

$$S\big(\mathbf{x}(t)\big) \le \prod_{k=i_1}^{i_2}(1-c_1 f^{\gamma}(t_k))S\big(\mathbf{x}^0\big) \le e^{-c_1\sum_{k=i_1}^{i_2} f^{\gamma}(t_k)}S\big(\mathbf{x}^0\big)$$

hence $S\big(\mathbf{x}(t)\big) \le \varepsilon$ and the proof of the first part is concluded.

In the case of switching connectivity we have $P_B^{(\sigma)}(t)$ and we need to show that it will be scrambling for $sigma$ sufficiently large. By Assumption, for any t and $s \in [t-B, t]$, $\gamma_{P_B(t)}, \gamma_{P_B(s)} \ge 1$ are the scrambling indexes of $P_B(t)$ and $P_B(s)$ and $\gamma_{P_B(s)} \ge 1$ for $s \in [t-B, t]$ so that $\gamma_{P_{B(s^*(t))}} = \max_{s\in[t-B,t]}\gamma_{P_B(s)}$ then $P_B^{(2)}(t) = \int_{t-B}^{t} C(t,s)P_B(s)ds \in \mathcal{T}$ as well and for it's scrambling index $\gamma_{P_B^{(2)}(t)}$ it holds that

$$\gamma_{P_B^{(2)}}(t) \le \begin{cases} \max\{\gamma_{P_B(t)}, \gamma_{P_{B(s^*(t))}}\} - 1, & \mathbb{G}_{P_B(t)} \subset \mathbb{G}_{P_B(s^*)} \\ & \text{or } \mathbb{G}_{P_B(s^*)} \subset \mathbb{G}_{P_B(t)} \\ \max\{\gamma_{P_B(t)}, \gamma_{P_{B(s^*(t))}}\}, & \text{o.w.} \end{cases}$$

Now, in the second case, there is not strict decrease of the scrambling index over $[t-2B, t]$. In this case, however, it holds that $\mathcal{E}^C_{P_B(t)} \cap \mathcal{E}_{P_B(s^*)} \ne \emptyset$, i.e. there exists $(i,j) \in \mathcal{V}\times\mathcal{V}$ such that $[\mathbb{G}_{P_B(t)}]_{ij} > 0$ and $[\mathbb{G}_{P_B(s^*)}]_{ij} = 0$ or vice versa. This element, however, will be a member of $\mathcal{E}_{P_B^{(2)}(t)}$ exactly because $P_B(\cdot)$ has strictly positive diagonal elements. From the discussion on the partitioning of $\mathcal{T}$ with respect to the scrambling indexes,

$$\gamma_{P_B^{l^*}(t)} \le \max\{\gamma_{P_B(t)}, \gamma_{P_{B(u^*(t))}}\} - 1$$

$u^* := \max_{s\in[t-(l^*-1)B,t]}\gamma_{P_B(s)}$. Consequently for $\sigma = l^*([N/2]+1)$ the matrix $P_B^{(\sigma)}(t)$ is scrambling and hence from Theorem 2.4, Lemma 3.3 and Proposition 3.7 we have

$$\tau\big(P_B^{(\sigma)}(t)\big) < 1 - c_2 f^{\sigma}(t)$$

where $c_2 = \frac{(1-e^{-m\epsilon})^{\gamma}}{m^{\gamma}} > 0$ for ϵ to have the meaning of Assumption 3.9 and the proof proceeds as in the first case. ■

If we strengthen Assumption 3.1 to $f(t) \ge \delta > 0$ then we have proved exponential rate of convergence with respect to Δ with an explicit rate of convergence. The next example illustrates this point.

Example 3.11: Consider the network consisted of $N = 4$ agents, with coupling defined by the following adjacency matrix

$$A(t) = \begin{bmatrix} 0 & a_{12}(t) & 0 & 0 \\ a_{21}(t) & 0 & a_{23}(t) & 0 \\ 0 & a_{32}(t) & 0 & a_{34}(t) \\ 0 & 0 & a_{43}(t) & 0 \end{bmatrix}$$

where for all $t \ge 0$ it holds that $a_{ij}(t) \ne 0 \Rightarrow 0 < a \le a_{ij}(t) < \frac{1}{2}$ and also

$$\begin{cases} a_{23}(t)=a_{32}(t)=a_{34}(t)=a_{43}(t)=0 \ \& \ a_{12}(t),a_{21}(t)\neq 0, \\ \hfill t\in[3l\epsilon,(3l+1)\epsilon) \\ a_{12}(t)=a_{21}(t)=a_{34}(t)=a_{43}(t)=0 \ \& \ a_{23}(t),a_{32}(t)\neq 0, \\ \hfill t\in[(3l+1)\epsilon,(3l+2)\epsilon) \\ a_{23}(t)=a_{32}(t)=a_{12}(t)=a_{21}(t)=0 \ \& \ a_{34}(t),a_{43}(t)\neq 0, \\ \hfill t\in[(3l+2)\epsilon,(3l+3)\epsilon) \end{cases}$$

for some fixed $\epsilon>0$ and $l\in\mathbb{Z}_+$. Then

$$C(t,s)= \begin{bmatrix} \bar{d}_1(t,s) & e^{-(t-s)}a_{12} & 0 & 0 \\ e^{-(t-s)}a_{21} & \bar{d}_2(t,s) & e^{-(t-s)}a_{23} & 0 \\ 0 & e^{-(t-s)}a_{32} & \bar{d}_3(t,s) & e^{-(t-s)}a_{34} \\ 0 & 0 & e^{-(t-s)}a_{43} & \bar{d}_4(t,s) \end{bmatrix}$$

where $\bar{d}_i(t,s)=e^{-3\epsilon}\delta(s-(t-3\epsilon))+e^{-(t-s)}(1-d_i(s))$ and $a_{ij}=a_{ij}(s)$. Now for any $t\geq 0$, $P_{3\epsilon}(t)=\int_{t-3\epsilon}^{t}[C(t,s)]_{ij}ds$

$$P_{3\epsilon}(t)=\begin{cases} 1-e^{-t}\int_{t-3\epsilon}^{t}e^{s}d_i(s)ds, & i=j \\ e^{-t}\int_{t-3\epsilon}^{t}e^{s}a_{ij}(s)ds, & (i,j)\in\mathcal{E} \\ 0, & o.w. \end{cases}$$

and by construction of the switching signal, it can be easily shown that $P(t)\geq J$ elementwise where

$$J:=\begin{bmatrix} \frac{1}{2}(1+e^{-\epsilon}) & a(1-e^{-\epsilon}) & 0 & 0 \\ a(1-e^{-\epsilon}) & e^{-\epsilon} & a(1-e^{-\epsilon}) & 0 \\ 0 & a(1-e^{-\epsilon}) & e^{-\epsilon} & a(1-e^{-\epsilon}) \\ 0 & 0 & a(1-e^{-\epsilon}) & \frac{1}{2}(1+e^{-\epsilon}) \end{bmatrix};$$

a remark made merely to prove that $P_{3\epsilon}(t)\in\mathcal{T}$ and that $\gamma_{P_{3\epsilon}(t)}=2$. Consequently, $P_{3\epsilon}^{(2)}(t)=\int_{t-3\epsilon}^{t}C(t,s)P_{3\epsilon}(s)ds$ is lower bounded by J^2 with J^2 corresponding to a matrix with at least one positive column (in fact the second and the third are all positive). Then $P_{3\epsilon}^{(2)}(t)$ is scrambling for any $t\geq 0$ and the lower bounded we are interested in is determined from J^2, being $\min^{+}[J^2]_{ij}$. It can be easily seen that for fixed ϵ and a small enough this number is in fact $a^2(1-e^{\epsilon})^2$. Let a attain such a small value. Since for any $t\geq 0$, there exists $l\in\mathbb{Z}_+$ such that $3l\epsilon\leq t\leq(3l+1)\epsilon$, we conclude that

$$\begin{aligned} S(\mathbf{x}(t)) &\leq S(\mathbf{x}(3(l-1)\epsilon))\leq(1-2a^2(1-e^{-\epsilon})^2)^{l-1}S(\mathbf{x}(0)) \\ &=\frac{S(\mathbf{x}(0))}{1-2a^2(1-e^{-\epsilon})^2}e^{-\theta t}S(\mathbf{x}(0)) \end{aligned}$$

where $\theta:=\frac{\ln(1-2\alpha^2(1-e^{-\epsilon})^2)}{2\epsilon}$, as it is dictated by Theorem 3.10 for $\{t_i\}$ any sub-sequence with 3ϵ interval and $f(t)$ to be lower bounded by a.

IV. NON-LINEAR FLOCKING NETWORKS

Also known as second order consensus systems these models are preferred to model the coordinating behavior of birds, as the latter are seen as autonomous individuals. A population of $N<\infty$ birds exchanges information and acts according to the algorithm

$$i\in\mathcal{V}:\begin{cases} \dot{x}_i(t) &= u_i(t) \\ \dot{u}_i(t) &= \sum_{j\in N_i}a_{ij}(t,\mathbf{x})\big(u_j(t)-u_i(t)\big) \end{cases} \tag{7}$$

with given initial position $\mathbf{x}^0=\mathbf{x}(0)$ and speed $\mathbf{u}^0=\mathbf{u}(0)$. In flocking dynamics the connectivity weights a_{ij} are state dependent. The first models assumed a symmetric dependence as a function of the relative distance between i and j, explicitly expressed [2]. In this work, for a solution $(\mathbf{x},\mathbf{u})$ of Eq. (7) we assume

$$a_{ij}(t,\mathbf{x}(t))\neq 0\Rightarrow a_{ij}(t,\mathbf{x}(t))\geq f(S(\mathbf{x}(t))) \tag{8}$$

for some positive non-increasing function f with the property that $f(s)\to 0$ as $s\to\infty$ and $\sup_{t,y}a_{ij}(t,y)<\infty$. Examples of symmetric coupling weights are proposed in [2] $\tilde{a}_{ij}:=\frac{B}{1+|x_i-x_j|^{2\beta}}$ of non-symmetric $a_{ij}=\frac{\psi(|x_i-x_j|)}{\sum_j\psi(|x_i-x_j|)}$ for some symmetric function $\psi(|x_i-x_j|)$ such as $\tilde{a}_{ij}$. It is easy to observe that these functions are not lower bounded away from zero as the distance of the birds x_i-x_j can become arbitrarily large, a hypothesis that fits in our framework. The problem is to derive initial conditions so that the flock remains connected and it aligns its speed to a common one, i.e. the solution $\big(\mathbf{x}(t),\mathbf{u}(t)\big)$ satisfies

$$\max_{i,j}|u_i(t)-u_j(t)|\to 0, \quad \sup_t\max_{i,j}|x_j(t)-x_i(t)|<\infty \tag{9}$$

i.e. the asymptotic flocking condition. For the next result the Assumptions of Theorem 3.10 hold. Since $a_{ij}(\cdot,\cdot)$ is uniformly upper bounded, we set

$$m:=\max_{i\in[N]}\sup_{t,y}\sum_j a_{ij}(t,y)<\infty. \tag{10}$$

The central result of this section reads:

Theorem 4.1: Consider the system (7) with condition (8) and its solution $\big(\mathbf{x}(t),\mathbf{u}(t)\big)$. The following conditions hold:

1) Static scrambling connectivity. The solution exhibits asymptotic flocking if
$$S(\mathbf{u}^0)<\int_{S(\mathbf{x}^0)}^{\infty}f(w)dw \tag{11}$$
2) Static routed-out branching connectivity. The solution exhibits asymptotic flocking if
$$S(\mathbf{u}^0)<\frac{(1-e^{-mB})^{\gamma}}{m^{\gamma}\gamma B}\int_{P_{\mathbf{x}^0,\mathbf{u}^0}^{\gamma,B}}^{\infty}f^{\gamma}(s)ds \tag{12}$$
where $P_{\mathbf{x}^0,\mathbf{u}^0}^{\gamma,B}=\max\{S(\mathbf{x}^0),|S(\mathbf{x}^0)-S(\mathbf{u}^0)\gamma B|\}$, $m=\max_i|N_i|f(0)$.
3) Switching connectivity. The solution exhibits asymptotic flocking if
$$S(\mathbf{u}^0)<\frac{(1-e^{-m\epsilon})^{\sigma}}{m^{\sigma}\sigma B}\int_{P_{\mathbf{x}^0,\mathbf{u}^0}^{\sigma,B}}^{\infty}f^{\sigma}(s)ds \tag{13}$$
where $\sigma=l^*([N/2]+1)$, l^* with the meaning of Remark 2.5 and $\epsilon>0$ with the meaning of Assumption 3.9.

Proof: We begin with the first connectivity condition, where there is at least one agent affecting the rest of the group. We follow the same path as in Theorem 3.5 for $\mathbf{u}$ and show that

$$\frac{d}{dt}S\big(\mathbf{u}(t)\big) \le -f\big(S(\mathbf{x}(t))\big)S\big(\mathbf{u}(t)\big) \Rightarrow$$
$$S\big(\mathbf{u}(t)\big) \le e^{-\int_0^t f(S(\mathbf{x}(w)))dw} S\big(\mathbf{u}^0\big)$$

so that asymptotic flocking will occur with exponential rate of convergence if $S\big(\mathbf{x}(t)\big) \le r$ for some $r > 0$. For this, we follow [14] and introduce the functional

$$V_1(\mathbf{x}, \mathbf{u}) = S(\mathbf{u}) + \int_0^{S(\mathbf{x})} f(w)dw \tag{14}$$

so that along a solution of (7) $\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ where $\mathbf{u}(t) = \dot{\mathbf{x}}(t)$ we have

$$\frac{d}{dt}V_1(t) = \frac{d}{dt}V_1\big(\mathbf{x}(t), \mathbf{u}(t)\big)$$
$$\le -f\big(S(\mathbf{x}(t))\big)S\big(\mathbf{u}(t)\big) + f\big(S(\mathbf{x}(t))\big)S\big(\mathbf{u}(t)\big) = 0$$

so that $V_1(t) \le V_1(0)$ which is equivalent to

$$S(\mathbf{u}(t)) + \int_0^{S(\mathbf{x}(t))} f(w)dw \le S(\mathbf{u}^0) + \int_0^{S(\mathbf{x}^0)} f(w)dw$$

From the imposed condition Eq. (11) on the initial data we deduce that there exists r' such that

$$S(\mathbf{u}^0) = \int_{S(\mathbf{x}^0)}^{r'} f(w)dw$$

so that $S\big(\mathbf{x}(t)\big) \ge S\big(\mathbf{x}^0\big)$

$$0 \le S\big(\mathbf{u}(t)\big) \le \int_{S(\mathbf{x}^0)}^{r'} f(w)dw - \int_{S(\mathbf{x}^0)}^{S(\mathbf{x}(t))} f(w)dw$$

which makes sense if $S\big(\mathbf{x}(t)\big) \le r'$. Pick $r = \max\{r', S(\mathbf{x}^0)\}$ to conclude that condition (11) ensures that the flock of birds will remain connected, hence they will coordinate their speeds exponentially fast.

For the second part, the flock is static and routed-out branching, hence it is routed-out branching over the interval $[t - B, t]$, for any $t > 0$ and $B > 0$. Let $m < \infty$ be defined as usual and

$$W(\mathbf{x}(s)) = mI - D(\mathbf{x}(s)) + A(\mathbf{x}(s)),$$
$$C(t, s) = e^{-mB}\delta(s - (t - B))I + e^{-m(t-s)}W(\mathbf{x}(s))$$

Finally for the scrambling index γ of the topological graph $\mathbb{G}_{P(\mathbf{x}(t))}$ (which is independent of time) $P_B^{(\gamma)}\big(\mathbf{x}(t)\big)$ is stochastic from Proposition 3.7 and has the same scrambling index as $P\big(\mathbf{x}(t)\big)$. Since the corresponding graph $\mathbb{G}_W$ is independent of time, so will be the scrambling index γ. We follow the first part of Theorem 3.5

$$\begin{aligned} S\big(\mathbf{u}(t)\big) &\le \tau\big(P_B^{(\gamma)}(\mathbf{x}(t))\big)S\big(\mathbf{u}(t - \gamma B)\big) \\ &\le \Big(1 - cf^{\gamma}\big(S(\mathbf{x}(t))\big)\Big)S\big(\mathbf{u}(t - \gamma B)\big) \end{aligned} \tag{15}$$

with $c := \frac{(1-e^{-mB})^{\gamma}}{m^{\gamma}}$ and $S\big(\mathbf{x}(t)\big) \ge r$ for r such that $f(r) = \frac{me^{-mB}}{1-e^{-mB}}$. We define the functional

$$V_2(\mathbf{x}, \mathbf{u}) = \int_{t-\gamma B}^{t} S\big(\mathbf{u}(s)\big)ds + c\int_0^{S(\mathbf{x})} f^{\gamma}(s)ds$$

the derivative of $\dot{V}_2$ along the solution of Eq. (7), $\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ is

$$\dot{V}_2(t) = S\big(\mathbf{u}(t)\big) - S\big(\mathbf{u}(t - \gamma B)\big) + cf^{\gamma}\big(S(\mathbf{x}(t))\big)S\big(\mathbf{u}(t)\big) \le 0$$

in view of Lemma 3.3 (from which it is deduced that $S\big(\mathbf{u}(t)\big) \le S\big(\mathbf{u}(t - \gamma B)\big), \forall t$ and Eq. (15). Then for $t \ge \gamma B$ we have $V_2(t) \le V_2(\gamma B)$ which is equivalent to

$$\int_{t-\gamma B}^{t} S\big(\mathbf{u}(s)\big)ds + c\int_0^{S(\mathbf{x}(t))} f^{\gamma}(s)ds \le$$
$$\int_0^{\gamma B} S\big(\mathbf{u}(s)\big)ds + c\int_0^{S(\mathbf{x}(\gamma B))} f^{\gamma}(s)ds$$

Let the following condition hold

$$\int_0^{\gamma B} S\big(\mathbf{u}(s)\big)ds < c\int_{S(\mathbf{x}(\gamma B))}^{\infty} f^{\gamma}(s)ds \tag{16}$$

and we pick r' such that

$$\int_0^{\gamma B} S\big(\mathbf{u}(s)\big)ds = c\int_{S(\mathbf{x}(\gamma B))}^{r'} f^{\gamma}(s)ds$$

then from the last inequality, $S\big(\mathbf{x}(t)\big) \le S\big(\mathbf{x}(\gamma B)\big)$ implies

$$0 \le \int_{S(\mathbf{x}(t))}^{r'} f^{\gamma}(s)ds \Rightarrow S\big(\mathbf{x}(t)\big) \le r'$$

so that the flock remains bounded and exponential speed alignment is ensured. Finally, we show that Eq. (12) implies Eq. (16). Indeed,

$$\int_0^{\gamma B} S\big(\mathbf{u}(s)\big)ds \le \gamma B S(\mathbf{u}^0)$$

from Lemma 3.3. Now we look for a lower bound of $S\big(\mathbf{x}(t)\big)$. If $S\big(\mathbf{x}(t)\big) \ge S(\mathbf{x}^0)$ from the form of Eq. (7) the rate at which $S\big(\mathbf{x}(t)\big)$ may shrink can be deduced from the extreme scenario of $\mathbf{x}^0 = (x^0, 0, \ldots, 0)$, $x^0 < 0$ so that $S(\mathbf{x}^0) = x^0$ and $\mathbf{u}^0 = (u^0, 0, \ldots, 0)$, $u^0 \ne 0$ with $S(\mathbf{u}^0) = |u^0|$. Neglecting the averaging effect which will inevitably diminish $S\big(\mathbf{u}(t)\big)$, $x^0 < 0$ implies that the first bird at t will have approached (or bypassed) the rest of the group by $-|x^0| + |u^0|t$. All in all, at $t = \gamma B$

$$S\big(\mathbf{x}(\gamma B)\big) \ge \max\big\{S(\mathbf{x}^0), |S(\mathbf{x}^0) - S(\mathbf{u}^0)\gamma B|\big\} = P_{\mathbf{x}^0,\mathbf{u}^0}^{\gamma,B}$$

so that

$$\int_{S(\mathbf{x}(\gamma B))}^{\infty} f^{\gamma}(s)ds \ge \int_{P_{\mathbf{x}^0,\mathbf{u}^0}^{\gamma,B}}^{\infty} f^{\gamma}(s)ds$$

then

$$S(\mathbf{u}^0) < \frac{(1 - e^{-mB})^{\gamma}}{m^{\gamma}\gamma B}\int_{P_{\mathbf{x}^0,\mathbf{u}^0}^{\gamma,B}}^{\infty} f^{\gamma}(s)ds.$$

The case of switching connectivity is treated as in Theorem 3.5 and the use of V_2 after substituting γ with σ. Then Eq. (13) substitutes Eq. (12) to ensure asymptotic flocking. ■

Remark 4.2: In the case of static connectivity, $\gamma = 1$ implies that Eq. (11) and Eq. (12) coincide as $B \downarrow 0$.

V. DISCUSSION

To the best of our knowledge, the use of the coefficient of ergodicity for continuous time consensus systems was limited to cases of increased connectivity [14]. An improved proof of that approach is provided in Section III-.1 whereas in the next section the full problem was attacked. The key idea is to invert the differential operator to an integral one so that concepts from the adapted result of Theorem 2.4 would apply. Indeed, the necessarily positive value $B > 0$ considered in §III-.2 (Assumption 3.8) has very similar meaning to the necessary positive time one needs to classify the communication classes in a continuous time Markov chain [3] .

Now, the combination of using the coefficient of ergodicity in integral equations provides two significant advantages: At first, one needs no symmetry conditions on the communication weights as unlike the Fielder number in algebraic graph theory [1], the contraction coefficient ρ has no applicability issues. Secondly, one is free to allow discontinuous jumps on the connectivity regime without the mobilization of elaborated generalized concepts of solution of differential equations. The proof we provided is elegant and concise to include vanishing communication weights and hence non-uniform types of convergence. Within our framework, the work of [11] is a special case where the coupling weights are assumed to be uniformly lower bounded. Moreover, the rate estimates are explicit and hence they are used for the study of the non-linear model (7), providing new initial conditions for asymptotic flocking.

The main drawback is that the contraction coefficient generally provides very conservative rate estimates. In particular much more conservative than its symmetric counterpart, the Fiedler number [1]. The latter tool gains its power from the spectral properties of topological graphs and incorporates much more information than τ.

In [17] a number of different non-linear as well as stochastic models is also studied with respect to the framework of Section III, but this discussion is omitted due to space limitation.

REFERENCES

[1] F. R K Chung, *Spectral graph theory*, vol. 92, American Mathematical Society, 1997.

[2] F. Cucker and S. Smale, *Emergent behavior in flocks*, IEEE Transactions on Automatic Control **52** (2007), no. 5, 852–862.

[3] G. Grimmett and D. Stirzaker, *Probability and random processes*, 3 ed., Oxford Univ. Press, 2003.

[4] D. J. Hartfiel, *Markov set-chains*, vol. 1695, Springer, 1998.

[5] I. C. F. Ipsen and T. M. Selee, *Ergodicity coefficients defined by vector norms*, SIAM J. Matrix Anal. Appl. **32** (2011), 153–200.

[6] A. Jadbabaie, J. Lin, and A. S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Transactions on Automatic Control **48** (2003), no. 6, 988–1001.

[7] U. Krause, *A discrete nonlinear and non-autonomous model of consensus formation.*, Communications in difference equations (2000), 227–236.

[8] Jan Lorenz and Dirk A. Lorenz, *On conditions for convergence to consensus.*, IEEE Trans. Automat. Contr. **55** (2010), no. 7, 1651–1656.

[9] A. A. Markov, *Extensions of the law of large numbers to dependent quantities (in russian)*, Izn. Fiz. - Matern. Obsch. Kazan Univ. **15** (1906).

[10] M. Mesbahi and M. Egerstedt, *Graph theoretic methods in multiagent networks*, Princeton University Press, 2010.

[11] L. Moreau, *Stability of continuous-time distributed consensus algorithms*, 43rd IEEE Conference on Decision and Control (2004) (2004), 21.

[12] R. Olfati-Saber and R. M. Murray, *Consensus problems in networks with switching topology and time-delays*, IEEE Transactions on Automatic Control **49** (2004), no. 09, 1520 – 1533.

[13] H. Qing and W. M. Haddad, *Distributed nonlinear control algorithms for network consensus*, Automatica **44** (2008), no. 9, 2375 – 2381.

[14] Motsch S. and Tadmor E., *Heterophilious dynamics enhances consensus*, SIAM: Review **56** (2014), no. 4, 577–621.

[15] E. Seneta, *Non-negative matrices and markov chains*, Springer, 2006.

[16] C. Somarakis and J. S. Baras, *The consensus problem under vanishing communication weights*, Proceedings of the 20th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2012) (2012), 1–8.

[17] C. Somarakis and J.S. Baras, *Towards a unified theory of consensus*, doi:10.13016/M2MW3B TR-2014-09, Digital Repository of ISR, UMD, 2014, http://hdl.handle.net/1903/15969.

[18] J. Tsitsiklis, D. Bertsekas, and M. Athans, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Transactions on Automatic Control **31** (1986), no. 9, 803–812.