THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2014-09

Towards a Unified Theory of Consensus

Christoforos Somarakis and John S. Baras

Technical Report submitted to
INSTITUTE FOR SYSTEMS RESEARCH
Digital Repository at the University of Maryland
Copyrighted Material , Number **0**, October **2014**

Website: http://drum.lib.umd.edu/

pp. **1–46**

TOWARDS A UNIFIED THEORY OF CONSENSUS

Christoforos Somarakis[1] and John S. Baras[2]
The Institute For Systems Research
Univeristy Of Maryland
College Park, MD 20740, USA

Abstract. We revisit the classic multi-agent distributed consensus problem. We adopt mild connectivity assumptions and with a novel application of the contraction coefficient we obtain simple yet general and unifying results both in discrete and continuous time. Furthermore, we extend the discussion to stochastic settings. We apply our approach to a wide variety of linear, non-linear consensus and flocking algorithms proposed in the literature and we derive new conditions for asymptotic consensus.

Contents

Key words and phrases. Consensus Dynamics, Flocking Networks.
[1]csomarak@umd.edu
[2]baras@umd.edu

1. **Introduction.**

Self-organized dynamics lie in the core of modern complex systems a most interesting branch of which is their application in networked control.

Examples of networks that illustrate a collective behavior as a result of local dynamic interaction among the nodes of the network are ubiquitous both in nature and in human societies. Ants cooperate together to form a nest or to transfer provisions and birds form flocks and fly together enhancing their hunting abilities. Humans interact and socialize by exchanging opinions and sometimes may converge to a fairly common view (especially after choosing a leader). Engineers build mobile communication or robotic networks which coordinate their behavior by local exchange of information. These are all examples of collaborative control of multi-agent complex systems. The self-organized aspect of these systems is usually understood by a decentralized, local exchange of information. The central phenomenon in these examples is the manner agents, as individuals, exchange information on a state of interest and update this state so that eventually all agents' states 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 the so-called *emergence of consensus* among agents.

In its classic version, a formal framework includes a finite number of agents $N \geq 2$, each agent $i = 1, \ldots, N$ of which possesses a value of interest. This value, denoted by $x_i \in \mathbb{R}$, evolves under the following averaging schemes, expressed either in discrete or continuous time:

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

The quantities a_{ij}'s are non-negative numbers that model the influence of agent j on i and essentially characterize the interdependence of agents, the connectivity regime and eventually the process of the asymptotic alignment. For the discrete model it holds that $\sum_j a_{ij} \equiv 1$ and for the continuous model $\sum_j a_{ij} \equiv 0$. In particular, the extensive amount of proposed frameworks, much of which is discussed below, are concerned with different versions of the connectivity weights a_{ij}. Equations of type (0.1) are also known as first order consensus schemes.

1.1. **Review of the existing literature.**

Being fundamental part of the large field of multi-agent self-organized dynamical system, consensus algorithms have continuously drawn the attention of researches among various scientific communities. In this section, we conduct a thorough, yet by no means complete, review of the existing models in the literature.

1.1.1. *Detereministic models.*

The interest in distributed iterative schemes has a long history in the literature [11], [41], [18]. In the control community, distributed computation over networks begins with the work of Tsitsiklis et. al. [47] where problems of asynchronous agreement and parallel computing were considered for Eq. (0.1).

A theoretical framework for solving consensus problems was introduced by Olfati-Saber et al. in [35] while in their seminal paper Jadbabaie et al. [22] studied a model of asymptotic alignment proposed by Viscek et al. [48]. Both these works consider populations of autonomous agents that exchange information under the assumption of symmetric communication, i.e. $a_{ij} = a_{ji}$. While Olfati-Saber et al. followed algebraic graph theory methods [6], Jadbabaie et al. based their results on the theory of non-negative matrices and non-homogeneous Markov Chains, [41]. The novelty of these works concern switching communication networks, i.e. communications weights $a_{ij}(t)$ that vary over time and may be positive or zero at each t. In [22] this switching connectivity regime asks for a connectivity condition to ensure asymptotic coordination, known as *recurrent connectivity*.

Nonlinear versions of Eq. (0.1) have also appeared in the literature, [33, 28, 38] both in discrete and continuous time versions, mainly as extensions to the linear ones. In [33, 34], Moreau studied non-linear discrete time and linear continuous time versions of consensus algorithms. In [33] the author used set valued Lyapunov functions and proved that in systems of type

$$x_i(t+1) = f_i(t, x_1(t), \ldots, x_N(t)), \qquad i = 1, \ldots, N$$

agreement among agents is reached as $t \to \infty$ on condition that each agent's state lies inside the convex hull of their neighbors' previous states. A similar argument was built for the linear continuous time version of the algorithm [33]. Another non-linear algorithm was analyzed in [28] very similar to this of Moreau, while in [38] the authors study the continuous time non-linear model

$$\dot{x}_i = \sum_j a_{ij}(t)\big(g_{ij}(x_j) - g_{ij}(x_i)\big)$$

as an extension to [35] and includes both static and switching connectivity conditions.

Another non-linear model for asymptotic consensus is with the use of the passivity property of the coupling functions. In [2, 36] the authors introduce and study the asymptotic properties of the model

$$\dot{x}_i = \sum_j g_{ij}\big(t, x_j(t) - x_i(t)\big)$$

with $g_{ij}(t, z)$ being a passive function in z. This model is another interesting generalization of Eq.(0.1) and very similar to the Kuramoto model [26] for synchronization but even more similar to Krause's opinion dynamics model [25].

In a series of papers, [9, 10, 7] Cucker et al. introduced a model for speed alignment among birds with connectivity weights that depend on their relative distance. In the most fundamental form the model reads

$$\begin{aligned} \dot{x}_i &= u_i \\ \dot{u}_i &= \sum_j a(|x_i - x_j|)(u_j - u_i) \qquad i = 1, \ldots, N \end{aligned} \tag{0.2}$$

where x_i denotes the position of the bird i and u_i denotes its speed. These systems are generally known in the literature as 2^{nd} order consensus models and have attracted an enormous attention from the Applied Mathematics community. The central objective is again the derivation of conditions under which the birds align their speed while the flock remains shaped. Mathematically, this translates to the following *asymptotic flocking* condition

$$|u_i(t) - u_j(t)| \to 0, \qquad \sup_t |x_i(t) - x_j(t)| < \infty, \qquad \forall \ \ i, j = 1, \ldots, N.$$

In the work of Cucker the communication between i and j is assumed to have the form

$$a(|x_i - x_j|) = \frac{K}{(\sigma + |x_i - x_j|^\beta)}.$$

Based on this particular form, sufficient conditions for asymptotic flocking, as functions of the parameters K, σ, β, are derived. The fact that this particular coupling under a_{ij} is instrumental in the stability analysis is very restricting. For this reason, simplified proofs, improvements and extensions both to microscopic and macroscopic level, were developed in the years to follow [17, 16, 40]. These works mainly consider continuous time versions of the Cucker-Smale model.

1.1.2. *Stochastic models.* Essentially any positive value of a_{ij} signifies the existence of connection between j and i (in the sense that j affects i e.g. by signal transmission). Real-world networked systems, however, suffer from various communication failures or creations between nodes. For example when agents are moving, some existing connections may fail as obstacles may appear between agents or assuming proximity graphs, one agent may enter the effective region of other agents. Therefore a standard abstraction is this when agents are connected via a network that changes with time due to link/node failures, packet drops etc. Such variations in topology can happen randomly, and this motivates the investigation of consensus problems under a stochastic framework.

1. Linear Consensus. Let us now discuss a number of studies based on stochastic versions of Eq. (0.1). Hatano et al. consider in [20, 32] an agreement problem over random information networks where the existence of an information channel between a pair of elements at each time instance is probabilistic and independent of other channels. In [37], Porfiri and Stilwell provide sufficient conditions for reaching consensus almost surely in the case of a discrete linear system, where the communication flow is given by a directed graph derived from a random graph process, independent of other time instances. Under a similar communication topology model, Tahbaz-Salehi and Jadbabaie in [45] provide necessary and sufficient conditions for almost sure convergence to consensus and in [46] they extend the applicability of their necessary and sufficient conditions to strictly stationary ergodic random graphs. In [31] Matei et al. consider the linear consensus system (0.1) under the assumption that the communication flow between agents is modeled by a randomly switching graph. The switching is determined by a homogeneous, finite-state Markov chain and each communication pattern corresponds to a state of the Markov process. Then necessary and sufficient conditions are provided to guarantee convergence to average consensus in the mean square sense and in the almost sure sense.
2. Nonlinear Consensus. In [29] the authors study local synchronization of nonlinear discrete-time dynamical networks with time varying couplings both in deterministic and stochastic varying connections, using variational stability methods. A different line of stochastic formulation of consensus algorithms is proposed for the 2^{nd} order model in [15, 1, 8], where additive Brownian noise is added to Eq. (0.2) and sufficient conditions are derived for consensus in the almost sure sense are derived. Both these works rely on the algebraic properties of the symmetric communication weights so that a Lyapunov stability argument for stochastic stability is developed.

1.2. **Motivation & contribution.** In their vast majority, all works rely on a fundamental assumption: The exchange of information among any two communicating nodes occurs under established connection with a time varying 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 etc. [41, 6, 4, 32, 18, 19] towards proving asymptotic consensus.

The following elementary example shows that if the uniform lower bound assumption is lifted, consensus is not ensured:

Example 1.1. Consider the 2-D dynamical system

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} 1-f(t) & f(t) \\ g(t) & 1-g(t) \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

where $f(t) = K_f/t^2$ and $g(t) = K_g/t^2$for $t \geq 1$ for $K_f, K_g < 1$. Then for $|x(0) - y(0)| = \delta \neq 0$ it can be shown that

$$\begin{aligned} |x(t+1) - y(t+1)| &= (1 - f(t) - g(t))(x(t) - y(t)) \\ &= \delta \prod_{i=0}^{t} (1 - f(i)) \to C \sin(\pi \sqrt{K_f + K_g}) \end{aligned}$$

for some constant $C > 0$ according to the Euler-Wallis formula. So for $\sqrt{K_f + K_g} \notin \mathbb{Z}$ consensus is not achieved.

The importance of this underlying assumption has been noted before [33] and we strenuously remark that whichever work does not explicitly state it, it should be subject to criticism. Distributed consensus systems that bear non-uniform positive weights have appeared in the literature [25], [9] and it is this condition that makes the corresponding stability problems particularly challenging.

This paper can be considered as a considerable outgrowth of [44] where the problem of vanishing communications was considered in a fairly basic level and only for discrete time linear consensus algorithms. The organization and contribution of this monograph is discussed in the rest of this section.

In §2, we state the main nomenclature to be used in this work and review elements on elements of algebraic graph theory, non-negative matrix theory, measure dynamical systems, stochastic differential equations and fixed point theory. We also provide preliminary results by extending parts of the theory of non-negative matrices that will come at hand in the following.

In §3, we revisit the classic deterministic linear consensus problem. More specifically in paragraph 3.1 the discrete time case is considered where we provide convergence results without the assumption of uniform lower-bound on connectivity weights. The results are stated progressively and they range from the strongest (elementary static, increased), to the mildest (recurrent) connectivity regime. Dropping the uniform lower bound imposes 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 paragraph 3.2 the case of continuous time is analyzed in a surprisingly similar vein, since the same mathematical machinery tools is essentially exploited: After turning the problem from a differential equation into an integral equation one, we manage to obtain bounds of contraction in just like the discrete case.

In §4, we consider the stochastic version of the problem from two different perspectives. In the first case, we impose uncertainty in the existence of connections and we establish probabilistic rules to control these particular dynamics and we propose a new framework based on measure preserving dynamical systems. The main contribution is the unified results this framework provides among a number of important relevant versions proposed in the literature. In particular we will show that convergence to consensus can happen only with a positive probability and not almost surely, whenever the weights are free to asymptotically vanish. The second stochastic approach deals with uncertainties in the form of equations. This leads to stochastically perturbed differential equations, where the noise is supplied by Brownian processes. We elaborate on the deterministic case and provide new results for asymptotic flocking in the almost sure and mean square sense.

A thorough discussion of the overall obtained results with concluding remarks is held in §5.

2. **Notations & Definitions.**

In this section, we will discuss the general theoretical framework in which we will establish our results. $\mathbb{Z}$ is the set of integers, $\mathbb{N}$ is the set of naturals and $\mathbb{R}$ the set of real numbers. For $N \in \mathbb{N}$, $\mathcal{V} = \{1, \dots, N\}$. We will work in the N-dimensional Euclidean space $\mathbb{R}^N$ any 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

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

A *rank*-1 is a $N \times N$ matrix M is such that it has identical rows and for which $M\mathbf{x} \in \Delta,\ \forall \mathbf{x} \in \mathbb{R}^N$. 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 and satisfies the triangular inequality, but $S(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in \Delta$. By $\mathbb{1}$ we understand the $N-$dimensional vector with all entries equal to 1 and obviously $S(c\mathbb{1}) = 0$ for any $c \in \mathbb{R}$. By I we understand the $N \times N$ identity matrix. By $||\cdot||_p$ we denote the p norm where in particular $\mathbf{x}^T\mathbf{x} = ||\mathbf{x}||_2^2$.

2.1. **Algebraic graph theory.**

By a *topological directed graph* $\mathbb{G}$ we understand the pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is the (static) set of vertexes, $\mathcal{E} = \{(i,j) : i,j \in \mathcal{V}\}$ 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 subset $N_i : \{j \in \mathcal{V}, (i,j) \in \mathcal{E}\}$, all the vertexes adjacent to i. 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 $l_m = j$. The graph $\mathbb{G}$ is *connected* if any vertex is a route. 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 be identically equal to zero. This results in the spectral property that 0 is a always an eigenvalue of L and for any other eigenvalue $\lambda \in \mathbb{C}$ of L, $\Re\{\lambda\} > 0$ if and only if $\mathbb{G}$ is connected. Two vertexes $i, j \in \mathcal{V}$ *communicate* if there is a path from i to j and a path from j to i. A vertex is *essential* if whenever there is a path from i to j then there is a path from j to i. A vertex is called *inessential* if it is not essential. All essential vertexes are divided into communication classes and all inessential vertexes that communicate with at least one vertex may be divided into inessential classes such that all vertexes within a class communicate. All such classes are *self-communicating*. Each remaining inessential vertex communicates with no vertexes and individually forms an inessential class called *non self-communicating*. By $\mathcal{S}$ we denote the family of graphs with fixed N vertexes and self-edges on every node, and by $\mathcal{T} \subset \mathcal{S}$ the set of graphs each of which is routed-out branching.

2.1.1. *Agreement dynamics.* By the term *agreement dynamics* we classify in this paper the elementary case of static time invariant linear consensus systems of the form (0.1) and in particular the continuous time model. Then in vector for

$$\dot{\mathbf{x}} = -L\mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{x}^0 \in \mathbb{R}^N \tag{1}$$

where $L = D - A$ is the Laplacian. It is very well known that $\mathbf{x}(t) \to \mathbb{1}\mathbf{c}^T\mathbf{x}^0$ where $\mathbf{c}$ belongs to the subspace of $\mathbb{R}^N$ spanned by the left eigenvectors of L associated with the zero eigenvalue. Consequently if $\mathbb{G}_A$ is routed-out branching then L has only one zero eigenvalue and $\mathbf{c} \in \mathbb{R}^N$ is such that $\mathbf{c}^T L = 0$, $c_i \geq 0$ and $\sum_i c_i = 1$, i.e. the (unique) normalized eigenvector, so that $\mathbf{c}^T\mathbf{x}^0 \in \Delta$. The rate of convergence is dictated by the real part of the second smallest eigenvalue of L. It is noted that in the symmetric case $a_{ij} = a_{ji}$, $\mathbf{c} = \mathbb{1}\frac{1}{N}$ and the limit point is the average of the initial states. The latter is true even if the symmetric weights depend on time.

For more on Graph Theory and relevant methods on multi-agent systems, the interested reader is referred to [13, 4, 32].

2.2. **Non-negative matrix theory.** A *non-negative* matrix $P = \{p_{ij}\}$ is such that $p_{ij} \geq 0$ for all i, j.[3] 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 $m = 1$ we have, of course, the well-known stochastic matrix. We will now introduce and discuss, in detail, the standard mathematical tool to deal with infinite products of stochastic matrices.

Given an m-stochastic matrix $P = [p_{ij}]$, the quantity

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

is *the coefficient of ergodicity* of P. A crucial set of properties of τ is discussed in the following result:

Theorem 2.1. [18] *For any m-stochastic matrix P and $\mathbf{z} \in \mathbb{R}^N$ the following properties hold:*

1. $||\boldsymbol{\delta} P|| \leq \tau(P)||\boldsymbol{\delta}||$, *for all real row vectors $\boldsymbol{\delta}$ such that $\boldsymbol{\delta}\mathbb{1} = 0$.*
2. $S(P\mathbf{z}) \leq \tau(P)S(\mathbf{z})$
3. $|\lambda| \leq \tau(P)$, *for any (possibly complex) eigenvalue λ of P with the property that $\lambda \neq m$.*

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 [30]. In the literature there exists an abundance of similar ideas: the coefficient of ergodicity is also known as contraction coefficient, Markov coefficient, Dobrushin coefficient, Birkhoff coefficient, Hajnal diameter, as each corresponding researcher has arrived at it independently and/or under different setups, [41, 19]. For a recent review on the coefficients of ergodicity we refer to [21].

Remark 2.2. Property (1) of Theorem 2.1 leads to the sub-multiplicative property: for P_1, P_2 m stochastic matrices, their product P_1P_2 constitutes and m^2 stochastic matrix and it satisfies

$$\tau(P_1P_2) \leq \tau(P_1)\tau(P_2).$$

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

The sub-multiplicative property becomes particularly useful when $m = 1$ exactly because at $m = 1$ the set of stochastic matrices becomes closed under matrix multiplication. As Theorem 2.1 the coefficient τ applies to dynamics of the type

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

with P being m-stochastic. A straightforward extension is this when P acts as an abstract linear operators on $\mathbb{Z}$ and it is summarized in the following Theorem which is actually the first result of this work:

Theorem 2.3. *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 $\mathbf{z}^* = \big(z_1(s_1), \dots z_N(s_N)\big)$ for $s_i \in I$ and

$$\begin{aligned}\tau(W_I) &= \frac{1}{2}\max_{h,h'}\sum_{k=1}^{N}\int_{s\in I}|p_{hk}(s) - p_{h'k}(s)|ds \\ &= m - \min_{h,h'}\sum_{k=1}^{N}\min\left\{\int_{s\in I}p_{hk}(s)ds, \int_{s\in I}p_{h'k}(s)ds\right\}\end{aligned} \tag{3}$$

The proof of this result relies on the first mean value theorem for integration and a technical lemma, both of which are cited below:

Lemma 2.4 (The first mean value theorem for integration)**.** *If $G \in C^0[J,\mathbb{R}]$ 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 2.5 (Lemma 1.1 of [18])**.** *Suppose $\boldsymbol{\delta} \in \mathbb{R}^N$ such that $\boldsymbol{\delta}^T\mathbb{1} = 0$ 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*

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

$\mathbf{e}_i$ is the row vector $(0,\dots,0,1,0,\dots,0)$ with 1 at the i^{th} position.

Proof of Theorem 2.3. 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

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

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 in for $s \in I_l$, $k \in \mathcal{V}$ and it is not identically zero. Then for fixed I_l we apply Lemma 2.4 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 $\boldsymbol{\delta}_l^T\mathbb{1} = 0$. Hence Lemma 2.5 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 [18]). Then if we let $S(\mathbf{z}*) = \max_l S(\mathbf{z}_l^*)$, we obtain the bound

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

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

$$\frac{1}{2}\max_{h,h'} \sum_k \int_{s\in I} |p_{hk}(s) - p_{h'k}(s)|ds = m - \min_{h,h'} \sum_k \min\left\{ \int_{s\in I} p_{hk}(s)ds, \int_{s\in I} p_{h'k}(s)ds \right\}$$

□

Similarly, for the expression

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

one can show, along the lines of the proof of Theorem 2.3 that if $W_I^{(2)} = \int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q)dqds$ is stochastic, then

$$S(\mathbf{w}) \leq \tau(W_I^{(2)})S(\mathbf{z}^*) \tag{4}$$

for some $\mathbf{z}^* = \left(z_1(s_{(ij)}^{(1)}), z_2(s_{(ij)}^{(2)}), \ldots, z_N(s_{(ij)}^{(N)})\right)$ 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

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

so long as $\int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q)dqds$ 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 which coupling elements p_{ij}, it holds that

$$\tau < m.$$

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 the extension in Eq. (3)) 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 [41, 19].

The properties of stochastic matrices and their products play a crucial role in the analysis to follow and the standard approach is through graph theory: Any non-negative (and in particular stochastic) matrix P can be represented as a graph $\mathbb{G}_P$ with its 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 and use the terminology of §2.1.

A non-negative matrix P is called *irreducible* if $\mathbb{G}_P$ consists of a single essential class and a stochastic 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 $P^t \to \mathbb{1}\mathbb{1}^T c$ 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 [18]. As the product of stochastic matrices is stochastic as well, the preceding notions can be extended to study the behavior of the non-homogeneous products of stochastic matrices. We exclusively study *backward products* of stochastic matrices defined as

$$P_{p,h} := P_{p+h}P_{p+h-1}\cdots P_{p+1} = [p_{ij}^{p,h}].$$

for $p \geq 0, h \geq 1$. 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 sub-graph 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 particular, 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 2.6. 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 $\bigcup_{v=0}^{i-1} \mathcal{Y}_v$, denoted by $l^* := \max_i\{l_{i,i-1}\}$.

For more on the dynamics of products of non-negative matrices the reader is referred to [19, 18, 41].

2.3. **Elements of dynamical system theory.** Let $(\mathbb{X}, \mathcal{B}, \mu)$ be a finite measure space (that is $\mu(\mathbb{X}) < \infty$) and for the rest of the paper we assume, without loss of generality, $\mu(\mathbb{X}) = 1$. We define a measurable transformation $T : \mathbb{X} \to \mathbb{X}$, as a map with the property that $T^{-1}(\mathcal{B}) \subset \mathcal{B}$. $T : \mathbb{X} \to \mathbb{X}$ is *measure preserving* if $\mu(T^{-1}B) = \mu(B)$ for any $B \in \mathcal{B}$. A measure preserving transformation is called *ergodic* if for any $B \in \mathcal{B}$ with the property that $T^{-1}B = B$ either $\mu(B) = 0$ or $\mu(B) = 1$.

For a collection of probability spaces, $\{(\mathbb{X}_t, \mathcal{B}_t, \mu_t)\}_{t\in\mathbb{N}}$, we define the *product* probability space in the natural way: $\mathbb{X} = \prod_{t\in\mathbb{N}} \mathbb{X}_t$ and a point $\chi \in \mathbb{X}$ is considered to be the sequence $\chi = \chi_0\chi_1\chi_2\ldots$ where $\chi_t \in \mathbb{X}_t$. The σ-algebra $\mathcal{B}(\mathbb{X})$ generated by subsets of $\mathbb{X}$ is the product of σ-algebras $\mathcal{B}_i$ and it is defined as the intersection of all σ-algebras that contain the collection of subsets of $\mathbb{X}$:

$$\mathcal{J} = \left\{ \prod_{j\leq n_1-1} \mathbb{X}_j \times \prod_{n_1\leq j\leq n_2} A_j \times \prod_{j\geq n_2+1} \mathbb{X}_j \right\} = \{\chi \in \mathbb{X} : \chi_j \in A_j, j \in [n_1, n_2]\}_{0\leq n_1\leq n_2}^{A_j\in\mathcal{B}_j}$$

each of which is a *measurable rectangle* (or a *cylinder*). On each of the above rectangles we attach the value $\prod_{t=n_1}^{n_2} \mu_t(A_t)$ and this can be extended to a probability measure μ on $(\mathbb{X}, \mathcal{B})$ in the standard way [49], concluding the definition of the product probability space $(\mathbb{X}, \mathcal{B}, \mu)$. A measurable transformation $T : \mathbb{X} \to \mathbb{X}$ on the product space, known as *shift*, is defined by $T(\chi_0\chi_1\chi_2\ldots) = \chi_1\chi_2$ and it may attain all the desired properties of measure preserving and ergodicity. By $T^t\chi$ we mean the element $\chi_t\chi_{t+1}\ldots$ and we will also use the projection map $\{T^t\chi\} = \chi_t$, $\chi_t \in \mathbb{X}_t$.

For more on dynamical systems and ergodic theory the reader is referred to [23, 49].

2.4. **Elements of stochastic differential equations.** For given $t_0 \in \mathbb{R}$ and a probability space $(\Omega, \mathcal{U}, \mathbb{P})$, a collection of random variables $\{\mathbf{Y}_t : t \geq t_0\}$, each of which $\mathbf{Y}_t : \Omega \to \mathbb{R}^N$ if $\mathcal{U}$-measurable, consists a stochastic process. The σ-algebra generated by $\mathbf{Y}_t$ is the smallest sub σ-algebra of $\mathcal{U}$ to which $\mathbf{Y}_t$ is measurable.

Let $\mathbf{B}$ be an N-dimensional Brownian motion defined on $[t_0, \infty)$ and $\mathbf{Y}_0$ is an N-dimensional random variable independent of $\mathbf{B}(t_0)$. The σ-algebra generated by $\mathbf{Y}_0$ and the history of the Brownian motion up to (and including) time $t \geq t_0$ is

$$\mathcal{U}_t := \mathcal{U}\big(\mathbf{B}(s)|_{t_0 \leq s \leq t}, \mathbf{Y}^0\big).$$

The family $\{\mathcal{U}_t\}$ is called a *filtration* and a process $\mathbf{Y}_t$ is *adapted* to $\mathcal{U}_t$ if $\mathbf{Y}_t$ is $\mathcal{U}_t$-measurable for all $t \geq t_0$. The set $(\Omega, \mathcal{U}, \mathcal{U}_t, \mathbb{P})$ consists a complete filtered probability space. Fix $T > t_0$ and let $\mathbf{b} : \mathbb{R}^N \times [t_0, T] \to \mathbb{R}^N$, $B : \mathbb{R}^N \times [t_0, T] \to M^{N\times N}$ are given vector valued and matrix valued deterministic functions, respectively. Then an $\mathbb{R}^N$ valued stochastic process $\mathbf{Y}_t$ is a solution of the Itô stochastic differential equation

$$\begin{cases} d\mathbf{Y}_t = \mathbf{b}(\mathbf{U}_t, t)dt + B(\mathbf{Y}_t, t)d\mathbf{B}_t \\ \mathbf{Y}_{t_0} = \mathbf{Y}^0 \end{cases} \qquad t_0 \leq t \leq T$$

provided:

1. $\mathbf{Y}_t$ is a $\mathcal{U}_t$-adapted process.
2. $\mathbb{E}\big[\int_{t_0}^{T} |b_i(\mathbf{Y}_t, t)| dt\big] < \infty$.
3. $\mathbb{E}\big[\int_{t_0}^{T} |B^2_{ij}(\mathbf{Y}_t, t)| dt\big] < \infty$.
4. $\forall t \in [t_0, T]$

$$\mathbf{Y}_t = \mathbf{Y}^0 + \int_{t_0}^{t} \mathbf{b}(\mathbf{Y}_s, s)ds + \int_{t_0}^{t} B(\mathbf{Y}_s, s)d\mathbf{B}_s, \qquad a.s.$$

The existence and uniqueness (in probability) of a solution to the above initial value problem is guaranteed after assuming a local Lipschitz condition on $\mathbf{b}$ and B and a linear sub-growth of $|\mathbf{b}(\mathbf{x}, t)|$ and $|\mathbf{B}(\mathbf{x}, t)|$ with respect to $\mathbf{x}$. For more on Itô calculus and explicit types of solutions in certain linear stochastic differential equations as well as in asymptotic behavior of stochastic processes the reader is referred to the excellent textbooks [14, 24] and especially to [3].

2.5. **Fixed point theory.** A pair $(\mathbb{M}, \rho)$ is a metric space if $\mathbb{M}$ is a set and $\rho : \mathbb{M} \times \mathbb{M} \to [0, \infty)$ such that $\rho(y, z) \geq 0$ with equality to hold if and only if $z = y$, $\rho(z, y) = \rho(y, z)$ and $\rho(y, z) \leq \rho(z, x) + \rho(z, y)$. A complete metric space is such that every Cauchy sequence in $\mathbb{M}$ converges in $\mathbb{M}$. The major result of Fixed Point Theory is Banach's Contraction Principle

Theorem 2.7. *Let $(\mathbb{M}, \rho)$ be a complete metric space and $\mathcal{Q} : \mathbb{M} \to \mathbb{M}$ to satisfy*

$$\rho(\mathcal{Q}y_1, \mathcal{Q}y_2) \leq \alpha \rho(y_1, y_2)$$

for any $y_1, y_2 \in \mathbb{M}$ and $\alpha \in [0, 1)$. Then there exists a unique $y \in \mathbb{M}$ such that $\mathcal{Q}y = y$.

In the study of stability of solutions of differential equations is occasionally desirable to derive estimates on the rate of convergence to an asymptotic state. In this case, stability by means of fixed point theory proves to be very convenient. The existence of a fixed point of a solution operator in a weighted complete metric space implies that the solution (i.e. the fixed point) attains the property of convergence with the prescribed rate (weight). For the purpose of this paper the weight is defined as a rate function $h : [t_0, \infty) \to [1, \infty)$ such that $h(t_0) = 1$, h is monotonically increasing and $\lim_t h(t) = \infty$. Examples of complete (weighted) metric spaces are provided in [5]. For more on fixed point theory the reader is referred to [39, 42, 5].

3. Deterministic Consensus.

We begin the discussion with dynamics that change with time in a purely deterministic fashion. In the first two subsections we study both discrete and continuous time variants of linear consensus algorithms, whereas in the last subsection we discuss applications in non-linear algorithms as those introduced in §1.1.1.

3.1. Discrete time.

Consider $N < \infty$ agents with values $x_i \in \mathbb{R}$. At each time t, agent i updates its value $x_i(t) \in \mathbb{R}$ according to

$$\begin{cases} x_i(t+\eta) - x_i(t) = \eta \sum_{j \neq i} a_{ij}(t)\big(x_j(t) - x_i(t)\big), \ t \geq 0 \\ x_i(0) = x_i^0 \end{cases} \tag{5}$$

for $i \in \mathcal{V}$ and $\eta > 0$ fixed which without loss of generality we will set $\eta = 1$.

We look for solutions of (5) with the property that $x_i(t) - x_j(t) \to 0$ as $t \to \infty$ $\forall i, j \in \mathcal{V}$. This is equivalent to $x_i(t) \to k$ as $t \to \infty$, $\forall i \in \mathcal{V}$. Indeed, note that for $\mathbf{x}(t) = \big(x_1(t), \ldots, x_N(t)\big)$ for $t \geq 0$, it holds that for $a_{ii}(t) = 1 - \sum_j a_{ij}(t) > 0$

$$x_i(t) \in [\min_i x_i^0, \max_i x_i^0]$$

then the $\omega(\mathbf{x}^0)$ limit set is non-empty, closed and invariant with respect to (5). Then if $x_i(t) - x_j(t) \to 0$, any point in the ω limit set will lie in Δ. Indeed for $\mathbf{x}(0) \in \omega$, we have $\mathbf{x}(0) \in \Delta$ as well and the solution $\mathbf{x}\big(t, 0, \mathbf{x}^0\big)$ will be in Δ for all $t \geq 0$. A similar argument for the continuous time counterpart of (5) is made in §3.2.

3.1.1. *Static connectivity I.*

Assumption 3.1. *The connectivity weights $a_{ij}(t) : \mathbb{Z}_+ \to \mathbb{R}_+$ are defined such that for any t the corresponding graph $\mathbb{G}_P$ is scrambling and that*

$$\sum_{j \in N_i} a_{ij}(t) \leq m < 1, \ a_{ij}(t) > 0 \Rightarrow a_{ij}(t) \geq f(t), \ i \neq j$$

where f is a positive function with the property that $\exists M \in [0, \infty)$ so that $f(t) \in (0, 1-m]$ for $t \geq M$.

Remark 3.2. It is an easy exercise to show that if $\sum_j a_{ij}(t) \leq m < 1$ then a sufficient condition for $f \leq 1 - m$ is $m \leq \frac{N}{N+1}$.

We rewrite (5) as

$$x_i(t+1) = a_{ii}(t)x_i(t) + \sum_{j \in N_i} a_{ij}(t)x_j(t)$$

where $a_{ii}(t) := 1 - \sum_{j \in N_i} a_{ij}(t)$ which is positive under Assumption 3.1. In vector form the solution reads

$$\mathbf{x}(t+1) = P(t)\mathbf{x}(t)$$

where $P(t)$ is the stochastic matrix

$$P(t) := \begin{pmatrix} 1 - \sum_{j \neq 1} a_{1j}(t) & a_{12}(t) & \cdots & a_{1N}(t) \\ a_{21}(t) & 1 - \sum_{j \neq 2} a_{2j}(t) & \cdots & a_{2N}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}(t) & a_{N2}(t) & \cdots & 1 - \sum_{j \neq N} a_{Nj}(t) \end{pmatrix} \tag{6}$$

This particular type of stochastic matrices will be used in the consensus dynamics both in discrete and continuous time.

Remark 3.3. An essential property of $P(t)$ is that the diagonal elements are strictly positive. For this reason, for any P_1, P_2, the product P_2P_1 is also stochastic with the same structure as (6).

Next we exploit the structure of $P(t)$ to obtain non-trivial upper bound estimates for $\tau\big(P(t)\big)$.

Lemma 3.4. *Let $P(t)$ be a stochastic matrix with the form of Eq. (6). Under Assumption 3.1, the coefficient of ergodicity, τ satisfies*

$$\tau\big(P(t)\big) < 1 - f(t), \qquad t >> 1.$$

Proof. Pick any $t \geq M$ for M having the meaning of Assumption 3.1. The definition of τ requires to find the two rows i, j that minimize the sum

$$\sum_k \min\{a_{ik}, a_{jk}\}.$$

The structure of $P(t)$, in Eq.(6) and Assumption 3.1 implies that there exists a column i^* with strictly positive elements. Then for arbitrary i, j

$$\sum_k \min\{a_{ik}, a_{jk}\} \geq \min\{a_{ii^*}, a_{ji^*}\}$$

from which two cases are to be considered

a. $i, j \neq i^*$ and $\min\{a_{ii^*}, a_{ji^*}\} = a_{ii^*}$ so that $a_{ii^*} \geq f(t)$.
b. $i = i^*$ or $j = i^*$ and $\min\{a_{ii^*}, a_{ji^*}\} = a_{ii^*}$ so that $a_{ji^*} \geq a_{i^*i^*} = 1 - d_i(t) \geq 1 - m$.

so that

$$\tau\big(P(t)\big) \leq 1 - \min\big\{1 - m, f(t)\big\} = 1 - f(t).$$

□

Theorem 3.5. *Under Assumption 3.1, the system (5) converges to consensus if $\sum_s f(s) = \infty$.*

Proof. The results is a straightforward application of Theorem 2.1 and Lemma 3.4. Given the initial vector $\mathbf{x}^0$, the general solution of (5) at time t is

$$\mathbf{x}(t) = P(t-1)P(t-2)\cdots P(0)\mathbf{x}^0 = P_{-1,t}\mathbf{x}^0$$

and for $t \geq M + 1$,

$$S\big(\mathbf{x}(t)\big) \leq \tau\big(P(t-1)\big)S\big(\mathbf{x}(t-1)\big) \leq \prod_{s=M}^{t-1} \tau\big(P(s)\big)S\big(\mathbf{x}(M)\big) \leq \prod_{s=M}^{t-1} \tau\big(P(s)\big)S\big(\mathbf{x}^0\big)$$

where we have used the sub-multiplicativity property of $\tau(\cdot)$. Then

$$\begin{aligned}
\lim_{t\to\infty} S\big(\mathbf{x}(t)\big) &\leq S\big(\mathbf{x}^0\big) \lim_{t\to\infty} \prod_{s=M}^{t-1} \tau\big(P(s)\big) \\
&\leq S\big(\mathbf{x}^0\big) \lim_{t\to\infty} \prod_{s=M}^{t-1} \big(1 - f(s)\big) \\
&\leq S\big(\mathbf{x}^0\big) \lim_{t\to\infty} \prod_{s=M}^{t-1} e^{-f(s)} \\
&\leq S\big(\mathbf{x}^0\big) \lim_{t\to\infty} e^{-\sum_{s=M}^{t-1} f(s)} \\
&= 0.
\end{aligned}$$

□

Assumption 3.1 allows for the connectivity graph $\mathbb{G}_{P(t)} = (\mathcal{V}, \mathcal{E}(t))$ to be time dependent only to the point where it is sufficiently connected, i.e. it must be scrambling for all sufficiently large t. We can generalize the above result, firstly to the case of deterministic switching connection topologies and secondly to the general case of the recurrent connectivity.

3.1.2. *Switching connectivity I.* We begin with the first generalization.

Assumption 3.6. *There exists $B > 0$ and $M \in [t_0, \infty)$ such that for any $t \geq M$, the family of stochastic matrices $\{P(s)\}_{s=t}^{t+B-1}$ contain at least one scrambling matrix.*

Assumption 3.6 is straightforward extension of Assumption 3.1 and it implies the following corollary of Theorem 3.5:

Corollary 3.7. *Under Assumption 3.6, the solution of Eq. (5) converges to a constant value if f satisfies, either*

$$\sum_{n=1}^{\infty} f(s_n) = \infty$$

where $\{s_l\}_{l\geq 1}$ is the sequence of times such that $s_l \in I_l = [M+(l-1)B, M+lB-1]$ and $P(s_l)$ is scrambling, or

$$0 < \sup_{l\geq 1} \max_{s\in I_l} \frac{\sum_{n\in I_l} f(n)}{f(s)} < \infty, \ \sum_t f(t) = \infty.$$

Proof. For any integer $n \geq \frac{B}{M}$, the solution at $t = nB$ can be written as

$$\mathbf{x}(t) = \prod_{l=n}^{1} P_{M+(l-1)B-1,B} \prod_{s=M}^{0} P(s)\mathbf{x}^0.$$

In view of Remark 3.3 and Assumption 3.6, for $l = 1, \ldots, n$ the matrix $P_{(l-1)B-1,B}$ is scrambling with $\tau = 1 - f(s_l)$ for $s_l \in [(l-1)B, lB-1]$. On the first imposed condition on f, the proof proceeds exactly as in Theorem 3.5. Finally, it is an easy exercise to show that the second imposed condition on f implies the first one and the proof is concluded. □

3.1.3. *Static connectivity II.* In this section we relax the connectivity assumptions considered in §3.1.1, to simple connectivity. In this case, it is by no means clear that $\tau(P(t)) < 1$ so that the contraction effect of $P(t)$ is captured by τ. It is true, however, that under certain conditions the left product $P_{t,h}$ may be scrambling. In the example to follow we illustrate the effect of non-uniform connectivity weights.

Example 3.8. Consider a network of 4 agents, characterized by the stochastic matrix

$$P(t) = \begin{pmatrix} 1-a_{12}(t) & a_{12}(t) & 0 & 0 \\ a_{21}(t) & 1-a_{21}(t)-a_{23}(t) & a_{23}(t) & 0 \\ 0 & a_{32}(t) & 1-a_{32(t)}-a_{34}(t) & a_{34}(t) \\ 0 & 0 & a_{43}(t) & 1-a_{43}(t) \end{pmatrix}$$

with $a_{ij}(t) \geq f(t) > 0$ and $f(t)$ a monotonically decreasing function. It is easy to check that $\tau\big(P(t)\big) = 1$ for all t. However, in the static case where $B = 1$, any product of such two matrices is scrambling. Indeed straightforward calculation reveals that $P(t)P(t-1)$ is scrambling with the first and fourth row to have nonzero entries that sum up to

$$a_{12}(t)a_{23}(t-1)+a_{32}(t-1)a_{43}(t)\geq 2f^2(t)$$

so that for $t >> 1$

$$\tau(P(t)P(t-1))=1-2f^2(t)$$

and

$$S\big(\mathbf{x}(t)\big)\leq\prod_{i=t}^{1}\tau\big(P(i)P(i-1)\big)S\big(\mathbf{x}^0\big)\leq S\big(\mathbf{x}^0\big)e^{-2\sum_{i=1}^{[\frac{t}{2}]-1}f^2(2i-1)}\to 0$$

on condition that $\sum_i f^2(2i-1)$ diverges, so that asymptotic consensus occurs.

We can make this argument more rigorous.

Assumption 3.9. *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 $f(t)\to 0$.*

Remark 3.10. Under Assumption 3.9, Lemma 3.4 implies that for any scrambling matrix $P(t)$ we have $\tau\big(P(t)\big)<1-f(t)$ for t large enough. The function f is appropriately chosen so that the non-trivial upper bound of τ is controlled by this function. Consequently, since any product $P_{t-\gamma,t}$ is stochastic with the same structure as well (see Remark 3.3) and scrambling, the lower bound of the off-diagonal elements of $P_{t-\gamma,t}$, yields $\tau\big(P_{t-\gamma,t}\big)<1-f^{\gamma+1}(t)$ when t is large enough. Again, how large should that t become depends on the choice of f which in its turn would depend on the connectivity structure of $P_{t-\gamma,t}$.

Lemma 3.11. *For fixed $p\geq 0, h\geq 1$, consider the matrix product $P_{p,h}$ for each $P(s)|_{s=p+1\ldots p+h}$ defined as in (6) with the property that $\tau\big(P_{p,s}\big)=1$ for $s\in 1\ldots h-1$ and $\tau\big(P_{p,h}\big)<1$. If $a:=\min_{s\in\{s=p+1\ldots p+h\}}\{a_{ij}(s)\}\in[0,1-m]$ where $\sum_j a_{ij}(s)\leq m<1,\forall s=p+1\ldots p+h$, then it holds that:*

$$\tau\big(P_{p,h}\big)\leq 1-a^h.$$

Proof. We will use induction on h. For $h=2$, $P_{p,2}=P_{p+2}P_{p+1}$. Since $\tau\big(P_{p+1}\big)=1$ and $\tau\big(P_{p,2}\big)<1$ then there exists a strictly positive column of $P(p+2)P(p+1)=[p_{ij}]$. Let i be this column. Then

$$p_{ii}=(1-d_i(p+2))(1-d_i(p+1))+\sum_j a_{ij}(p+2)a_{ji}(p+1)\geq(1-m)^2$$

$$p_{ji}=\sum_k a_{jk}(p+2)a_{kj}(p+1)\geq\min\{(1-m)a,a^2\}=a^2$$

from the bound on a the result follows. If the statement is true for $h=l$, then for $h=l+1$, similar calculations yield the bounds $p_{ii}\geq(1-m)^{l+1},\ p_{ji}\geq a^{l+1}$ so that $\tau\big(P_{p,h}\big)\leq 1-a^{l+1}$. □

With this in mind we state the next result that concerns static topological connectivity.

Theorem 3.12. *Let Assumption 3.9 hold. If there exists $M>0$ such that for any $t_1,t_2\geq M$, we have $P(t_1)\sim P(t_2)$ so that the corresponding graph $\mathbb{G}_{P(t_1)}=\mathbb{G}_{P(t_2)}$ is routed-out branching, the solutions of (5) reach global consensus if f satisfies*

$$\sum_t f^\gamma(M+t\gamma-1)=\infty$$

or if f satisfies

$$0 < \sup_{t \geq M} \frac{\sum_{i=t\gamma}^{(t+1)\gamma-1} f^{\gamma}(i)}{f^{\gamma}((t+1)\gamma-1)} < \infty, \ \sum_{t \geq M} f^{\gamma}(t) = \infty$$

where γ is the (time independent) scrambling index of $P(t)$.

Proof. $P(t_1) \sim P(t_2)\ \forall t_1, t_2$ implies that the corresponding communication graph is static routed-out branching. This is equivalent to $P(t)$ being regular. Then the backward product $P_{t-\gamma,\gamma}$ is a scrambling stochastic matrix with

$$\tau\big(P_{t-\gamma,\gamma}\big) \leq 1 - f^{\gamma}(t)$$

when t is large, according to Lemma 3.11. From the sub-multiplicativity property of τ, we can proceed as in Theorem 3.5 for $t \geq n\gamma$, to obtain the estimate

$$S(\mathbf{x}(t)) \leq S(\mathbf{x}^0) e^{-\sum_{i=1}^{n-1} f^{\gamma}(M+i\gamma-1)}$$

so that asymptotic consensus occurs in view of the first condition. It is, again, an easy exercise to show that the second condition implies the first. □

3.1.4. *Switching connectivity II.* We will consider now the mildest connectivity condition for the discrete case, i.e. for any $t \geq 0$, $P(t) \in \mathcal{S}$ but any product of matrices must belong in $\mathcal{T}$, over a uniformly bounded interval of time. In particular we have the following condition:

Assumption 3.13. *There exist $M > 0$ and $B \geq 1$, such that for all $t \geq M$, $\mathbb{G}_{P_{t,B}} \in \mathcal{T}$.*

Assumption 3.13 is the well-known condition of recurrent connectivity [22].

Theorem 3.14. *Let Assumptions 3.9 and 3.13 hold. Then we have unconditional asymptotic consensus for the solution of (5) if f satisfies: either*

$$\sum_t f^{\sigma}(M + t\sigma - 1) = \infty,$$

or

$$0 < \sup_{t \geq M} \frac{\sum_{i=t\sigma}^{(t+1)\sigma-1} f^{\sigma}(i)}{f^{\sigma}((t+1)\sigma-1)} < \infty \text{ and } \sum_{t \geq M} f^{\sigma}(t) = \infty$$

where $\sigma = l^*([N/2]+1)B$ and l^* with the meaning of Remark 2.6.

Proof. We have that $P_{t,B} \in \mathcal{T}$ with $\gamma = \gamma_{t,B}$. Then $P_{t+B,B} \in \mathcal{T}$ as well and

$$\gamma_{t,2B} \leq \begin{cases} \max\{\gamma_{t,B}, \gamma_{t+B,B}\} - 1, & \mathbb{G}_{P_{t,B}} \subset \mathbb{G}_{P_{t+B,B}} \text{ or } \mathbb{G}_{P_{t+B,B}} \subset \mathbb{G}_{P_{t,B}} \\ \max\{\gamma_{t,B}, \gamma_{t+B,B}\}, & o.w. \end{cases}$$

If $\mathbb{G}_{P_{t,B}}$ is not a sub-graph of $\mathbb{G}_{P_{t+B,B}}$ and $\mathbb{G}_{P_{t,B}}$ is not a sub-graph of $\mathbb{G}_{P_{t+B,B}}$ or vice-versa, it holds that $\mathcal{E}^C_{P_{t,B}} \cap \mathcal{E}^C_{P_{t+B,B}} \neq \emptyset$. An element of this set is the pair (i,j) such that $[\mathbb{G}_{P_{t,B}}]_{ij} = 0$ and $[\mathbb{G}_{P_{t+B,B}}]_{ij} > 0$ or vice versa. This element, however, will be a member of $\mathcal{E}_{P_{t,2B}}$ since $[\mathbb{G}_{P_{t,2B}}]_{ij} \geq [\mathbb{G}_{P_{t+B,B}}]_{ij}[\mathbb{G}_{P_{t,B}}]_{jj} > 0$. From the discussion on the partitioning of $\mathcal{T}$ with respect to the scrambling indexes and for $l^* = \max_i\{l_{i,i-1}\}$,

$$\gamma_{t,l^*B} \leq \max\{\gamma_{t,B}, \gamma_{t+B,B}\} - 1$$

Consequently $P_{t,l^*([N/2]+1)B}$ will be scrambling. Set $\sigma = l^*([N/2]+1)B$. Consider the solution

$$\mathbf{x}(t) = P(t)P(t-1)\cdots P(0)\mathbf{x}^0$$

Since $P_{M,\sigma-1}$ will be scrambling at time $M + \sigma B t$ the estimate

$$S\big(\mathbf{x}(M+\sigma Bt)\big) \leq e^{-\sum_t f^{\sigma}(M+t\sigma-1)} S(\mathbf{x}^0) \tag{7}$$

holds. The proof is concluded by the non-summability of $\sum_t f^\sigma(M+t\sigma-1)$. Again it is only an easy exercise to show that the second condition implies the first. $\square$

Remark 3.15. This first set of theorems on discrete time linear consensus is a generalization of existing results in the literature [22]. In case of uniformly bounded weights all the relevant results are recovered in a more concise manner whereas in the vanishing communication topology, the interdependence of the connectivity regime and the rate at which the connections vanish is illustrated.

As a first application of the Theorems 3.5, 3.12 and 3.14 we have the following example.

Example 3.16. Consider the system (5) and its solution $\mathbf{x}(t)$ with with $f(t) \geq \omega(t^{-\alpha})$ i.e. for large t , f dominates a function that vanishes as slow as $t^{-\alpha}$. Then under Assumption 3.1 and Theorem 3.5 or under Assumption 3.6 and Corollary 3.7 convergence to consensus is guaranteed for $\alpha \in [0,1]$. On the other hand the simple routed-out branching condition gives the sufficient condition $\alpha \in [0,1/\gamma]$ for Assumption 3.9 and Corollary 3.12 or $\alpha \in [0,1/\sigma]$ under Assumption 3.13 and Theorem 3.14.

Example 3.17 (Application in flocking dynamics.)**.** Let us review now a second order consensus model, as the one's proposed in the literature [9, 10]. For $N < \infty$ and $i = 1,\dots,N$ we consider N birds with positions x_i and velocities v_i to coordinate their speeds according to the following algorithm:

$$\begin{cases} x_i(t+\eta) &= x_i(t) + \eta v_i(t) \\ v_i(t+\eta) &= v_i(t) + \eta \sum_j a_{ij}\big(\mathbf{x}(t)\big)\big(v_j(t) - v_i(t)\big) \end{cases} \tag{8}$$

with $\mathbf{x}(0) = \mathbf{x}^0$, $\mathbf{v}(0) = \mathbf{v}^0$ as given initial data and $\eta > 0$ the fixed mesh value. We assume $a_{ij}(\mathbf{x}) \geq f\big(S(\mathbf{x})\big)$ whenever $a_{ij} \neq 0$ for f an upper bounded non-negative and non-increasing function. To simplify notation we write $\mathbf{x}(t)$ for $\mathbf{x}(th)$ and the same for $\mathbf{v}$. The next result stems out of Theorem 3.5.

Corollary 3.18. *Consider Eq.* (8) *and its solution* $\big(\mathbf{x}(t),\mathbf{v}(t)\big) \in \mathbb{R}^{N\times N}$. *Let Assumption 3.1 to hold for the connectivity weights* a_{ij} *for* $t_0 = 0$, $M = 0$. *Set* $C := \limsup_{x\to\infty} xf(x) \leq \infty$. *If* $\eta < \frac{1}{\max_i |N_i| f(0)}$ *and*

$$S(\mathbf{v}^0) < C - f\big(S(\mathbf{x}^0)\big)S(\mathbf{x}^0)$$

then $(\mathbf{x}(t),\mathbf{v}(t))$ *satisfies*

$$\sup_{t\geq 0} S\big(\mathbf{x}(t)\big) < \infty \text{ and } \lim_{t\to\infty} S\big(\mathbf{v}(t)\big) = 0.$$

Proof. The smallness on the mesh η is imposed to make the second part of Eq. (8) a well posed consensus algorithm.[4] Then, under the Assumption 3.1 the corresponding graph is $\mathbb{G}_{P(\mathbf{x}(t))}$ is scrambling for every t and hence it justifies the contraction bound

$$S\big(\mathbf{v}(t)\big) \leq \big(1 - \eta f\big(S\big(\mathbf{x}(t)\big)\big)\big)S\big(\mathbf{v}(t)\big)$$

Proving that $\sup_t S\big(\mathbf{x}(t)\big) < \infty$ implies that the flock will always remain sufficiently connected so that a_{ij} will be uniformly lower bounded and a direct application of Theorem 3.5 suffices to prove speed coordination. Indeed let t^* be the first time that $S\big(\mathbf{x}(t^*)\big) \geq S\big(\mathbf{x}(t^*-1)\big)$. Then for the "functional"

$$V(t) = S\big(\mathbf{v}(t)\big) + f\big(S\big(\mathbf{x}(t)\big)\big)S\big(\mathbf{x}(t)\big)$$

[4]Indeed a large h may lead to instability,[9]

we have

$$
\begin{aligned}
&V(t^*) - V(t^*-1) \leq \\
&\leq \big(1 - \eta f\big(S\big(\mathbf{x}(t^*-1)\big)\big)\big) S\big(\mathbf{v}(t^*-1)\big) - S\big(\mathbf{v}(t^*-1)\big) + \\
&+ f\big(S\big(\mathbf{x}(t^*)\big)\big) S\big(\mathbf{x}(t^*)\big) - f\big(S\big(\mathbf{x}(t^*-1)\big)\big) S\big(\mathbf{x}(t^*-1)\big) \\
&\leq -\eta f\big(S\big(\mathbf{x}(t^*-1)\big)\big) S\big(\mathbf{v}(t^*-1)\big) + f\big(S\big(\mathbf{x}(t^*-1)\big)\big)\big[S\big(\mathbf{x}(t^*)\big) - S\big(\mathbf{x}(t^*-1)\big)\big] \\
&\leq -\eta f\big(S\big(\mathbf{x}(t^*-1)\big)\big) S\big(\mathbf{v}(t^*-1)\big) + \eta f\big(S\big(\mathbf{x}(t^*-1) S\big(\mathbf{v}(t^*-1)\big) = 0
\end{aligned}
$$

but $V(t^*-1) \leq S(\mathbf{v}^0) + \eta f\big(S(\mathbf{x}^0)\big) S(\mathbf{x}^0) = V(0)$. Consequently, $V(t^*) \leq V(0)$ or equivalently

$$0 \leq S\big(\mathbf{v}(t^*)\big) \leq S(\mathbf{v}^0) + f\big(S(\mathbf{x}^0)\big) S(\mathbf{x}^0) - f\big(S\big(\mathbf{x}(t^*)\big)\big) S\big(\mathbf{x}(t^*)\big)$$

choosing $C' < C$ small enough so that $S(\mathbf{v}^0) = C' - S(\mathbf{x}^0)$ and this implies

$$0 \leq C' - f\big(S\big(\mathbf{x}(t^*)\big)\big) S\big(\mathbf{x}(t^*)\big)$$

and this implies that $S\big(\mathbf{x}(t^*)\big)$ is bounded. □

Notice that if $\lim_x x f(x) = \infty$ then we have asymptotic flocking without any condition on the initial data.

Remark 3.19. The above example provides a generalization of the discrete models known in the literature [9, 10] in the sense that a_{ij} are not assumed to be either symmetric or to have any particular expression. Moreover the above result can be generalized to the case of simple or even switching connectivity (Theorems 3.12 and 3.14). These cases are to be discussed next.

3.2. **Continuous time.** For $N < \infty$ number of agents we consider the following initial value problem:

$$
\begin{cases}
\dot{x}_i(t) &= \sum_{j \in N_i} a_{ij}(t)\big(x_j(t) - x_i(t)\big), \ t \geq t_0 \\
x_i(t_0) &= x_i^0
\end{cases}
\tag{9}
$$

where $i \in \mathcal{V}$.

When modeling failing signals and an overall switching connectivity regime, the researcher must consider the connectivity weights $a_{ij}(t)$ to "jump" from a positive value to zero in a discontinuous fashion. In this case Eq. (9) turns a differential equation with discontinuous right hand-side and a generalized notion of solution is usually considered, [12]. This system is simple enough to allow for a simple alternative. Assuming $a_{ij}(t)$ to be right continuous, a solution $\mathbf{x}(t, t_0, \mathbf{x}^0)$ of (9) is a continuous function with a right t-derivative that satisfies the differential equation for every $t \geq t_0$.[5] This solution also satisfies any classical integral equation which occurs after inverting Eq.(9) in the classical manner.

Assumption 3.20. *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 yet crucial remarks:

[5]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 (9) in almost every t and the same analysis applies.

Lemma 3.21. *Under Assumption 3.20 the solution* $\mathbf{x}\big(t, t_0, \mathbf{x}^0\big)$ *of Eq.* (9) *then* $x^0_{min} \le x_i(t) \le x^0_{\max}$, $\forall\ t \ge t_0, i \in \mathcal{V}$.

Proof. We will make a fixed point theory argument. The solution $\mathbf{x}\big(t, t_0, \mathbf{x}^0\big)$ satisfies

$$\mathbf{x}\big(t, t_0, \mathbf{x}^0\big) = e^{-\int_{t_0}^t D(s)ds}\mathbf{x}^0 + \int_{t_0}^t e^{-\int_s^t D(w)dw} A(s)\mathbf{x}(s)ds$$

Define the space of functions

$$\mathbb{M} = \big\{\mathbf{y}(t) \in C^0\big([t_0, \infty), \mathbb{R}^N\big) : \mathbf{y}(t) = \mathbf{x}^0, x^0_{min} \le y_i(t) \le x^0_{\max}\big\}$$

together with the weighted metric

$$\rho(\mathbf{y}_1, \mathbf{y}_2) = \sup_{t \ge t_0} e^{-\gamma(t-t_0)}||\mathbf{y}_1(t) - \mathbf{y}_2(t)||_1$$

where $e^{\gamma(t-t_0)}$ serves as rate function with $\gamma > 0$ to be determined. The pair $(\mathbb{M}, \rho)$ is a complete metric space. Define the vector valued function

$$(\mathcal{L}\mathbf{y})(t) := e^{-\int_{t_0}^t D(s)ds}\mathbf{x}^0 + \int_{t_0}^t e^{-\int_s^t D(w)dw} A(s)\mathbf{y}(s)ds$$

is such that $\mathbf{y} \in \mathbb{M}$ implies

$$\begin{aligned}
(\mathcal{L}y)_j(t) &\le e^{-\int_{t_0}^t d_j(s)ds} x^0_{\max} + \int_{t_0}^t e^{-\int_s^t d_j(w)dw} \sum_l a_{jl}(s) ds x^0_{\max} \\
&\le e^{-\int_{t_0}^t d_j(s)ds} x^0_{\max} + \big(1 - e^{-\int_{t_0}^t d_j(s)ds}\big) x^0_{\max} \\
&= x^0_{\max}
\end{aligned}$$

and similarly for the lower bound $(\mathcal{L})_j(t) \ge x^0_{\min}$. It follows that $\mathcal{L} : \mathbb{M} \to \mathbb{M}$. Set $m = \max_i \sup_{t \ge t_0} d_i(t) < \infty$ from the Assumption 3.20. Finally, for $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{M}$

$$\begin{aligned}
\rho(\mathcal{L}\mathbf{y}_1, \mathcal{L}\mathbf{y}_2) &\le \sup_{t \ge t_0} e^{-\gamma(t-t_0)} \int_{t_0}^t ||e^{-\int_s^t D(w)dw} A(s)||_1 ||\mathbf{y}_1(s) - \mathbf{y}_2(s)||_1 ds \\
&\le \sup_{t \ge t_0} e^{-\gamma(t-t_0)} \sum_j \int_{t_0}^t e^{-\int_s^t d_i(w)dw} d_j(s) e^{\gamma(s-t_0)} ds \rho(\mathbf{y}_1, \mathbf{y}_2) \\
&\le \sup_{t \ge t_0} M e^{-\gamma t} \int_{t_0}^t e^{\gamma s} ds \rho(\mathbf{y}_1, \mathbf{y}_2) \\
&\le \frac{m}{\gamma} \rho(\mathbf{y}_1, \mathbf{y}_2).
\end{aligned}$$

Then for $\gamma > m$ the operator $\mathcal{L}$ is a contraction and by Theorem 2.7 it attains a unique fixed point in $\mathbb{M}$. □

Lemma 3.22. *If* $\mathbf{x}(t, t_0, \mathbf{x}^0)$ *is the solution of Eq.* (9) *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.

Proof. From Lemma 3.21 we have that $\omega(\mathbf{x}^0)$ is non-empty, compact and connected and any element of which must lie in Δ. Take $\mathbf{x}^\omega \in \omega(\mathbf{x}^0)$ and consider the solution $\mathbf{x}(t, t_0, \mathbf{x}^\omega)$. Since $\mathbf{x}^\omega \in \Delta$ as well, we have that $\dot{\mathbf{x}} \equiv 0$, i.e. the solution is constant. □

3.2.1. *Static & switching networks I.* We begin the first round of results in this section 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 3.23. *Let Assumption 3.20 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 (9) to a constant value.

Proof. We write Eq. (9) in vector form

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

so that from Theorem 2.1 we obtain the bound

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

then

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

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 3.22. □

Remark 3.24. This is a generalization of the results obtained in [40] concerning continuous time consensus algorithms. In addition, furtherly 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 $\int^{\infty} f(s)ds = \infty$. The non-integrability condition is the continuous time counterpart of the non-summability of f imposed on Theorem 3.5.

3.2.2. *Static & switching networks II.* We will escalate the analysis with the study the dynamics of Eq. (9) under the recurrent connectivity condition. Define for $B \geq 0$, $s \in [t-B, t]$

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

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

Proposition 3.25. *Let Assumption 3.20 hold. For any* $B > 0$, $l \geq 1$, *the matrix*

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

whenever defined, is stochastic.

Proof. The matrix

$$P_B(t) := \int_{t-B}^{t} \left(e^{-mB}\delta(s_1-(t-B))I + e^{-m(t-s_1)}W(s_1) \right) ds_1$$

is stochastic. Indeed, the i^{th} row of $P_B(t)$ consists of the positive diagonal element

$$e^{-mB} + \int_{t-B}^{t} e^{-m(t-s_1)}(m-d_i(s_1))ds_1 = 1 - \int_{t-B}^{t} e^{-m(t-s_1)} d_i(s_1)ds_1$$

and the non-negative off-diagonal elements

$$\int_{t-B}^{t} e^{-m(t-s_1)} a_{ij}(s_1)ds_1.$$

since $d_i(s_1) = \sum_j a_{ij}(s_1)$, $P_B(t)$ is stochastic.

We proceed with induction: For $l = 2$,

$$\begin{aligned}
&P_B^{(2)}(t) = \\
&= \int_{t-B}^{t}\int_{s_1-B}^{s_1} \left(e^{-mB}\delta(s_1-(t-B))I + e^{-m(t-s_1)}W(s_1) \right)\cdot \\
&\qquad\qquad \cdot \left(e^{-mB}\delta(s_2-(s_1-B))I + e^{-m(s_1-s_2)}W(s_2) \right) ds_2 ds_1 \\
&= \int_{t-B}^{t}\int_{s_1-B}^{s_1} e^{-2mR}\delta(s_1-(t-B))\delta(s_2-(s_1-B))ds_2ds_1 I + \\
&+ \int_{t-B}^{t}\int_{s_1-B}^{s_1} e^{-mR}\delta(s_1-(t-B))e^{-m(s_1-s_2)}W(s_2)ds_2ds_1 + \\
&+ \int_{t-B}^{t}\int_{s_1-B}^{s_1} e^{-m(t-s_1)}W(s_1)e^{-mR}\delta(s_2-(s_1-B))ds_2ds_1 + \\
&+ \int_{t-B}^{t}\int_{s_1-B}^{s_1} e^{-m(t-s_1)}W(s_1)e^{-m(s_1-s_2)}W(s_2)ds_2ds_1
\end{aligned}$$

and straightforward calculations yield

$$\begin{aligned}
P_B^{(2)}(t) = e^{-2mB}I &+ \int_{t-2B}^{t-B} e^{-m(t-s_2)}W(s_2)ds_2 + e^{-mB}\int_{t-B}^{t} e^{-m(t-s_1)}W(s_1)ds_1 \\
&+ \int_{t-B}^{t}\int_{s_1-B}^{s_1} e^{-m(t-s_2)}W(s_1)W(s_2)ds_2ds_1
\end{aligned}$$

Now, every element of $P_B^{(2)}(t)$ is non-negative as a sum of non-negative matrices. It is only left to verify that $\sum_j [P_B^{(2)}(t)]_{ij} = 1$ for any i. Indeed, the first matrix contributes e^{-2mB}, the second and the third $e^{-mB} - e^{-2mB}$ and the fourth $(1 - e^{-mB})^2$, so eventually

$$e^{-2mB} + 2(e^{-mB} - e^{-2mB}) + (1 - 2e^{-mB} + e^{-2mB}) = 1$$

Let $P_B^{(l)}(t)$ be stochastic. Then the elements of $P_B^{(l+1)}(t)$ are non-negative by the same reasoning as above and finally, since

$$\begin{aligned}
&P_B^{(l+1)}(t)\\
&= \int_{t-B}^{t} C(t,s_1)\cdots\int_{s_l-B}^{s_l} C(s_l,s_{l-1})ds\\
&= e^{-mB}P_B^{(l)}(t) + (1-e^{-mB})P_B^{(l)}(t) -\\
&- \int_{t-B}^{t}\int_{s_1-B}^{s_1}\cdots\int_{s_l-B}^{s_l} C(t,s_1)C(s_1,s_2)\cdots\big(D(s_{l+1})-A(s_{l+1})\big)ds_{l+1}\dots d_{s_1}\\
&= P_B^{(l)}(t) - \int_{t-B}^{t}\int_{s_1-B}^{s_1}\cdots\int_{s_l-B}^{s_l} C(t,s_1)C(s_1,s_2)\cdots\big(D(s_{l+1})-A(s_{l+1})\big)ds_{l+1}\dots d_{s_1}
\end{aligned}$$

the sum of the i^{th} row of $P_B^{(l+1)}(t)$ equals 1 because the corresponding sum in the final integrand is zero (as it is a left multiplication of a matrix with a Laplacian matrix). ☐

Assumption 3.26. *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.*

In addition to the upper boundedness of $a_{ij}(t)$, the weights are also assumed to satisfy the dwelling time condition [22]. This ensures that in any subset of $\mathbb{R}_+$ with positive and bounded measure, the number of discontinuities must be finite. More rigorously:

Assumption 3.27. *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 3.28. *Let Assumptions 3.9,3.20, 3.26 and 3.27 hold. Unconditional asymptotic consensus for the solution of the system (9) 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 2.6.

Proof. We begin with the static case. The solution $\mathbf{x}$ of (9) satisfies

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

$$\begin{aligned}\mathbf{x}(t) &= \int_{t-B}^{t} \Big(e^{-mB}\delta(s-(t-B))I + e^{-m(t-s)}\big(mI - D(s) + A(s)\big)\Big)\mathbf{x}(s)ds \\ &= \int_{t-B}^{t} C(t,s_1)\mathbf{x}(s_1)ds_1 \\ &= \int_{t-B}^{t}\int_{s_1-B}^{s_1}\cdots\int_{s_{\gamma-1}-B}^{s_{\gamma-1}} C(t,s_1)C(s_1,s_2)\cdots C(s_{\gamma-1},s_\gamma)\mathbf{x}(s_\gamma)ds_\gamma\cdots ds_1\end{aligned}$$

Consequently from Theorem 2.3, Lemma 3.21 and Proposition 3.25 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 we can execute similar calculations as in Lemma 3.11:

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

with $s_0 = t$ and for $i \ne j^*$

$$\begin{aligned}&[P_B^{(\gamma)}(t)]_{j^*i} \ge \\ &\ge \int_{s_0-B}^{s_0}\int_{s_1-B}^{s_1}\cdots\int_{s_{\gamma-1}-B}^{s_{\gamma-1}}\sum_{l_0,\dots,l_{\gamma-1}} e^{-m(s_0-s_\gamma)}a_{il_0}(s_1)a_{l_0l_1}(s_2)\dots a_{l_{\gamma-1}j^*}(s_\gamma)ds_\gamma\cdots ds_1 \\ &> \int_{s_0-B}^{s_0}\int_{s_1-B}^{s_1}\cdots\int_{s_{\gamma-1}-B}^{s_{\gamma-1}} e^{-m(s_0-s_\gamma)}ds_\gamma\cdots ds_1 f^\gamma(t) = \frac{(1-e^{-mB})^\gamma}{m^\gamma}f^\gamma(t)\end{aligned}$$

For $t' \ge M$ large enough so that $f(t) \le \frac{me^{-mB}}{1-e^{-mB}}$ whenever $t \ge t'$, we obtain from Lemma 3.4 the estimate:

$$\tau\big(P_B^{\gamma}(t)\big) \le 1 - c_1 f^\gamma(t) \tag{10}$$

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

$$S\big(\mathbf{x}(t)\big) \le S\big(\mathbf{x}(t_i)\big) \le \big(1-c_1f^\gamma(t_i)\big)S\big(\mathbf{x}(t_i-\gamma B)\big) \le \big(1-c_1f^\gamma(t_i)\big)S\big(\mathbf{x}(t_{i-1})\big)$$

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 $t \ge t_i$

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

and the proof of the first part is concluded.

In the case of switching connectivity we proceed as above but with the remarks in the proof of Theorem 3.14. Then $P_B^{(\sigma)}(t)$ is scrambling for any t and the contraction estimate

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

where now $c_2 := \frac{(1-e^{-m\epsilon})^\sigma}{m^\sigma} > 0$ for $\epsilon > 0$ as defined in Assumption 3.27. □

Remark 3.29. Unconditional consensus is the term used to clarify that convergence is independent of the initial conditions.

Example 3.30. 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 $t \geq 0$ it holds that $a_{ij}(t) \neq 0 \Rightarrow 0 < a \leq a_{ij}(t) < \frac{1}{2}$ and also

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

for some fixed $\epsilon > 0$.

Then

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

where $\bar{d}_i(t,s) = e^{-3\epsilon}\delta(s-(t-3\epsilon)) + e^{-(t-s)}(1-d_i(s))$. Now for any $t \geq 0$,

$$P_{3\epsilon}(t) = \int_{t-3\epsilon}^{t} [C(t,s)]_{ij} ds =$$
$$= \begin{cases} 1 - e^{-t}\int_{t-3\epsilon}^{t} e^s d_i(s)ds, & i = j \\ e^{-t}\int_{t-3\epsilon}^{t} e^s a_{ij}(s)ds, & (i,j) \in \{(1,2),(2,1),(2,3),(3,2),(3,4),(4,3)\} \\ 0, & o.w. \end{cases}$$

and by construction of the switching signal, it can be easily shown that

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

elementwise; 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

$$S(\mathbf{x}(t)) \leq S(\mathbf{x}(3l\epsilon)) \leq (1 - 2a^2(1-e^{-\epsilon})^2)^l S(\mathbf{x}(0)) = Ke^{-\theta t}S(\mathbf{x}(0))$$

where $K := e^{\frac{2}{3}a^2(1-e^{-\epsilon})^2}$ and $\theta := 2a^2\frac{(1-e^{-\epsilon})^2}{3\epsilon}$, as it is dictated by Theorem 3.28 for $\{t_i\}$ any sub-sequence with 3ϵ interval and $f(t)$ to be lower bounded by a.

3.3. **Necessary conditions.** So far, we have discussed only sufficient conditions for consensus and it is exactly the effect of the contraction coefficient that points to such direction. Necessary conditions are rather rare. Here we will discuss necessary conditions for asymptotic consensus and we shall conclude on the discrepancy between the sufficient conditions partially because we used the contraction coefficient.

Theorem 3.31. *Consider the system* (9) *and its solution* $\mathbf{x}$ *with Assumption 3.20 to hold and the communication graph to be routed-out branching. Assume that over a population of* N *autonomous agents there is a cut of* $\mathcal{V}$ *such that* $\mathcal{V} = \mathcal{V}_1 \sqcup \mathcal{V}_2$ *so that for any* $(i,j) \in \mathcal{V}_1 \times \mathcal{V}_2$ *or* $(j,i) \in \mathcal{V}_1 \times \mathcal{V}_2$, $a_{ij}(t) > 0$ *implies* $\int^{\infty} a_{ij}(s)ds < \infty$. *If for any* $l_1, l_2 \in \mathcal{V}$, $x_{l_1}(t) - x_{l_2}(t) \to 0$ *implies* $|x_{l_1}(t) - x_{l_2}(t)| \leq e^{-\gamma t} S(\mathbf{x}^0)$ *then there exist initial conditions such that* $S\big(\mathbf{x}(t)\big) > 0$ *for any* $t \geq t_0$

Proof. Let the initial conditions be set such $x_l(0) < x_n(0)$ for $l \in \mathcal{V}_1$ and $m \in \mathcal{V}_2$. Consider then the subset $\mathcal{V}_{11}$ of $\mathcal{V}_1$ and accordingly the subset $\mathcal{V}_{22}$ of $\mathcal{V}_2$ which by assumption they must have connections between them. Let $i \in \mathcal{V}_{11}$ such that $x_i(t) \leq x_{i^*}(t)$ for any $i^* \in \mathcal{V}_{11}$ and $j \in \mathcal{V}_{22}$ such that $x_j(t) \geq x_{j^*}(t)$ for any $j^* \in \mathcal{V}_{22}$.

$$\dot{x}_i(t) \leq d_{ij}\big(x_j(t) - x_i(t)\big) + z_i(t)$$
$$\dot{x}_j(t) \geq d_{ji}\big(x_i(t) - x_j(t)\big) + z_j(t)$$

where $z_i(t) = \sum_{l \in \mathcal{V}_1} a_{il}\big(x_l(t) - x_i(t)\big)$, $z_j(t) = \sum_{l \in \mathcal{V}_2} a_{jl}\big(x_l(t) - x_j(t)\big)$ are functions that signify the interconnections among agents on the separated subsets. Now,

$$\frac{d}{dt}\big(x_j(t) - x_i(t)\big) \geq -\big(d_{ij}(t) + d_{ji}(t)\big)\big(x_j(t) - x_i(t)\big) + z_j(t) - z_i(t)$$

if either $z_i(t)$ or $z_j(t)$ do not vanish then $S(\mathbf{x}(t))$ will not converge to zero and there is nothing to prove. On the other hand we have by assumption that $|z_i(t) - z_j(t)| \leq 2(N-1)Ce^{-\gamma t}S(\mathbf{x}^0)$ for $(N-1)C$ to play the role of the uniform upper bound of $a_{ij}(t)$ according to Assumption 3.20. Next we set for simplicity $Q(s) = \big(d_{ij}(s) + d_{ji}(s)\big)$

Now, $\int_{t_0}^{\infty} Q(s)ds < \infty$ and this means that there is a sequence $\{t_n\}_{n\geq 1}$ and a constant $J_1 > 0$ such that

$$\int_0^{t_n} Q(s)ds \geq J_1.$$

Since

$$x_j(t) - x_i(t) \geq e^{-\int_0^t Q(s)ds}(x_j^0 - x_i^0) + \int_0^t e^{-\int_w^t Q(s)ds}\big(f_j(w) - f_i(w)\big)dw,$$

we have that

$$|x_j(t_n) - x_i(t_n)| \geq \left| e^{-J_1}|x_{ji}^0| - \int_0^{t_n} e^{-\int_w^{t_n} Q(s)ds} 2(N-1)Ce^{-\gamma w}dw S(\mathbf{x}^0) \right|.$$

Choosing $\left|e^{-J_1}|x_{ji}^0| - \frac{2(N-1)CS(\mathbf{x}^0)}{\gamma}\right| > \epsilon$ we obtain $|x_{ij}(t)| > \epsilon$ for infinitely many t and the proof is concluded. □

The exponential convergence assumption taken in the theorem above is a moderate condition that it can be dropped if the corresponding coupling weights are uniformly bounded away from zero. Then one is allowed to combine the uniform convergence of solutions and the linearity of the system to conclude on the exponential rate.

3.4. **Applications in non-linear systems.** The aim of this section is to apply the results of §3.2 in three different non-linear systems proposed in the literature.

The first model is a consensus scheme, with a non-linear coupling based on passivity and it is discussed under different contexts [26, 2, 36, 25]. We derive a simple convergence result with a direct linearization technique.

Next, we focus on the continuous time counterpart of the previously discussed model in Eq. (8) and in the related literature [9, 10, 40, 16, 17]. We weaken the assumptions beyond the asymmetric weight condition, to the simple static and switching connectivity and derive conditions on the initial values for convergence to asymptotic flocking.

Finally, motivated by [38], we introduce a generic non-linear perturbation to our non-linear system and inspect the condition to consensus by stability in variation and fixed point theory.

3.4.1. *Non-linear passive coupling.* A network of $N < \infty$ agents exchanges information according to the following algorithm:

$$\begin{cases} \dot{x}_i(t) = \sum_{j \in N_i} g_{ij}\big(t, x_j(t) - x_i(t)\big), & t \geq t_0 \\ \mathbf{x}(t_0) = \mathbf{x}^0 \end{cases} \tag{11}$$

for $i \in \mathcal{V}$. For any $t \geq t_0$ there may or may not exist a connection between j and i. This defines a connectivity regime that can be described by a graph $\mathbb{G}_g(t) = (\mathcal{V}, \mathcal{E}(t))$ with $(i,j) \in \mathcal{E}(t)$ if and only if $g_{ij}(t, \cdot) \neq 0$.

Let $\mathbb{G}_g(t)$ be the graph that is associated with existence or not of a connection between j and i. The passivity assumptions of $g_{ij}(\cdot)$ are summarized in the next statement:

Assumption 3.32. *For any $i, j \in \mathcal{V}$, $g_{ij}(t,x) : [t_0, \infty) \times \mathbb{R} \to \mathbb{R}$ is continuous in x and right-continuous in t and for $t \geq t_0$ it satisfies the following properties:*

1. $g_{ij}(\cdot, x) : [t_0, \infty) \to [0, m)$ *uniformly in x,*
2. $g_{ij}(t, 0) = 0$ *for any $t \geq t_0$,*
3. $g_{ij}(t, \cdot) \neq 0 \Rightarrow \frac{g_{ij}(t,x)}{x} > 0 \ \forall \ x \neq 0$,
4. $g_{ij}(t, \cdot) \neq 0 \Rightarrow \lim_{x \to 0} \frac{g(t,x)}{x} \in \mathbb{R}_+$ *independent of t.*

The form of g_{ij} incorporates two features of the consensus algorithms discussed so far. The first is that by construction g_{ij} are compatible with the previously discussed connectivity regimes (switching connectivity) and the second is the passivity property which makes the solutions to behave in a qualitative similar way to the linear case (9). For example, the boundedness of the solutions of Eq.(9) as explained in Lemma 3.21 is preserved in Eq.(11) under the passivity condition.

Lemma 3.33. *Under Assumption 3.32, the solution of Eq. (11) satisfies*

$$S\big(\mathbf{x}(t)\big) \leq S(\mathbf{x}^0).$$

Proof. Let $t^* > t_0$ be the first time that the solution $\mathbf{x}(t)$ of Eq. (11) escapes $[\min_i x_i^0, \max_i x_i^0]$. Then there exists $i \in \mathcal{V}$ such that $x_i(t^*) = \max_i x_i^0$ and $\dot{x}_i(t^*) > 0$, a contradiction according to Assumption 3.32. Similarly for $\min_i x_i^0$. □

Now we are ready to apply Theorem 3.23.

Corollary 3.34. *Consider the initial value problem (11) and let Assumption 3.32 hold. Define*

$$f(t) := \max_{x \in [0, S(\mathbf{x}^0)]} \frac{g(t,x)}{x}.$$

Unconditional asymptotic consensus of Eq. (11) is achieved if

1. $\mathbb{G}_g$ *is static and* $\int^{\infty} f(s)ds = \infty$
2. $\mathbb{G}_g(t)$ *is time varying the f satisfies one of the properties of Theorem 3.23.*

Proof. The passivity assumption obviously ensures that the solution of (11) exists in the large. Let $\mathbf{x}(t, t_0, \mathbf{x}^0) \in \mathbb{R}^N$ an arbitrary but fixed solution of Eq. (11), we define

$$a_{ij}(t) := \frac{g_{ij}(t, x_j(t) - x_i(t))}{x_j(t) - x_i(t)}$$

and effectively rewrite the initial value problem (11) as

$$\dot{y}_i(t) = \sum_i a_{ij}(t)\big(y_j(t) - y_i(t)\big)$$
$$y_i(0) = x_i^0$$

so that the solutions $\mathbf{y}$ and $\mathbf{x}$ are indistinguishable. Thus one can study the behavior of $\mathbf{y}$ to conclude about $\mathbf{x}$.[6] Under Assumption 3.32 we see that $a_{ij}(t)$ satisfies

$$f(t) := \min_{x \in [0, S(\mathbf{x}^0)]} \frac{g(t,x)}{x} \leq a_{ij}(t) \leq \max_{x \in [0, S(\mathbf{x}^0)]} \frac{g(t,x)}{x}$$

and that by the continuity assumptions the right hand-side is in turn uniformly bounded from above. Then all the results of the previous section can be applied for the new system, accordingly. □

The form $g(t, x_i - x_j)$ is quite general. Apart from the standard Kuramoto setting where $g_{ij}(x) = \sin x$ [26] it can also include the opinion dynamics framework proposed by Krause [25] with $g_{ij}(\mathbf{x}) = a(|x_i - x_j|)(x_j - x_i)$.

The authors acknowledge that the actual difficulty in the latter model lies in the fact that the dynamics are deployed in a proximity graph without g_{ij}'s being uniformly bounded from below and the dynamics a_{ij} are non-linear and autonomous. A definitive analysis of Krause's dynamics is yet to be achieved since it is precisely the derivation of an appropriate Lyapunov functional a quite challenging problem.

3.4.2. *Flocking models.* Also known as second order consensus systems these algorithms were proposed for modeling the velocity coordination and flock formation populations of birds, whenever the latter were seen as autonomous individuals. A population of $N < \infty$ birds is the collection of the values (x_i, u_i) where $i = 1, \dots, N$. Each bird i evolves its position x_i and its speed u_i according to

$$\begin{cases} \dot{x}_i(t) &= u_i(t) \\ \dot{u}_i(t) &= \sum_{j \in N_i} a_{ij}(\mathbf{x})\big(u_j(t) - u_i(t)\big), \qquad t \geq 0 \end{cases} \tag{12}$$

for $i \in \mathcal{V}$ and given vectors of initial position $\mathbf{x}^0 = \mathbf{x}(0)$ and velocity $\mathbf{u}^0 = \mathbf{u}(0)$. In flocking dynamics the connectivity weights a_{ij} are state dependent making the algorithm essentially non-linear. The first models assumed symmetric coupling effects $a_{ij} = a_{ji}$ to be decreasing functions of the relative distance between i and j. More specifically these rates were also assumed to have an explicit closed form. [9, 10]. Since then, a number of generalizations improved this restriction to the case of asymmetric weights preferring topologically defined relative distances and not only Euclidean [40]. This means that a_{ij} is not necessarily a function only of x_i, x_j but it may depend on the positions of the whole group, i.e. $a_{ij} = a_{ij}(\mathbf{x})$. The derived results, so far demand the restricting condition $\sum_i a_{ij} < 1$ and increased (scrambling) connectivity) in the sense of Theorem 3.23. In this work, for a solution $(\mathbf{x}, \mathbf{u})$ of Eq. (12) we take

$$a_{ij}(\mathbf{x}(t)) \geq f(S(\mathbf{x}(t)))$$

with the property that $f(s) \to 0$ as $s \to \infty$ and $\sum_j a_{ij} \leq (N-1)f(0) =: m < \infty$. The difficulty of these systems lies on the fact that a_{ij} are not uniformly bounded

[6] This is known in the literature as a direct linearization technique (see for example [5]).

from below. The problem is to derive sufficient conditions on initial position and velocity 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

$$\lim_{t} S\big(\mathbf{u}(t)\big) = 0 \quad \& \quad \sup_{t\geq 0} S\big(\mathbf{x}(t)\big) < \infty \tag{13}$$

i.e. the asymptotic flocking condition. For the next result the Assumptions of Theorem 3.28 hold.

Theorem 3.35. *Consider the system* (12) *and its solution* $\big(\mathbf{x}(t), \mathbf{u}(t)\big) \in \mathbb{R}^N \times \mathbb{R}^N$. *The following conditions hold:*

1. *Static scrambling connectivity. The solution exhibits asymptotic flocking if*
$$S(\mathbf{u}^0) < \int_{S(\mathbf{x}^0)}^{\infty} f(w)dw \tag{14}$$
2. *Static routed-out branching connectivity. The solution exhibits asymptotic flocking if*
$$S(\mathbf{u}^0) < \frac{(1-e^{-mB})^\gamma}{m^\gamma \gamma B} \int_{P^{\gamma,B}_{\mathbf{x}^0,\mathbf{u}^0}}^{\infty} f^\gamma(s)ds \tag{15}$$
where $P^{\gamma,B}_{\mathbf{x}^0,\mathbf{u}^0} = \max\big\{S(\mathbf{x}^0), |S(\mathbf{x}^0) - S(\mathbf{u}^0)\gamma B|\big\}$, $m = \max_i |N_i| f(0)$.
3. *Switching connectivity. The solution exhibits asymptotic flocking if*
$$S(\mathbf{u}^0) < \frac{(1-e^{-m\epsilon})^\sigma}{m^\sigma \sigma B} \int_{P^{\sigma,B}_{\mathbf{x}^0,\mathbf{u}^0}}^{\infty} f^\sigma(s)ds \tag{16}$$
where $\sigma = l^*([N/2]+1)$, l^* *with the meaning of Remark 2.6 and* $\epsilon > 0$ *with the meaning of Assumption 3.27.*

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 3.23 for $\mathbf{u}$ and show that

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

so that asymptotic flocking will occur with exponential rate of convergence if $S\big(\mathbf{x}(t)\big) \leq r$ for some $r > 0$. For this, we follow [17, 40] and introduce the functional

$$V_1(\mathbf{x}, \mathbf{u}) = S(\mathbf{u}) + \int_0^{S(\mathbf{x})} f(w)dw \tag{17}$$

so that along a solution of (12) $\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ where $\mathbf{u}(t) = \dot{\mathbf{x}}(t)$ we have

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

then

$$V_1(t) \leq V_1(0) \Leftrightarrow S(\mathbf{u}(t)) + \int_0^{S(\mathbf{x}(t))} f(w)dw \leq S(\mathbf{u}^0) + \int_0^{S(\mathbf{x}^0)} f(w)dw$$

From the imposed condition Eq. (14) 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 $S\big(\mathbf{x}(t)\big) \geq S\big(\mathbf{x}^0\big)$

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

which makes sense if $S\big(\mathbf{x}(t)\big) \leq r'$. Pick $r = \max\{r', S(\mathbf{x}^0)\}$ to conclude that condition (14) 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

$$W(\mathbf{x}(s)) = mI - D(\mathbf{x}(s)) + A(\mathbf{x}(s))$$

and next

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

Finally for the scrambling index γ of the topological graph $\mathbb{G}_{P(\mathbf{x}(t))}$ (which is independent of time)

$$P_B^{(\gamma)}\big(\mathbf{x}(t)\big) = \int_{t-B}^{t}\int_{s_1-B}^{s_1}\int_{s_2-B}^{s_2}\cdots\int_{s_{\gamma-1}-B}^{s_{\gamma-1}} C(t,s_1)C(s_1,s_2)\cdots C(s_{\gamma-1},s_\gamma)ds_\gamma\dots ds_1$$

which is stochastic from Proposition 3.25 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 3.23

$$S\big(\mathbf{u}(t)\big) \leq \tau\big(W_B^{(\gamma)}(\mathbf{x}(t))\big)S\big(\mathbf{u}(t-\gamma B)\big) \leq \Big(1 - cf^\gamma\big(S\big(\mathbf{x}(t)\big)\big)\Big)S\big(\mathbf{u}(t-\gamma B)\big) \tag{18}$$

with $c := \frac{(1-e^{-mB})^\gamma}{m^\gamma}$ and $S\big(\mathbf{x}(t)\big) \geq r$ 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. (12), $\big(\mathbf{x}(t),\mathbf{u}(t)\big)$ is

$$\begin{aligned}\dot{V}_2(t) &= S\big(\mathbf{u}(t)\big) - S\big(\mathbf{u}(t-\gamma B)\big) + cf^\gamma\big(S(\mathbf{x}(t))\big)S\big(\mathbf{u}(t)\big)\\ &\leq \big(1 - cf^\gamma(S\mathbf{x}(t))\big)S\big(\mathbf{u}(t-\gamma B)\big) - S\big(\mathbf{u}(t-\gamma B)\big) + cf^\gamma\big(S(\mathbf{x}(t))\big)S\big(\mathbf{u}(t-\gamma B)\big)\\ &\leq 0\end{aligned}$$

in view of Lemma 3.21 (from which it is deduced that $S\big(\mathbf{u}(t)\big) \leq S\big(\mathbf{u}(t-\gamma B)\big), \forall t$. Then for $t \geq \gamma B$

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

Let the following condition hold

$$\int_0^{\gamma B} S\big(\mathbf{u}(s)\big)ds = c\int_{S(\mathbf{x}(\gamma B))}^{\infty} f^\gamma(s)ds \tag{19}$$

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, $S\big(\mathbf{x}(t)\big) \leq S\big(\mathbf{x}(\gamma B)\big)$ implies

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

so that the flock remains bounded and exponential speed alignment is ensured. Finally, we show that Eq. (15) implies Eq. (19), Indeed,

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

from Lemma 3.21. Now we look for a lower bound of $S\big(\mathbf{x}(t)\big)$. If $S\big(\mathbf{x}(t)\big) \geq S(\mathbf{x}^0)$ from the form of Eq. (12) the rate at which $S\big(\mathbf{x}(t)\big)$ may shrink can be deduced from the extreme scenario of $\mathbf{x}_0 = (x^0, 0, \dots, 0)$, $x^0 \neq 0$ so that $S(\mathbf{x}^0) = x^0$ and $\mathbf{u}^0 = (u^0, 0, \dots, 0)$, $u^0 \neq 0$ 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 $t = \gamma B$

$$S\big(\mathbf{x}(\gamma B)\big) \geq \max\left\{S(\mathbf{x}^0), |S(\mathbf{x}^0) - S(\mathbf{u}^0)\gamma B|\right\} = P^{\gamma,B}_{\mathbf{x}^0,\mathbf{u}^0}$$

so that

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

then

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

The case of switching connectivity is treated as in Theorem 3.23 and the use of V_2 after substituting γ with σ. Then Eq. (16) substitutes Eq. (15) to ensure asymptotic flocking. □

Remark 3.36. In the case of static connectivity, $\gamma = 1$ implies that Eq. (14) and Eq. (15) coincide as $B \downarrow 0$.

3.4.3. *A fully-nonlinear model.* The sign of the coupling weights $a_{ij}(t)$ is instrumental in the analysis. All the reviewed literature of §1 discusses models which with one way or another assume this condition. Adopting the terminology of [43], condition $a_{ij} \geq 0, \forall i, j$ classifies (9) as a *cooperative* system, while $a_{ij} \leq 0$ as *competitive* system. Although the literature in consensus algorithms as cooperative systems is nearly complete, to the best of our knowledge there are no results on multi-agent competitive systems. As it is explained in the general theory of monotone dynamical systems [43], competitive dynamical systems exhibit a much richer behavior than the cooperative ones. As such analysis is far beyond the scope of this paper, in this section we will consider a non-linear perturbation of Eq. (9) that sustains consensus solutions. Then by stability in variation we will derive sufficient conditions under which it can behave as a co-operative one, converging to a common constant value.

For $N < \infty$, a population of N agents exchanges values according to the initial value problem:

$$\begin{cases} \dot{x}_i = \sum_{j \in N_i} a_{ij}(t)\big(x_j(t) - x_i(t)\big) + g_{ij}(t, x_j) - g_{ij}(t, x_i), & t \geq t_0 \\ x_i(t_0) = x_i^0 \end{cases} \tag{20}$$

for $i \in \mathcal{V}$. We study the asymptotic behavior of Eq. (20) by approximating it with (9). Indeed for

$$\mathbf{G}(t, \mathbf{x}) = \left[\sum_j g_{1j}(t, x_j) - g_{1j}(t, x_1), \dots, \sum_j g_{Nj}(t, x_j) - g_{Nj}(t, x_N)\right]^T$$

we rewrite (20) as

$$\dot{\mathbf{x}} = -L(t)\mathbf{x} + \mathbf{G}(t, \mathbf{x})$$

where $L(t) = D(t) - A(t)$ is the Laplacian matrix.

The variation of constants formula implies that the solution $\mathbf{x}$ of (20) satisfies

$$\mathbf{x}(t) = \Phi(t,t_0)\mathbf{x}^0 + \int_{t_0}^{t} \Phi(t,s)\mathbf{G}(s,\mathbf{x}(s))ds \tag{21}$$

Now, $\mathbf{y}(t) = \Phi(t,t_0)\mathbf{x}^0$ is the solution of (9), for which we know that under the Assumptions 3.9 and 3.13 $\Phi(t,t_0)\mathbf{x}^0 \to \boldsymbol{\alpha}^T\mathbf{x}^0$ for some $\boldsymbol{\alpha} \in \mathbb{R}^N$ with the properties that $\alpha_i \geq 0$, $\sum_i \alpha_i = 1$. In terms of the semi-norm, there exists a bounded rate function $h(t,t_0)$ such that $\lim_t h(t,t_0) = 0$ for any fixed t_0 and

$$S\big(\mathbf{y}(t)\big) = S\big(\Phi(t,t_0)\mathbf{x}^0\big) \leq h(t,t_0)S(\mathbf{x}^0) \tag{22}$$

For the sake of simplicity we will also assume that

$$\int^{\infty} h(t,t_0)dt < \infty \tag{23}$$

for the unperturbed system. Next we will impose conditions on g_{ij}. There is an abundance of conditions one can impose on the functions $g_{ij}(t,x)$ to proceed and for the sake of simplicity we will assume a global lipschitz condition.

Assumption 3.37. *For any $t \geq t_0$ it holds that:*

$$|g_{ij}(t,x_1) - g_{ij}(t,x_2)| \leq k_{ij}(t)|x_1 - x_2|$$

for some integrable bounded function $k_{ij}(t)$.

Under the preceding assumption we obtain $S\big(\mathbf{G}(t,\mathbf{x})\big) \leq k(t)S\big(\mathbf{x}\big)$ for

$$k(t) = \max_{i,j} \sum_l k_{il}(t) + k_{jl}(t)$$

Proposition 3.38. *The solution $\mathbf{x}(t)$ (20), satisfies*

$$S\big(\mathbf{x}(t)\big) \leq h(t,t_0)S(\mathbf{x}^0) + \int_{t_0}^{t} h(t,s)k(s)S\big(\mathbf{x}(s)\big)ds$$

Proof. Let $\mathbf{y}$ be the solution of $\dot{\mathbf{y}} = -L(t)\mathbf{y}$. Then a variation of parameters of the functional $S(\mathbf{x})$, yields the form

$$S\big(\mathbf{x}(t)\big) = S\big(\mathbf{y}(t)\big) + \int_{t_0}^{t} \frac{\partial}{\partial \mathbf{x}} S\big(\mathbf{y}(t,s,\mathbf{x}(s))\big)^T \mathbf{G}\big(s,\mathbf{x}(s)\big)ds$$

(see also Theorem 1.3.1 of [27]) and the result follows from the bound on $S\big(\mathbf{y}(t)\big)$ and Assumption 3.37. □

From the Proposition above we obtain that $S\big(\mathbf{x}(t)\big)$ is in fact upper bounded by $q(t)$, which satisfies the integral equation

$$q(t) = h(t,t_0)q(t_0) + \int_{t_0}^{t} h(t,s)k(s)q(s)ds, \qquad q(t_0) = S(\mathbf{x}^0) \tag{24}$$

Theorem 3.39. *Consider Eq. (9) with its solution to satisfy (22) and also let condition (23) hold. Under Assumption 3.37 the solution of (20) converges to a consensus value as fast as $\frac{1}{w(t)}$ if:*

$$\sup_{t \geq t_0} w(t)h(t,t_0) < \infty, \qquad \textit{and} \qquad \sup_{t \geq t_0} w(t) \int_{t_0}^{t} h(t,s)\frac{k(s)}{w(s)}ds \leq \alpha < 1$$

Proof. From the discussion above the stability of (20) to a consensus value is reduced to the stability to zero of the integral equation (24). For the latter, we will build a fixed point theory argument. Recall the discussion in §2.5 and consider the space

$$\mathbb{M} = \{y \in \mathbb{B} : y(t^0) = S(\mathbf{x}^0), \sup_{t \geq t_0} w(t)|y(t)| < \infty\}$$

which together with the weighted metric $\rho(y_1, y_2) = \sup_{t \geq t_0} w(t)|y_1(t) - y_2(t)|$ constitute a complete metric space [5]. In this space we will apply the Contraction Mapping Principle (Theorem 2.7) as follows: Define the operator

$$(\mathcal{Q}y)(t) := \begin{cases} S(\mathbf{x}^0), & t = t_0 \\ h(t,t_0)S(\mathbf{x}^0) + \int_{t_0}^t h(t,s)k(s)y(s)ds, & t \geq t_0 \end{cases}$$

and note that under for any $y \in \mathbb{M}$, $(\mathcal{Q}y)(t) \to 0$ as the first term vanishes in view of the unperturbed system to satisfy the conditions of Theorem 3.28 and the second term vanishes as the convolution of an L^1 function with a function that goes to zero. The same holds for the weighted quantity $w(t)|(\mathcal{Q}y)(t)|$ in view of the imposed conditions. Finally, it is easy to see that $\mathcal{Q}$ is a contraction in $(\mathbb{M}, \rho)$ since

$$\rho(\mathcal{Q}y_1, \mathcal{Q}y_2) \leq \sup_{t \geq t_0} w(t) \int_{t_0}^t h(t,s)\frac{k(s)}{w(s)} ds \rho(y_1, y_2) \leq \alpha\rho(y_1, y_2).$$

Hence, by the contraction mapping principle $\mathcal{Q}$ attains a unique fixed point in $\mathbb{M}$ and the proof is concluded. □

We conclude this section with an illustrating example.

Example 3.40. Consider the non-linear system

$$\dot{x} = a(t)(y-x) + \kappa_1(t)\left(\frac{\sin y}{1+y^2} - \frac{\sin x}{1+x^2}\right)$$
$$\dot{y} = b(t)(x-y) + \kappa_2(t)\left(\frac{\cos x}{1+x^4} - \frac{\cos y}{1+y^4}\right)$$

for $t \geq 0$, $x(0) = x^0, y(0) = y^0$, $a(t), b(t) \geq 0$. At first, we observe that for $\kappa_1 = \kappa_2 \equiv 0$ and $\mathbf{x}(t) = \big(x(t), y(t)\big)$

$$S(\mathbf{x}(t)) \leq e^{-\int_0^t a(s)+b(s)ds} S(\mathbf{x}^0)$$

and convergence occurs if for $f(t,0) = \int_0^t a(s) + b(s)ds$, with $f(t) \to \infty$. Define a function w such that $\sup_t w(t)e^{-f(t)} < \infty$. Then basic algebra yields approximations of the Lipschitz constants $k_1(t) = \kappa_1(t)$ and $k_2(t) = 1.231\kappa_2(t)$ and $\kappa(t) = \kappa_1(t) + \kappa_2(t)$. By a simple continuity argument, requiring

$$\sup_t \int_0^t f(t,s)\kappa(s)ds < 1,$$

one can always find w to satisfy the conditions of Theorem 3.39.

As a numerical application take $a(t) = t, b(t) = 1.1 + \sin t$, $k_1(t) = \sin t^2$, $k_2(t) = 1.231\sqrt{t}$ and $w(t) = e^{0.05t}$. Then it can be calculated (using for example MAPLE) the bound

$$0 \leq \sup_{t \geq 0} \frac{\sin t^2 + 1.231\sqrt{t}}{1.1 + t + \sin t - 0.05} \leq 0.9$$

so that

$$\int_0^t e^{-\int_s^t a(w)+b(w)dw} \leq 0.8 \int_0^t e^{-\int_s^t a(w)+b(w)-0.05dw}\big(a(s) + b(s) - 0.05\big)ds \leq 0.9$$

so that Theorem 3.39 applies.

4. **Stochastic Consensus.** As already discussed in §1.1.2, the stochastic part in consensus systems was initially implemented for the purpose of modeling the uncertainty in the inter-connection regime among agents. In the mathematical world, this statistical regularity smooths out the, characterized by several researchers as stringent, condition of recurrent connectivity. Imposing mild statistical regularity on the dynamics of inter-connections, the assumption of connectivity over uniformly bounded intervals of time is satisfied almost surely and it can be essentially omitted.

The purpose of this section is to re-formulate the consensus problem with emphasis on discrete time, modeling the communication topology in a measure theoretic framework. For this we recall the discussion §2.3. We will show that, our setting is general enough to unify many results proposed in the literature. We propose a dynamical shift, T, that generates stochastic matrices so that asymptotic consensus is an event with a probability induced by the measure that preserves T. If in addition the shift T is ergodic then any T-invariant event is of either full or zero measure. In the examples subsection we will review results of the literature that are based on the close relationship of our framework to stationary stochastic processes.

Similarly to the deterministic case we separate the existence of a connection among agents from the weight of this connection, exactly because we are under the non-uniform lower bound condition. It is this condition that plays a critical role in the analysis of the system. In fact, unless the probabilistic regime concerning the connection failures is trivial, asymptotic consensus is never guaranteed in full probability whenever the weights are free to vanish.

4.1. **Topology driven by measure preserving dynamical systems.** For $N < \infty$ we define the discrete set $\mathbb{Y}$ of all possible (directed) connections among N nodes. The cardinality of $\mathbb{Y}$ is finite. Then we define the product measure space $(\mathbb{X}, \mathcal{B}, \mu) = \prod_{t\geq 0}(\mathbb{Y}, 2^{\mathbb{Y}}, m)$ for some measure function m and the induced product measure as it was discussed in §2.3. The shift operator $T : \mathbb{X} \to \mathbb{X}$ is a measure preserving transformation since for any $A \in \mathcal{J}$, $\mu(T^{-1}A) = \mu(A)$ (see also [49]). We understand χ_i as an $N \times N$, $0-1$ matrix with all diagonals zero and the off-diagonal values to attain value 1 if there is a connection from the agent of the column to the agent of the row, otherwise attain the zero value, as well.

4.1.1. *Discrete time.* Recall the system (5) and its solution

$$\mathbf{x}(t) = P(t-1)P(t-2)\ldots P(0)\mathbf{x}(0) \tag{25}$$

for $\mathbb{G}_{P(t)} \in \mathcal{S}$. We would want the steering force that generates the matrices $P(t)$ at every instant t, to be the shift $T : \mathbb{X} \to \mathbb{X}$. Let the family of functions $\{a_{ij}(t)\}_{i\neq j}$, so that for $t \geq 0$ and any $i \neq j \in \mathcal{V}$, $a_{ij}(t) \in [f(t), \infty)$ for some fixed non-increasing positive function f that vanishes as $t \to \infty$. Let the stochastic matrix

$$P(t) = \phi(T^t\chi) \tag{26}$$

to be defined through the following measurable function $\phi : \mathbb{X} \to \mathcal{S}$:

$$[\phi(T^t\chi)_{ij}] = \begin{cases} \frac{a_{ij}(t)}{\varepsilon\sum_{j:[\{T^tx\}_{ij}]=1} a_{ij}(t)} & \text{if } \left[\{T^tx\}_{ij}\right] = 1 \text{ and } i \neq j \\ 0 & \text{if } [\{T^tx\}_{ij}] = 0 \text{ and } i \neq j \end{cases}$$

for some fixed $\varepsilon > 1$ so that $[\phi(T^t\chi)_{ii}] := 1 - \sum_j[\phi(T^t\chi)_{ij}]$. We are interested in the set $P_B \in \mathcal{B}$ as follows:

$$P_B = \left\{\chi \in \mathbb{X} : \mathbb{G}_{P_{t,B}} \text{ is routed-out branching } \forall t \geq 0\right\}.$$

Due to the uncertainty of connectivity we chose to state a slightly different, from the deterministic, version on the weights. We re-scaled the connectivity weights, in order to preserve the stochastic structure for $P(t)$, regardless of the probabilistic generator that controls the existence of connections. It should be noted that $P(t)$ is in this case not necessarily symmetric.

The setting clearly proposes that the solution $\mathbf{x}(t)$ is a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the consensus problem becomes equivalent to the convergence of $\mathbf{x}(t)$ to Δ. We are basically interested in convergence in the almost sure sense.[7]

Consensus may be achieved if μ assigns a positive value to P_B for some B and in particular, it is this value of $\mu(P_B)$ the probability with which consensus occurs, in exactly the same way as in Theorem 3.14 and therefore we can readily state the following result without proof:

Theorem 4.1. *Let $(\mathbb{X}, \mathcal{B}, \mu)$ be the direct product measure space on products of stochastic matrices and $T : \mathbb{X} \to \mathbb{X}$ a shift. Consider the system (5) and the solution given by (25) with $P(t)$ as in (26) and also consider the set P_B. The solution $\mathbf{x}(t)$ of (5) converges to Δ with probability $\mu(P_B)$ if f satisfies either*

$$\sum_t f^\sigma(t\sigma) = \infty,$$

or

$$0 < \sup_{t\geq 1} \frac{\sum_{i=t\sigma}^{(t+1)\sigma-1} f^\sigma(t)}{f^\sigma((t+1)\sigma - 1)} < \infty \text{ and } \sum_{t\geq 1} f^\sigma(t) = \infty$$

where $\sigma = l^*([N/2] + 1)B$ and l^* with the meaning of Remark 2.6.

This theorem is simply the measure theoretic analogue of Theorem 3.14 and little does it contribute to our discussion. It illustrates, however, the interdependence between the non-uniform lower bounds of a_{ij}, the induced statistical regularity and it is only of theoretical interest. Almost sure convergence is ensured if the event $\bigcup_{B\geq 1} P_B$ is of full measure. The most common processes in the literature (e.g. i.i.d, markov or stationary) obey probability laws that are invariant in time and they yield almost sure consensus only under the uniform bound condition (i.e. $a_{ij}(t) \neq 0 \Rightarrow a_{ij}(t) \geq \delta > 0$). It is exactly this case where there is no difference between the existence of connection and its weight, when one studies the asymptotic convergence to Δ.

For this reason, in the rest of this section we will strengthen to $a_{ij}(t) \in \{0\} \cup (0, 1)$ uniformly in t so that we can focus on the processes, produced by the shift T, which guarantee the asymptotic behavior of $P_{0,t}$ to a rank-1 matrix.

Corollary 4.2. *Let $T : \mathbb{X} \to \mathbb{X}$ be an ergodic shift on the product space $(\mathbb{X}, \mathcal{B}, \mu)$, $P(t)$ with the form of Eq. (26) and $a_{ij}(t) \in \{0\} \cup (0, 1)$ for $i \neq j$. Then the solution $\mathbf{x}$ of (5) converges to consensus with probability one if $\mu(P_B) > 0$ for some $B > 0$.*

Proof. At first we show that the set $W = \bigcup_B P_B$ is T-invariant. Indeed, for fixed $B > 0$ and any $\chi \in P_B$, we have $T\chi \in P_B$, $P_B \subset T^{-1}P_B$ and this is true over the union, i.e. $W \subset T^{-1}W$. Next it is easy to see that $T^{-1}P_B \subset W_{B+1} \subset W$ which finally means that $T^{-1}W = W$. The ergodicity condition makes T an indecomposable transformation on T invariant sets, i.e. $\mu(W) = 0$ or $\mu(W) = 1$ but the first case is excluded because $\mu(W) \geq \mu(P_B) > 0$. Then the only realization of

[7] Since $\mathbb{P}\big(|\mathbf{x}(t)| \leq N \max_i x_i(0)\big) = 1$ we have $E[|\mathbf{x}(t)|^r <]\infty$ and from these two facts we have that almost sure convergence implies convergence in the r^{th} mean for any $r \geq 0$.

shifting over $\mathbb{X}$ is this concerning processes with routed-out branching graphs over B intervals for some $B < \infty$. Any other event occurs with zero probability and the result follows in view of the uniformly bounded weights. □

It should be noted here that $P(t)$ is *not* a stationary process as by construction the measure μ does not concern the weights $a_{ij}(t)$. The stationarity property can be observed in the quantity $\mathbb{G}_{P(t)}$ which as we mentioned above is the only key feature for the stability analysis.

Example 4.3 (Stationary Ergodic processes [46]). The problem of consensus over stationary ergodic processes assumes that the matrix $P(t)$ is essentially such a process. It is very well known that a measure preserving shift can be used to generate stationary processes and, conversely, that any stationary process is equal (in distribution) to a process generated by a measure preserving shift [24]. Given a stationary ergodic process one can easily verify whether this particular shift is ergodic after applying Birkhoffs ergodic theorem: If $\frac{1}{n}\sum_{t=0}^{n-1}\mathbf{1}_{W_B}(P_{t,B})$[8] is positive then the shift is ergodic and consensus is proved in the almost sure sense. Thus, Corollary 4.2 reproduces the results of [46] but in a broader setting as not only does it allow for connectivity over B intervals of time, but it is also not concerned with the stationarity of the weighted graph. It exclusively describes the existence of a connection and not the strength of it.

Remark 4.4. The approach above can of course include stationary processes that occur from deterministic systems which exhibit a non-trivial stochastic behavior, such as chaotic maps or non-linear differential equations, so long as their solutions produce a (natural) invariant measure on the state space (see [23]). Then one can read these dynamics as stochastic and consider the consensus problem with communication topology driven by chaotic signals.

Example 4.5 (IID processes [20], [45]). One of the first works on the topic of probabilistic consensus in [20], formulated Eq. (5), as a stochastic linear equation with symmetric connectivity weights ($a_{ij} = a_{ji}$) to randomly take values at each time $t \in \mathbb{N}$. The partition of interest would be $a_{ij}(t) \neq 0$ with probability p and $a_{ij}(t) = 0$ with probability $1-p$, independently of the rest of the connections and times.

Let us digress for a moment and see $P(t) = P_t(y)$ as a random process defined on a probability space $(\mathbb{Y}, 2^{\mathbb{Y}}, \mathbb{P})$. Then $P(t)$ takes values in the space of stochastic matrices with positive diagonals and uniformly bounded weights. Then the backward product $P_{t,B}(y)$ is a homogeneous sequence of independent random trails and it forms a stationary process. By the independence assumption it is easy to directly calculate the probability of the event the corresponding graph $\mathbb{G}_{P_{t,B}}$ to be routed-out branching: If p is the probability that $a_{ij}(t) \neq 0$ then the probability of j affecting i through a B time interval is by the binomial theorem $1-(1-p)^B$. For $\mathbb{G}$ a graph on N nodes, let $q \in [1, (N-1)^2)$ denote the minimal number of edges so that each additional edge will keep $\mathbb{G}$ routed-out branching. Then for P_B as defined

[8] Here $\mathbf{1}_s(A)$ is a dual function that takes value 1 if $s \in A$ and 0 otherwise

before

$$\begin{aligned}\mathbb{P}(P_B) &> \sum_{l=q}^{(N-1)^2} \binom{(N-1)^2}{l} \big(1-(1-p)^B\big)^l (1-p)^{B((N-1)^2-l)} \\ &= 1 - \sum_{l=0}^{q-1} \binom{(N-1)^2}{l} \big(1-(1-p)^B\big)^l (1-p)^{B((N-1)^2-l)} \\ &= 1 - \mathcal{O}\big((1-p)^B\big) \to 1, \text{ as } B >> 1\end{aligned}$$

To see why the event $E = \big\{ \sup_B \mathbb{G}_{P_{t,B}} \text{is not connected}, \forall t \geq 0 \big\}$ is a zero probability event, note that P_B are nested for B decreasing and for this reason $\mathbb{P}(E^c) = \lim_{B\to\infty} \mathbb{P}(P_B)$.

To adapt this example to our framework we work as follows. Let the set $\{0,1\}$ and $(p, 1-p)$ the probability vector for some fixed $p \in (0,1)$, so that $\{0\}$ is assigned to $1-p$ and $\{1\}$ is assigned to p. This is an elementary measure space. On this space, we define the triplet $(\mathbb{Y}, 2^{\mathbb{Y}}, m)$ over $(N-1)^2$ pairs of nodes (i.e. without self-connections) each fixed pair of which will be considered connected and take values in an open subset of $[0,1]$ with probability p or it will be zero with probability $1-p$, independently of the rest of the pairs. Eventually, $(\mathbb{X}, \mathcal{B}, \mu) = \prod_{j=0}^{\infty} (\mathbb{Y}, 2^{\mathbb{Y}}, m)$ is the product space of interest on which the shift $T : \mathbb{X} \to \mathbb{X}$ is defined, as $T(\chi_0\chi_1\chi_2\dots) = \chi_1\chi_2\dots$. If $\mathcal{J}$ is the semi-algebra of all measurable rectangles then $\mu(T^{-1}A) = \mu(A)$ for any $A \in \mathcal{J}$ and by Theorem 1.1 of [49], T is measure preserving. It is a standard exercise to show that T is ergodic [49]. It is only left to show that for some $B > 0$, $\mu(P_B) > 0$, a calculation very similar to the one carried before and Corollary 4.2 applies.

Example 4.6 (Markov processes [31])**.** The authors considered Eq. 5 with a switching communication topology driven by a Markovian jump process and in particular a process on a homogeneous Markov chain over l states defined by a stochastic matrix Z, each state of which, corresponds to a connectivity regime among N nodes. The result is summarized as follows: Unconditional asymptotic consensus is achieved if and only if Z is irreducible and the union of states of the chain correspond to a routed-out branching graph. We note that the irreducibility of Z implies the existence of an invariant measure $\boldsymbol{\pi} \in \mathbb{R}^l > 0$ with $\sum_i \pi_i = 1$ with the property that $\boldsymbol{\pi}^T Z = \boldsymbol{\pi}$. In the shift oriented framework, we have a transformation T on $(\boldsymbol{\pi}, Z)$ known as Markov shift which is ergodic if and only if Z is irreducible [49]. Then the event of connectivity over a B-interval of times is dictated by the invariant measure to be of positive measure and Corollary 4.2 applies.

4.1.2. *Continuous time.* The results of the previous section can be modified to deal with the problem in continuous time and there are numerous different settings to choose upon. Let us recall the deterministic case and the system (9). The stability of its solution with respect to Δ is decided upon the product of the matrices $P_B^{\gamma}(t)$. Just as in the discrete time case, the stochastic nature is implemented exclusively to model the communication failure. Taking into account the necessary dwelling time condition (Assumption 3.27) we will use the same measure preserving shift on the same product space which will operate every $\epsilon > 0$ time: More specifically, consider continuous deterministic functions a_{ij} that satisfy Assumptions 3.20 and $m := \sup_{t\geq t_0} \max_i \sum_j a_{ij}(t) < \infty$. For fixed $\epsilon > 0$, $q \in \mathbb{N}$ and $t \in [t_0 + (q-1)\epsilon, t_0 + q\epsilon)$, we define the $m-$stochastic matrix $W(t) = mI - D(t) + A(t)$ with

$$[W_{ij}(t)] = \begin{cases} a_{ij}(t)[\{T^q\chi\}_{ij}], & i \neq j \\ m - \sum_j w_{ij}(t), & i = j \end{cases}$$

where $T : \mathbb{X} \to \mathbb{X}$ is the measurable transformation defined in the preceding section. Consequently the matrices $P_B^{(l)}(t) : \mathbb{X} \to \mathcal{S}$ from Proposition 3.25 are well-defined processes as measurable mappings. Clearly, the properties of T reflect the properties of $P_B^{(l)}(t)$ and hence Theorem 3.28 as well as the non-linear results of §3.4.1 and §3.4.3 can be restated in their probabilistic version.

Example 4.7. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and for $\omega \in \Omega$, $\mathbf{x}(t,\omega) \in \mathbb{R}^N \times \mathcal{F}$ a random process such that $\mathbf{x}(t_0) = \mathbf{x}^0$, it is sample continuous, product measurable, it has a sample right derivative and it is a solution of the stochastic differential equation

$$\dot{x}_i(t,\omega) = \sum_j a_{ij}(t,\omega)\big(x_j(t,\omega) - x_i(t,\omega)\big), \qquad x_i(t_0) = x_i^0$$

if it satisfies this equation with probability one for all $t \geq t_0$. The stochastic part of this equation lies in $a_{ij}(t,\omega)$ which are assumed to be stochastic processes generated by the shift $T : \mathbb{X} \to \mathbb{X}$ and in particular we assume to be just as the one described in the Example 4.5. Then we are interested in the integral

$$\int_{t_0}^{\infty} f(s,\omega)ds$$

where $f(t,\omega) = \min_{i,j} \sum_l \min\{a_{il}(t,\omega), a_{jl}(t,\omega)\}$, since

$$\mathbb{P}\Big(\omega \in \Omega : \lim_{t\to\infty} S\big(\mathbf{x}(t,\omega)\big) \neq 0\Big) = 1 - \mathbb{P}\Big(\omega \in \Omega : \int_{t_0}^{\infty} f(s,\omega)ds = \infty\Big)$$

and it can be easily calculated that at every $I_q = [t_0 + (q-1)\epsilon, t_0 + q\epsilon)$,

$$\mu\Big(x \in \mathbb{X} : \mathbb{G}_{P(t)} \text{ scrambling}, t \in I_q\Big) > p^{N-1} > 0$$

and by the non-summability and independence of the above events, the sum over q diverges and the Borel-Cantelli Lemma assures that $P(t)$ will be scrambling for infinitely many ϵ intervals of time. Hence $\mathbb{P}\big(\omega \in \Omega : \int_{t_0}^{\infty} f(s,\omega)ds = \infty\big) = 1$ and almost sure asymptotic consensus occurs.

4.2. **Noisy flocking dynamics.** A standard application where a system of stochastic differential equations of Itô type occurs is being a stochastic perturbation of a deterministic nominal system. In this section we will study two flocking models, both of which are simplifications of a general non-linear system of stochastic differential equations.

Consider the set $\mathcal{V}$ of autonomous agents and fix $t_0 \in \mathbb{R}$ and $T \geq t_0$. The two vector valued stochastic processes $\mathbf{X}_t = (X_t^{(1)}, \dots, X_t^{(N)}), \mathbf{U}_t = \big(U_t^{(1)}, \dots, U_t^{(N)}\big)$ stand for the positions and the velocities of the members of the flock. $(\mathbf{X}_t, \mathbf{U}_t)$ are the solution of the system of Itô stochastic differential equations

$$\begin{aligned} dX_t^{(i)} &= U_t^{(i)}dt \\ dU_t^{(i)} &= \sum_j a_{ij}(t,\mathbf{X}_t)\big(U_t^{(j)} - U_t^{(i)}\big)dt + \sum_j g_{ij}(t,\mathbf{X}_t,\mathbf{U}_t)dB_t^{(ij)}, \end{aligned} \qquad i \in \mathcal{V} \qquad (27)$$

for $t \in [t_0, T]$, subject to initial data

$$X_{t_0}^{(i)} = X_i^0, U_{t_0}^{(i)} = U_i^0$$

or, equivalently, they are the solution of

$$\begin{aligned} d\mathbf{X}_t &= \mathbf{U}_t dt \\ d\mathbf{U}_t &= -L(t,\mathbf{X}_t)\mathbf{U}_t dt + \sum_{i=1}^{N} \mathbf{G}_i(t,\mathbf{X}_t,\mathbf{U}_t) d\mathbf{B}_t^{(i)} \end{aligned} \tag{28}$$

for $t \in [t_0, T]$, subject to initial data

$$\mathbf{X}_{t_0} = \mathbf{X}^0, \mathbf{U}_{t_0} = \mathbf{U}^0$$

provided $\mathbf{X}_t, \mathbf{U}_t$ are $\mathcal{U}_t$-measurable, $L(\mathbf{X},t)\mathbf{U}_t \in L^1_N(0,T)$, $G \in L^2_{N\times N}(0,T)$.

$$\begin{cases} \mathbf{X}_t = \mathbf{X}^0 + \int_{t_0}^t \mathbf{U}_s ds \\ \mathbf{U}_t = \mathbf{U}^0 - \int_{t_0}^t L\big(s,\mathbf{X}_s\big)\mathbf{U}_s ds + \sum_{i=1}^N \int_{t_0}^t G_i(s,\mathbf{X}_s,\mathbf{U}_s) d\mathbf{B}_s^{(i)} \end{cases} \quad a.s. \tag{29}$$

$\mathbf{B}^{(i)}(\cdot) = \big(B^{(i1)}(\cdot), B^{(i2)}(\cdot), \dots, B^{(iN)}(\cdot)\big)$ is an N-dimensional Brownian motion, $\mathbf{X}^0, \mathbf{U}^0$ are two N-dimensional random variables independent of $\mathbf{B}(\cdot)$.

Since we study the asymptotic behavior of solutions, we are essentially interested in the collection $\big\{(\mathbf{X}_t, \mathbf{U}_t)\big\}_{t\geq t_0}$ as solution of the above system of SDE's.

The stochastic system (27) exhibits *asymptotic strong stochastic* flocking if and only if the position-velocity processes $X_t^{(i)}, U_t^{(i)}$, $i \in \mathcal{V}$ satisfy the conditions

$$\lim_{t\to\infty} |U_t^{(i)} - U_t^{(j)}| = 0, \quad a.s. \quad \text{and} \quad \sup_{t\geq t_0} |X_t^{(i)} - X_t^{(j)}| < \infty, \quad a.s.$$

Additionally, the stochastic system exhibits *asymptotic strong stochastic* flocking in the mean square sense if the aforementioned processes converge accordingly.

Our aim is to discuss two simplifications of Eq. (27):

1. Time invariant flocking model with a state-independent multiple diffusions,

$$\begin{cases} dX_t^{(i)} = U_t^{(i)} dt \\ dU_t^{(i)} = \sum_j a_{ij}(U_t^{(j)} - U_t^{(i)})dt + \sum_j g_{ij}(t) dB_t^{(j)} \end{cases} \tag{26.1}$$

2. Time varying linear model with state-dependent stochastic disturbance and uniform time-varying diffusion coefficient

$$\begin{cases} dX_t^{(i)} = U_t^{(i)} dt \\ dU_t^{(i)} = \sum_j a_{ij}(t)(U_t^{(j)} - U_t^{(i)})dt + g(t)\sum_j \big(U_t^{(j)} - U_t^{(i)}\big) dB_t^{(j)} \end{cases} \tag{26.2}$$

4.2.1. *Time invariant flocking.* We begin with the study of Eq. (26.1) subject to initial data $\mathbf{X}^0, \mathbf{U}^0$. In the absence of noise, Eq. (26.1) reduces to Eq. (1), the dynamics of which are fully understood.

Assumption 4.8. *The associated graph $\mathbb{G}_A$ of the adjacency matrix $A = [a_{ij}]$, is routed-out branching.*

We recall the discussion in §2.1 and from Assumption 4.8 we deduce the existence and uniqueness of a the normalized eigenvector of the Laplacian with respect to the zero eigenvalue, $\mathbf{c} \in \mathbb{R}^N$. The solution of Eq.(1) with initial data $\mathbf{U}^0$ is $e^{-Lt}\mathbf{U}^0$ and it satisfies

$$|e^{-Lt}\mathbf{u}^0 - \mathbb{1}\mathbf{c}^T\mathbf{u}^0| \leq Ke^{-\Re\{\lambda\}t}$$

for some $K > 0$ that depends both on the norm $|\cdot|$ and the parameters a_{ij} and $\Re\{\lambda\}$ the second smallest real part of the eigenvalues of L that is strictly positive by Assumption 4.8.

Proposition 4.9. *The solution of* (26.1) $\big(\mathbf{X}_t, \mathbf{U}_t\big)$ *satisfies*

$$\mathbf{X}_t = \mathbf{X}^0 + \int_{t_0}^{t} \mathbf{U}_s ds$$
$$\mathbf{U}_t = e^{-L(t-t_0)}\mathbf{U}^0 + \int_{t_0}^{t} e^{-L(t-s)} G(s) d\mathbf{B}_s$$

for $t \in [t_0, T]$.

Proof. The form of $\mathbf{X}_t$ is the definition of the process so we will only prove the expression of $\mathbf{U}_t$. Define the process

$$\mathbf{V}_t := \mathbf{U}^0 + \int_{t_0}^{t} e^{L(s-t_0)} G(s) d\mathbf{B}_s$$

the differential of which is $d\mathbf{V}_t = e^{L(t-t_0)}G(t)d\mathbf{B}_t$. We will use Itô's product rule to calculate the differential of $e^{-L(t-t_0)}\mathbf{V}_t$ which is identical to $\mathbf{U}_t$:

$$d\big(e^{-L(t-t_0)}\mathbf{V}_t\big) = G(t)d\mathbf{B}_t - Le^{-L(t-t_0)}\mathbf{V}_t dt = -L\mathbf{U}_t dt + G(t)d\mathbf{B}_t.$$

Then the result follows. □

We see that in this simple case, the solution $\mathbf{U}_t$ is expressed in closed form. Asymptotic stochastic flocking is determined by the asymptotic behavior of the local martingales $\int_{t_0}^{t} g_{ij}(s)dB_s^j$ as $t \to \infty$.

Theorem 4.10. *Let Assumption 4.8 hold. If* $\mathbb{E}[(\mathbf{U}^0)^2], \mathbb{E}[(\mathbf{X}^0)^2] < \infty$ *and for any* $i, j \in \mathcal{V}$, *the functions* g_{ij} *satisfy*

$$\lim_{t\to\infty} \int_{t_0}^{t} g_{ij}^2(s)ds < \infty \qquad \text{and} \qquad \int_{t}^{\infty} g_{ij}^2(s)ds \in L^1_{[t_0,\infty]}$$

then asymptotic stochastic flocking occurs. In particular the agents align their speed around the $\mathcal{U}_\infty$*-measurable random variable*

$$k := \mathbf{c}^T\mathbf{U}^0 + \sum_{i,j} \int_{t_0}^{\infty} c_i g_{ij}(s) dB_s^{(j)}.$$

and they exhibit asymptotic stochastic flocking in the almost sure and in the mean square sense.

Proof. At first we clarify that k is well-defined since $\int_{t_0}^{\infty} g_{ij}(s)dB_s^{(j)}$ is almost surely finite exactly because the first imposed condition on g_{ij} yields almost sure finiteness by the Martingale Convergene Theorem [24]. Next,

$$\mathbf{U}_t - \mathbb{1}k = \big(e^{-L(t-t_0)} - \mathbb{1}\mathbf{c}^T\big)\mathbf{U}^0 + \int_{t_0}^{t} \big(e^{-L(t-s)} - \mathbb{1}\mathbf{c}^T\big)G(s)d\mathbf{B}_s + \mathbb{1}\mathbf{c}^T \int_{t}^{\infty} G(s)d\mathbf{B}_s$$

From the Properties of Itô's integral and the Cauchy-Schwarz inequality we obtain the following bound:

$$
\begin{aligned}
&\mathbb{E}\big[||\mathbf{U}_t - \mathbb{1}k||_2^2\big] \leq \\
&\leq K^2 e^{-2\Re\{\lambda\}(t-t_0)}\mathbb{E}[||\mathbf{U}^0||_2^2] + \mathbb{E}\bigg[\bigg(\int_{t_0}^{t}(e^{-L(t-s)} - \mathbb{1}\mathbf{c}^T)G(s)d\mathbf{B}_s\bigg)^2\bigg] + \\
&\qquad + \mathbb{E}\bigg[\bigg(\int_{t}^{\infty}\mathbb{1}\mathbf{c}^T G(s)d\mathbf{B}_s\bigg)^2\bigg] \\
&\leq K^2 e^{-2\Re\{\lambda\}(t-t_0)}\mathbb{E}[||\mathbf{U}^0||_2^2] + \sum_{i,j}\int_{t_0}^{t} K^2 e^{-2\Re\{\lambda\}(t-s)} g_{ij}^2(s)ds + \\
&\qquad + \sum_{i,j}\int_{t}^{\infty} c_i^2 g_{ij}^2(s)ds
\end{aligned} \tag{30}
$$

We have that by assumption $g_{ij}^2(t)$ vanishes. Now, $\mathbb{E}\big[||\mathbf{U}_t - \mathbb{1}k||_2^2\big]$ is bounded from above by three terms, each of which converges to zero as $t \to \infty$: the first, by Assumption 4.8, the third by the imposed condition on $g_{ij}(s)$'s and the second as a convolution of an L^1 function with a function that goes to zero. Then the random variable $\mathbf{U}_t$ converges asymptotically to Δ in the mean square sense. To prove almost sure speed coordination we first see that from the Chebyshev inequality for any, $\varepsilon > 0$

$$
\mathbb{P}\big(|U_t^{(i)} - U_t^{(j)}| \geq \varepsilon\big) \leq \frac{1}{\varepsilon^2}\mathbb{E}\big[|U_t^{(i)} - U_t^{(j)}|^2\big] \leq \frac{1}{\varepsilon^2}\mathbb{E}\big[||\mathbf{U}_t - \mathbb{1}k||_2^2\big]
$$

it is an easy exercise to show that all of the terms that bound $\mathbb{E}\big[||\mathbf{U}_t - \mathbb{1}k||_2^2\big]$ from above in Eq. (30) are integrable over $[t_0, \infty]$ (the second term can be proved by a simple change in the order of integration). Then because $\mathbb{P}\big(|U_t^{(i)} - U_t^{(j)}|\big)$ is summable, almost sure convergence to $\mathbb{1}k \in \Delta$ follows (see Theorem 4(c) of §7.2 in [14]).

Finally,

$$
|X_t^{(i)} - X_t^{(j)}| \leq |X_{t_0}^{(i)} - X_{t_0}^{(j)}| + \int_{t_0}^{t} |U_s^{(i)} - U_s^{(j)}|ds < \infty \ a.s.
$$

and hence $X_t^{(i)} - X_t^{(j)}$ is bounded in probability, therefore it is bounded in the 2^{nd}-mean (see Theorem 4(b) of §7.2 in [14]). □

It is noted that since g_{ij} are deterministic functions, k is a normally distributed random variable with mean $\sum_i c_i \mathbb{E}[U_i^0]$ and variance $\sum_{i,j} c_i^2 \int_{t_0}^{\infty} g_{ij}^2(s)ds$.

Remark 4.11. The results of this section can be trivially generalized to the case of time-varying connectivity weights $a_{ij}(t)$. In this case the kernel $\Phi(t, t_0)$ behaves similarly in the case of Sec. 3.2, whatever the connectivity regime may be.

4.2.2. *Time varying flocking.* Algebraic methods do not apply in general linear systems whereas stability in variation can effectively work in the case of state-independent noise as it was analyzed above. When a version with state-dependent noisy compartments is considered one would not want to disregard its contribution with respect to consensus. Given Eq. (26.2) subject to initial data $\mathbf{X}^0, \mathbf{U}^0$ we will derive expressions based on the coefficient of ergodicity. We will now state two assumptions; a very mild and a very strong one:

Assumption 4.12. *For $A = [a_{ij}(t)]$ it holds that $\max_{i\in\mathcal{V}} \sup_{t\geq t_0} \sum_j a_{ij}(t) < \infty$.*

Assumption 4.13. *The adjacency matrix of A_g of the diffusion compartment, corresponds to a complete graph.*

The necessity of Assumption 4.13 stems from the fact that the white noise $d\mathbf{B}_t$ as an integrator obeys no rules of monotonicity with respect to its integrand. For simplicity we introduce the notation $S(\mathbf{U}_t^2) = \max_{i,j}(U_t^{(i)} - U_t^{(j)})^2$

Proposition 4.14. *Under Assumptions 4.12 and 4.13, the solution $(\mathbf{X}_t, \mathbf{U}_t)$ of (26.2) satisfies*

$$d\big(S(\mathbf{U}_t^2)\big) \leq 2\big(g^2\frac{N}{2} - f(t)\big)S(\mathbf{U}_t^2)dt - g(t)S(\mathbf{U}_t^2)\sum_l dB_t^{(l)}$$

where $f(t) = \min_{i,j}\sum_l \min\{a_{il}, a_{jl}\}$.

Proof. We fix $t \geq t_0$ and we will always consider the elements $i = i_t, j = j_t \in \mathcal{V}$ that maximize the process $U_t^{(ij)} := \big(U_t^{(i)} - U_t^{(j)}\big)$, We need to derive an expression of the differential $d\big(e^{2mt}(U_t^{(ij)})^2\big)$ with Next we compute the differentials $dU_t^{(ij)}$ and $d\big(U_t^{(ij)}\big)^2$ using Itô calculus, as follows:

$$\begin{aligned}
&dU_t^{(ij)} = \sum_l \big(a_{il}U_t^{(li)} - a_{jl}U_t^{(lj)}\big)dt + g(t)\sum_l U^{(li)}dB_t^{(l)} - g(t)\sum_l U_t^{(lj)}dB_t^{(l)} = \\
&= \sum_l \big(a_{il}U_t^{(li)} - a_{jl}U_t^{(lj)}\big)dt - g(t)U^{(ij)}\sum_l dB_t^{(l)} \\
&d\big(U_t^{(ij)}\big)^2 = 2U_t^{(ij)}\big[\sum_l \big(a_{il}U_t^{(li)} - a_{jl}U_t^{(lj)}\big) + Ng^2U_t^{(ij)}\big]dt - 2g\big(U_t^{(ij)}\big)^2\sum_l dB_t^{(l)}
\end{aligned}$$

so eventually,

$$\begin{aligned}
&d\big(e^{2mt}(U_t^{(ij)})^2\big) = \\
&= 2me^{2mt}(U_t^{(ij)})^2dt + e^{2mt}d\big(U_t^{(ij)}\big)^2 \\
&= 2e^{2mt}U_t^{(ij)}\big[\big(m + \frac{Ng^2}{2}\big)U_t^{(ij)} + \sum_l \big(a_{il}U_t^{(li)} - a_{jl}U_t^{(lj)}\big)\big]dt - 2ge^{2mt}\big(U_t^{(ij)}\big)^2\sum_l dB_t^{(l)} \\
&= 2e^{2mt}U_t^{(ij)}\sum_l \big(a_{il} - a_{jl}\big)U_t^l dt - 2ge^{2mt}\big(U_t^{(ij)}\big)^2\sum_l dB_t^{(l)}
\end{aligned}$$

where $a_{ii} = m + \frac{Ng^2}{2} - \sum_l a_{il}$ which is positive for m large enough. At this point we shall focus on $\sum_l(a_{il} - a_{jl})U_t^{(l)}$ for which we notice that $a_{ij} > 0$ and $Q = \sum_l(a_{il} - a_{jl}) = 0$ for all $i, j \in \mathcal{V}$. Then if we let $w_l := a_{il} - a_{jl}$, we note that

$$\theta = \sum_{l:w_l>0} w_l = -\sum_{l:w_l<0} w_l$$

so

$$\begin{aligned}
Q &= \sum_{l:w_l>0} w_lU_t^{(l)} + \sum_{l:w_l<0} w_lU_t^{(l)} \\
&= \sum_{l:w_l>0} w_lU_t^{(l)} - \sum_{l:w_l<0} |w_l|U_t^{(l)} \\
&= \theta\Big(\frac{\sum_{l:w_l>0} w_lU_t^{(l)}}{\theta} - \frac{\sum_{l:w_l<0} |w_l|U_t^{(l)}}{\theta}\Big) \\
&\leq \theta\big(\max_l U_t^{(l)} - \min_l U_t^{(l)}\big) \\
&= \theta U_t^{(ij)}
\end{aligned}$$

then since $\theta \leq \frac{1}{2}\max_{i,j}\sum_l |a_{il}-a_{jl}| = m + g^2(t)\frac{N}{2} - \min_{i,j}\sum_l \min\{a_{il},a_{jl}\}$ we obtain the bound for $f(t) = \min_{i,j}\sum_l \min\{a_{il}(t),a_{jl}(t)\}$

$$d\big(e^{2mt}(U_t^{(ij)})^2\big) \leq 2e^{2mt}\big(m + \frac{Ng^2(t)}{2} - f(t)\big)\big(U_t^{(ij)}\big)^2 dt - 2g(t)e^{2mt}\big(U_t^{(ij)}\big)^2 \sum_l dB_t^{(l)}$$

and finally we obtain

$$\begin{aligned} d(U_t^{(ij)})^2 &= d\big(e^{-2mt}e^{2mt}(U_t^{(ij)})^2\big) \\ &= e^{-2mt}d\big(e^{2mt}(U_t^{(ij)})^2\big) - 2m(U_t^{(ij)})^2 dt \\ &\leq \big(Ng^2(t) - 2f(t)\big)\big(U_t^{(ij)}\big)^2 dt - 2g(t)\big(U_t^{(ij)}\big)^2 \sum_l dB_t^{(l)} \end{aligned}$$

as i,j are chosen to be the maximizers of $\big(U_t^{(i)} - U_t^{(j)}\big)^2$ the proof is concluded. □

Now we are ready to prove the flocking result :

Theorem 4.15. *Under Assumptions 4.12 and 4.13, asymptotic stochastic flocking for the system (26.2) occurs if*

$$e^{\int_{t_0}^t Ng^2(s)-2f(s)ds} \in L^1_{[t_0,\infty)}.$$

Proof. From Proposition 4.14 the solution $(\mathbf{X}_t, \mathbf{U}_t)$ satisfies [3]

$$S(\mathbf{U}_t^2) \leq e^{\int_{t_0}^t \left(\frac{N}{2}g^2(s)-2f(s)\right)ds + \sum_l \int_{t_0}^t g(s)dB_t^{(l)}} S(\mathbf{U}_{t_0}^2), \qquad a.s.$$

so that the second moment is calculated from the properties of the martingales $\int_{t_0}^t g(s)dB_s$ as :

$$\mathbb{E}\big[S(\mathbf{U}_t^2)\big] \leq e^{\int_{t_0}^t Ng^2(s)-2f(s)ds}\mathbb{E}\big[S(\mathbf{U}_{t_0}^2)\big]$$

and the rest of the proof is identical to this of Theorem 4.10. □

We note here that if $g^2 \in L^1$ then the non-summability of f suffices to prove consensus and in addition $e^{\int^t f(s)ds} \in L^1$ suffices to prove flocking. Other wise, $\int_{t_0}^\infty g^2(s)ds = \infty$ implies that $|\int_{t_0}^t g(s)dB_s|$ behaves asymptotically as

$$\sqrt{2\int_{t_0}^t g^2(s)ds \log\log \int_{t_0}^t g^2(s)ds}, \qquad a.s.$$

by the iterated logarithm for martingales [24, 14]. With this in mind it is possible that the assumption of the integrability of g^2 can be relaxed.

Remark 4.16. The two results above are improvements of [15, 1] as we allow weights to be non-symmetric and the diffusion coefficient to be time-varying. In the particular case of time independent weights we also allow minimal connectivity as well as we are able to identify the consensus point of the velocities. Furthermore, one is free to assume non-linear state-dependent connections on condition that they are lower-bounded away from zero. Initial conditions as for example Eq. (14) that automatically imply a bounded distance among agents as in the deterministic case do, are yet to be established.

5. **Discussion.** We considered several consensus and flocking models beyond the point of symmetric or uniform coupling, both in discrete and continuous time. Our analysis is based on the mathematical tool that measures the contraction rate of products of stochastic matrices with respect to the agreement sub-space Δ. The appropriate use of this concept allowed for a thorough and unified approach of several, seemingly different, models so that in most cases the convergence results were either recovered in a more concise manner or they were extended.

5.1. **Simple convergence analysis.** For the linear deterministic algorithms, we revealed the role of uniform bounds on positive connectivity weights with respect to the static or switching networks. The idea behind the unconditional exponential coordination in switching connectivity is the assumption of uniformly lower bound on coupling weights together with the uniform recurrent connectivity. Additional information on the values of the weights and the established connections are of interest only in the question on the rate of convergence to consensus.

The uniformly lower bound condition was replaced by sufficient conditions, on the rate at which the weights are allowed to vanish, for agreement among agents to asymptotically occur. The main result is stated both under static and under switching connectivity communication regime.

This case was considered in the continuous time dynamics, surprisingly enough, with the use of the same mathematics. To the best of our knowledge, the use of the coefficient of ergodicity to continuous systems was limited to cases of increased connectivity [40]. An insightful comparison between the discrete and continuous time dynamics revealed the way to extend the use of this valuable tool to the case of continuous time, as well. The basic problem was inverted from a differential equation to an integral equation one so that concepts from the theory of continuous time Markov Chains would be activated. Indeed, the necessarily positive value $B > 0$ considered in §3.2.2 (with Assumption 3.26) has nearly equivalent meaning to the necessary positive time one needs to classify the communication classes in a continuous time markov chain [14] .

Additionally, a major advantage in working with integral equations is that we would be free to allow discontinuous jumps of the connectivity weights without the mobilization of elaborated generalized concepts of solution of differential equations.

5.2. **The effect of symmetry.** Models with symmetry are more easy to use and analyze but are less realistic. In the case of consensus dynamics, symmetric coupling is very easily examined under the L^2 metric $\sum_{i<j}(x_i - x_j)^2$ and not the one used here $S(\mathbf{x})$. This is because in the former case, a very useful concept from topological algebraic theory, known as the *Fiedler number* plays the role of the contraction coefficient [4, 13] yielding better results. It can be shown for example that in the case of symmetric weights for a continuous time linear algorithm a condition $\int^\infty \min^+_{ij} a_{ij}(s)ds = \infty$ is enough to ensure consensus under a routed-out branching condition, whereas in the non-symmetric case we showed that we need $\int^\infty \min^+_{ij} a^\gamma_{ij}(s)ds = \infty$ where γ is the scrambling index.

This difference is best illustrated in the analysis of non-linear flocking models, the symmetric case of which is fully understood [9, 10] as compared to the non-symmetric one where increased connectivity was necessary [40]. In our work, we managed to extend the non-symmetric version to the point the symmetric versions are (i.e. the simple connectivity and switching connectivity regimes).

5.3. **The consensus point.** In §2.1 we saw that the time invariant nature of the system allows the existence of an integral of motion, i.e. $\mathbf{c}^T\mathbf{x}$. Then the consensus point $k = \mathbf{c}^T\mathbf{x}^0$. In general systems (time varying linear or non-linear) such integrals do not exist, unless one is willing to assume time invariant symmetry conditions. For instance, if $\mathbf{c}^T\mathbb{R}^N$ is a row vector such that $\mathbf{c}^T L(t) \equiv 0$, i.e. a common left eigenvector of the Laplacian matrices $L(t)|_{t\geq t_0}$. A straightforward example is the case $a_{ij} = a_{ji}$ where $\mathbf{c} = \frac{1}{N}\mathbb{1}$.

5.4. **More on non-linearity.** Beyond flocking models, we also exploited the coefficient of ergodicity as a contraction kernel to two important generalizations of the linear case. The first one explains that the the underlying concept of convergence in linear algorithms is that these systems have the property of passive co-operative coupling. Based on this property we showed that a simple direct linearization is enough to reduce the study of the original non-linear problem to a linear one with the solutions of these two to be indistinguishable.

Next, we exploited the contraction coefficient as a kernel to study non-linear and essentially non-monotonic consensus algorithms, by stability in variation. We created a fixed point theory argument and established conditions under which the non-linear system converges to consensus by also providing estimates on the rate of convergence.

5.5. **Necessary conditions revisited.** Apart from Theorem 3.31 in §3.3 we systematically avoided to state necessary conditions for consensus and we had a good reason to do so.

We know that uniformly lower bounds together with the recurrent connectivity condition suffice to ensure exponential convergence. Heuristically speaking, the positive diagonal elements of $P(t)$ ensure that if the latter condition is not met for a single connected component then it may be met for two components over uniformly bounded intervals of time. Hence there are only specific initial conditions that ensure global consensus, whereas there are other who do not. If, in turn, the latter condition is not met, then recurrent connectivity condition may be met for three connected components and so forth, until the number of components meets the number of agents. These types of necessary conditions are presented for example in [29, 46, 31].

Whenever the weights are free to vanish, however, the above argument is of little use and the case of non-convergence is at the moment, only discussed by examples. On the other hand, it may be the contraction coefficient appears that it is unfit for providing necessary conditions. We reported above for example that it gives more conservative estimates than the Fiedler number in the symmetric case.

In this work, we stated another type of necessary conditions to consensus, mainly to highlight the discrepancy with the sufficient conditions. Indeed the necessary conditions are significantly milder than the sufficient ones due to the, already mentioned above, fact that the coefficient of ergodicity typically provides very conservative estimates of contraction.

5.6. **Stochastic regularity and noise.** The developed framework for the deterministic case was carried through to the case of stochastic sources. In the literature, uncertainty is supplied either on the part of connections or by considering pure-noise disturbances. In the first case, the time-invariant deterministic rules that ensure sufficient connectivity to asymptotic agreement, were transported to a measure theoretic framework where the dynamics were dictated by measure preserving dynamical systems, a fairly unifying and simple approach which covers several proposed models. We showed that whenever the weights are lower bounded away from zero, the event of asysmptotic consensus can be proved with probability one. On the other hand, whenever the weights are allowed to vanish, the event of asymptotic consensus can only happen with a positive probability and not almost surely.

Finally, in the case of noisy perturbations, the coefficient of ergodicity was mobilized once again to produce new results in this direction. We remark yet again, how symmetry plays a crucial role in the analysis of Itô's stochastic differential equations. In particular, whenever the noisy part is dictated by a Brownian motion, the dynamics are too irregular to sustain general asymmetric connections on the part of the noise, whenever the latter depends on the state. The analysis provides a number of fruitful problems for the future.

REFERENCES

[1] S.M Ahn and S-Y. Ha. Stochastic flocking dynamics of the cucker-smale type with multiplicative white noises. *Journal of Mathematical Physics*, 51:103301–1–103301–17, 2010.

[2] M. Arcak. Passivity as a design tool for group coordination. *IEEE Transactions on Automatic Control*, 52(8):1380–1390, 2007.

[3] L. Arnold. *Stochastic Differential Equations: Theory and Applications*. Wiley-Interscience publication, 1974.

[4] B. Bollobás. *Modern Graph Theory*, volume 184. Springer, 1998.

[5] T.A. Burton. *Stability by Fixed Point Theory for Functional Differential Equations*. Dover, 2006.

[6] F. R K Chung. *Spectral Graph Theory*, volume 92. American Mathematical Society, 1997.

[7] F. Cucker and J. G. Dong. A general collision-avoiding flocking framework. *IEEE Transactions on Automatic Control*, 56(5):1124–1129, 2011.

[8] F. Cucker and E. Mordecki. Flocking in noisy environments. *J. Math. Pures Appl.*, 89:278–296, 2008.

[9] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Transactions on Automatic Control*, 52(5):852–862, 2007.

[10] F. Cucker and S. Smale. On the mathematics of emergence. *Japan Journal of Mathematics*, 2:197–227, 2007.

[11] M. H. DeGroot. Reaching a consensus. *Journal of American Statistical Association*, 69:118–121, 1974.

[12] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides*. Kluwer Academic Publishers, 1988.

[13] C. Godsil and G. Royle. *Algebraic Graph Theory*, volume 207. Springer, 2001.

[14] G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Oxford Univ. Press, 3 edition, 2003.

[15] K. Ha, S-Y. Lee and D. Levy. Emergence of time-asymptotic flocking in a stochastic cucker-smale system. *Commun. Math. Sci.*, 7(2):453–469, 2009.

[16] S-Y. Ha and Tadmor E. From particle to kinetic and hydrodynamic description of flocking. *Kinetic and Related Models*, 1(3):415–435, 2008.

[17] S-Y. Ha and J. G. Liu. A simple proof of the cucker-smale flocking dyanics and mean-field limit. *Commun. Math. Sci.*, 7(2):297–325, 2009.

[18] D. J. Hartfiel. *Markov Set-Chains*, volume 1695. Springer, 1998.

[19] D. J. Hartfiel. *Nonhomogeneous Matrix Products*. World Scientific, 2002.

[20] Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Transactions on Automatic Control*, 50(11):1867–1872, 2005.

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

[22] 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(6):988–1001, 2003.

[23] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*, volume 54. Cambridge University Press, 1996.

[24] L. B. Koralov and Y. G. Sinai. *Theory of Probability and Random Processes*. Springer, 2 edition, 2007.
[25] U. Krause. A discrete nonlinear and non-autonomous model of consensus formation. *Communications in difference equations*, pages 227–236, 2000.
[26] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer-Verlag, 1984.
[27] V. Lakshminkantham and Deo S.G. *Method of Variation of Parameters for Dynamic Systems*, volume 1. Gordon and Breach Science, 1998.
[28] Jan Lorenz and Dirk A. Lorenz. On conditions for convergence to consensus. *IEEE Trans. Automat. Contr.*, 55(7):1651–1656, 2010.
[29] W. Lu, F. M. Atay, and J. Jost. Synchronization of discrete-time dynamical networks with time-varying couplings. *SIAM J. Math. Anal.*, 39(4):1231–1259, 2007.
[30] A. A. Markov. Extensions of the law of large numbers to dependent quantities (in russian). *Izn. Fiz. - Matern. Obsch. Kazan Univ.*, 15, 1906.
[31] I. Matei, J. S. Baras, and C. Somarakis. Convergence results for the linear consensus problem under markovian random graphs. *SIAM J. Control Optim.*, pages 1574–1591, 2013.
[32] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.
[33] L. Moreau. Stability of continuous-time distributed consensus algorithms. *43rd IEEE Conference on Decision and Control (2004)*, page 21, 2004.
[34] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(02):169–182, 2005.
[35] R. Olfati-Saber and R. M. Murray. Consensus problems in networks with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(09), 2004.
[36] A. Papachristodoulou, A. Jadbabaie, and U. Mnz. Effects of delay in multi-agent consensus and oscillator synchronization. *IEEE Trans. Automat. Contr.*, 55(6):1471–1477, 2010.
[37] M. Porfiri and D. L. Stilwell. Consensus seeking over random weighted directed graphs. *IEEE Transactions on Automatic Control*, 52(9):1767–1773, 2007.
[38] H. Qing and W. M. Haddad. Distributed nonlinear control algorithms for network consensus. *Automatica*, 44(9):2375 – 2381, 2008.
[39] H. L. Royden. *Real Analysis*. McMillan Publishing Company, 3 edition, 1989.
[40] Motsch S. and Tadmor E. Heterophilious dynamics enhances consensus. *arXiv:1301.4123*, 2013.
[41] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer, 2006.
[42] D.R. Smart. *Fixed Point Theorems*. Cambridge University Press, 1980.
[43] H.L Smith. *Monotone Dynamical Systems, An Introduction to the Theory of Competitive and Cooperative Systems*, volume 41. AMS, 1995.
[44] 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)*, pages 1–8, 2012.
[45] A. Tahbaz-Salehi and A. Jadbabaie. A necessary and sufficient condition for consensus over random networks. *IEEE Transactions on Automatic Control*, 53(3):791–795, 2008.
[46] A. Tahbaz-Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. *IEEE Transactions on Automatic Control*, 55(1):225–230, 2010.
[47] J. Tsitsiklis, D. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9):803–812, 1986.
[48] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75(6):1226–1229, 1995.
[49] P. Walters. *An introduction to Ergodic Theory*, volume 79. Springer, 1982.