New Results On Stochastic Consensus Networks*

Christoforos Somarakis† John S. Baras‡

Abstract

We revisit the general linear consensus problem and provide new conditions for asymptotic consensus on two types of stochastic versions of the algorithm, in discrete and continuous time. We show that our method unifies a number of proposed models in the literature as well as it extends and generalizes existing results.

1 Introduction

Self-organized networks are an important class of complex systems and a central topic of networked control theory. Examples of such networks typically illustrate a global collective behavior as a result of a local interaction among autonomous agents.

Perhaps the simplest control algorithm for coordination is the linear consensus algorithm. In its classic version, it involves 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 ($n \in \mathbb{N}$) or continuous time ($t \in \mathbb{R}$):

$$x_i(n+1) = \sum_j p_{ij}^n x_j(n), \quad \dot{x}_i = \sum_j p_{ij}^t (x_j - x_i) \tag{0.1}$$

The non-negative numbers p_{ij}^n or p_{ij}^t model the influence of agent j on i. They essentially characterize the connectivity regime and eventually the process of the asymptotic alignment. For the discrete model the basic assumption is $\sum_j p_{ij}^n \equiv 1$ and for the continuous model $\sum_j p_{ij}^t \equiv 0$. These types of algorithms have been extensively studied in the literature a recent review of which can be found in [15].

The central objective for the asymptotic analysis of systems like (0.1) is the derivation of sufficient conditions on the coupling weights p_{ij} in order for the solution $\mathbf{x}$ to satisfy asymptotic convergence to a constant value, i.e. say for the discrete time $x_i(n) \to k$ as $n \to \infty$, $\forall\ i = 1, \ldots, N$. In the context of these algorithms this is equivalent to

$$|x_i(n) - x_j(n)| \to 0 \text{ as } n \to \infty,\ \forall\ i, j,$$

(see [15]).

In their seminal work, Jadbabaie et al. [8] proved that a sufficient condition for asymptotic consensus occurs if the union of the communication graphs over uniformly bounded intervals of time corresponds to a routed-out branching graph. This is known as the recurrent connectivity condition and it has been consolidated as the mildest sufficient connectivity condition which ensures asymptotic consensus [13].

1.1 Related literature It didn't take long until that type of connectivity was substituted by its stochastic counterpart.

One reason is that real-world networked systems suffer from various communication failures or creations between nodes. Such variations in topology can happen randomly, and this motivates the investigation of consensus problems under a stochastic framework.

Another reason is that researchers criticized the recurrent connectivity condition as a too stringent deterministic scheme. It seems that asking a priory an internal overall connectivity every a uniformly bounded period of time and unacceptably strong assumption. Adding statistical regularity on the dynamics of interconnections, this assumption is satisfied almost surely. Thus it can be omitted. Along these lines, numerous modifications have been introduced [7, 14, 17, 18, 11].

A fundamentally different approach of probabilistic formulation of consensus algorithms is proposed for the flocking models where the dynamic alignment occurs in $\dot{\mathbf{x}}$ rather than in $\mathbf{x}$. This stochastic counterpart of (0.1) has been studied by noisy perturbation of the nominal deterministic models [5, 1, 3]. This perturbation is the additive multiplicative White Noise and sufficient conditions are derived for almost sure consensus.

1.2 This work. Recent advances in the linear consensus problem have extended the result to more general conditions and types of convergence to a common value [16, 15]. The present paper is an excerpt from [15]

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

†Dept. of Applied Mathematics, the Institute for Systems Research, University of Maryland, College Park, MD, USA csomarak@umd.edu

‡Dept. of Electrical Engineering, the Institute for Systems Research, University of Maryland, College Park, MD, USA baras@umd.edu

where a general theory for consensus systems is developed. In this monograph, we discuss a unified approach based on the universal use of the contraction coefficient (the central mathematical tool for estimating the averaging effect of stochastic matrices on vectors, [6]) both for the discrete and the continuous time systems under mild connectivity assumptions. In particular, we establish generalized conditions for consensus that demand neither symmetry nor uniformly lower bounds on the connectivity weights. Hence we are led to non-uniform type of convergence to consensus.

In this paper, we discuss the stochastic counterpart of these models from both the aforementioned perspectives. In the first part, we impose uncertainty in the existence of connections. We establish probabilistic rules to control these particular dynamics and we propose a new framework based on measure preserving dynamical systems. We will show that convergence to consensus can happen only with a positive probability strictly less than one exactly because of the non-uniform connectivity condition. We develop our theory in discrete time only and our main contribution is the unified result this framework provides among a number of important stochastic versions proposed in the literature.

The second part, stochastic approach deals with uncertainties in the equations. This leads to "noisy" 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. Here our analysis concerns continuous time dynamics.

The paper is organized as follows: Basic notations and definitions are provided in §2. In §3 we state our models and the leading hypotheses with preliminary results of the deterministic version taken from [15]. In §4 we state and prove the first result of this paper and we work on examples from the literature to highlight the unifying perspective of our results. In §5 we proceed with convergence results on flocking networks and stochastic differential equations in continuous time. A thorough discussion of the overall obtained results with concluding remarks is held in §6.

2 Notations & Definitions

We set $\mathbb{Z}$ for the integers, $\mathbb{N}$ is for the natural numbers and $\mathbb{R}$ for the real numbers. For $N \in \mathbb{N}$ we define $\mathcal{V} := \{1, \dots, N\}$. Any vector $\mathbf{x} \in \mathbb{R}^N$ is considered as a column vector, unless otherwise stated. The *agreement* or *consensus* space Δ is defined 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 $M\mathbf{x} \in \Delta$, $\forall \mathbf{x} \in \mathbb{R}^N$. By $\mathbf{1}$ we understand the $N-$dimensional vector with all entries equal to 1. By I we understand the $N \times N$ identity matrix. By $||\cdot||_p$ we denote the p-norm where in particular $\mathbf{x}^2 = \mathbf{x}^T\mathbf{x} = ||\mathbf{x}||_2^2$. Also $\delta(\cdot)$ takes the place of the delta function, L^1_S denotes the space of absolutely integrable functions that are defined in S and $[v]$ denotes the integer part of v.

2.1 Elements of algebraic graph theory By a topological directed graph $\mathbb{G}$ we understand the pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{E} = \{(i,j) \in \mathcal{V} : i,j \in \mathcal{V}\}$ is the set of edges where in principal $(i,j) \neq (j,i)$. By $\mathcal{S}$ we denote the family of graphs with fixed N nodes and self-edges one every node. Also $\mathcal{T} \subset \mathcal{S}$ the subset of graphs each of which is routed-out branching, i.e. each member of $\mathcal{T}$ has at least one node $i \in \mathcal{V}$ (the root) through which the whole graph can be traversed. Each member $\mathbb{G}_i$ of it, has a scrambling index γ_i, i.e. the minimum number of steps from a root node to reach all the other nodes of graph. 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

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

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 2.1. *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 $l^* := \max_i\{l_{i,i-1}\}$.*

A non-negative matrix $P = [p_{ij}]$ is such that $p_{ij} \geq 0$. P is stochastic if it is non-negative and $\sum_j p_{ij} \equiv 1$. Any non-negative matrix P can be represented as a graph $\mathbb{G}_P = (\mathcal{V}, \mathcal{E})$ so that $p_{ij} \neq 0 \Rightarrow (i,j) \in \mathcal{E}$. We are interested in backward products $P_{t,h} = P_{t+h}P_{t+h-1} \dots P_{t+1}$ for $t \geq 0, h \geq 1$.

2.2 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 $\mu(B) = 1$.

Given a collection of spaces, $\{(\mathbb{X}_n, \mathcal{B}_n, \mu_n)\}_{n\in\mathbb{N}}$, we define the *product* space in the natural way: $\mathbb{X} = \prod_{n\in\mathbb{N}} \mathbb{X}_n$ and a point $\chi \in \mathbb{X}$ is considered to be the sequence $\chi = \chi_0\chi_1\chi_2\dots$ 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 $\mathbb{X}$:

$$\begin{aligned}\mathcal{J} &= \Big\{ \prod_{j\le n_1-1} \mathbb{X}_j \times \prod_{n_1\le j\le n_2} A_j \times \prod_{j\ge n_2+1} \mathbb{X}_j \Big\} \\ &= \big\{\chi \in \mathbb{X} : \chi_j \in A_j, j \in [n_1, n_2]\big\}^{A_j\in\mathcal{B}_j}_{0\le n_1\le n_2}\end{aligned}$$

each of which is a *measurable rectangle* (or a *cylinder*). On each of the above rectangles we attach the value $\prod_{n=n_1}^{n_2} \mu_n(A_n)$ and this can be extended to a probability measure μ on $(\mathbb{X}, \mathcal{B})$ in the standard way [19], 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 $T(\chi_0\chi_1\chi_2\dots) = \chi_1\chi_2\dots$ and it may attain all the desired properties of measure preserving and ergodicity. By $T^n\chi$ we mean the element $\chi_n\chi_{n+1}\dots$ and we will also use the projection map $\{T^n\chi\} = \chi_n$, $\chi_n \in \mathbb{X}_n$. For more on dynamical systems and ergodic theory the reader is referred to [19, 9].

2.3 Elements of SDEs and stochastic flocking

Let $(\Omega, \mathcal{U}, \mathbb{P})$ be a probability space. Fix $t_0 \in \mathbb{R}$ and let $\{\mathbf{B}(t)\}_{t\ge t_0} \in \mathbb{R}^N$ be a Brownian motion. Let $\mathbf{X}_0, \mathbf{U}_0 \in \mathcal{U} \to \mathbb{R}^N$ be two random variables independent of $\mathbf{B}(t_0)$. If the σ-algebra generated by $\mathbf{X}_0, \mathbf{U}_0$ and the history of the Brownian motion up to (and including) time t is $\mathcal{U}_t$ then $(\Omega, \mathcal{U}, \mathcal{U}_t, \mathbb{P})$ is a complete filtered probability space. Consider the set $\mathcal{V}$ of agents and fix $T \ge t_0$. The two processes $\mathbf{X}_t = (X_t^{(1)}, \dots, X_t^{(N)}), \mathbf{U}_t = (U_t^{(1)}, \dots, U_t^{(N)})$ stand for the positions and the velocities of the members of the flock and they are the solution of the system of Itô stochastic differential equations

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

for $i \in \mathcal{V}$, $t \in [t_0, T]$ and subject to initial data $X_{t_0}^{(i)} = X_i^0, U_{t_0}^{(i)} = U_i^0$, provided $\mathbf{X}_t, \mathbf{U}_t$ are $\mathcal{U}_t$ adapted processes, $L(t)$, stands for the time-varying laplacian matrix, G is square integrable in (t_0, T) and

$$\begin{cases} \mathbf{X}_t = \mathbf{X}^0 + \int_{t_0}^t \mathbf{U}_s ds \\ \mathbf{U}_t = \mathbf{U}^0 - \int_{t_0}^t P(s)\mathbf{U}_s ds + \sum_i \int_{t_0}^t G_i(s) d\mathbf{B}_s^{(i)} \end{cases} \tag{2.1}$$

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

Definition 2.1. *Equation (2.1) exhibits asymptotic strong stochastic flocking if $(X_t^{(i)}, U_t^{(i)})$ satisfy*

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

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

3 The model setup.

In this section, we present our models and a first set of assumptions. We will discuss two very closely related versions of the discrete time problem: the "deterministic" and the "probabilistic" one. Our aim is to state without proof the result of the former type and then to continuous with the probabilistic formulation on the second one basing the proof of our main result on the deterministic theorems.

3.1 The deterministic case

Fix $n_0 \in \mathbb{N}$ and consider the initial value problem is:

$$\begin{cases} \mathbf{x}(n+1) = P(n)\mathbf{x}(n), \quad n \ge n_0 \\ \mathbf{x}(n_0) = \mathbf{x}^0 \end{cases} \tag{3.2}$$

where $\mathbf{x} = (x_1, \dots, x_N)^T$ is the state vector and $P(n) = [p_{ij}^n]$ with $p_{ij}^n \ge 0$ and in particular $p_{ii}^n = 1 - \sum_j p_{ij}^n$.

Assumption 3.1. *$p_{ij}^n = p_{ij}(n) : \mathbb{Z}_+ \to \mathbb{R}_+$ are defined such that $\forall\, n \ge n_0$, $i, j \in \mathcal{V}$ with $j \ne i$*

1. *$\sum_j p_{ij}^n \le m < 1$.*
2. *$p_{ij}^n > 0 \Rightarrow p_{ij}^n \ge f(n)$ where f has the properties: $\exists M \in [n_0, \infty)$ s.t. $f(n) \in (0, 1-m]$ $\forall n \ge M$ and $f(n) \to 0$.*

Assumption 3.2. *There exist $M > n_0$, $B \ge 1$ such that $\mathbb{G}_{P_{n,B}} \in \mathcal{T}, n \ge M$.*

Theorem 3.1. *[15] Under Assumptions 3.1 and 3.2, the solution $\mathbf{x}$ of (3.2) satisfies*

$$\mathbf{x}(n) \to \Delta \ \text{as}\ n \to \infty,$$

if

$$\sum_l f^\sigma(M + l\sigma - 1) = \infty$$

where $\sigma = l^*([N/2] + 1)B$ and l^* as in Remark 2.1.

For the continuous time case we fix $t_0 \in \mathbb{R}$ and consider the initial value problem

$$(3.3) \qquad \begin{cases} \dot{\mathbf{x}}(t) = -P(t)\mathbf{x}(t), t \geq t_0 \\ \mathbf{x}(t_0) = \mathbf{x}^0 \end{cases}$$

where $P(t) = [p_{ij}^t]$ is known as the Laplacian matrix with the property that $p_{ij}^t \geq 0$ for $i \neq j$ and $p_{ii}^t = -\sum_{j\neq i} p_{ij}^t$. The corresponding connectivity conditions are now stated:

Assumption 3.3. *1. $\forall\ i, j \in \mathcal{V}$, $p_{ij}^t \geq 0$ are upper bounded, right continuous functions of time.*

2. $\forall\ i \neq j \in \mathcal{V}$ and $p_{ij}^t \neq 0$ implies $p_{ij}^t \geq f(t)$ where $f(\cdot) > 0$, non-increasing with the property that $f(t) \to 0$ as $t \to \infty$.

3. $\forall\ t \geq t_0, \exists\ \epsilon > 0$ independent of t such that $p_{ij}(t) \neq 0 \Rightarrow p_{ij}(s) \geq f(s)$ for $s \in S_\epsilon(t^*) = [t^* - \epsilon, t^* + \epsilon]$ for some $t^* \in \mathbb{R}$ such that $t \in S_\epsilon(t^*)$.

We take $m > \max_i \sup_{t\geq t_0} \sum_j p_{ij}^t$ finite. This number exists in view of the assumption right above. We fix $B > 0$ and define the matrix:

$$W_B^t = \int_{t-B}^{t} e^{-mB}\delta(s-(t-B))I + e^{-m(t-s)}\big(mI - P(s)\big)\, ds$$

which it can be shown to be stochastic [15].

Assumption 3.4. *There exist $M > t_0$, $B > 0$ such that $\mathbb{G}_{W_B^t} \in \mathcal{T}$, $t \geq M$.*

Theorem 3.2. *[15] Under Assumptions 3.3 and 3.4 the solution $\mathbf{x}$ of (3.3) satisfies*

$$\mathbf{x}(t) \to \Delta \text{ as } t \to \infty,$$

if there exists a sequence $\{t_i\}_{i\geq 0} \geq t_0$ 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^* as in Remark 2.1.

It should be noted that whenever the non-zero weights p_{ij}^n or p_{ij}^t are uniformly bounded from below, the convergence of (3.2) or (3.3) to Δ is exponential. In [15] explicit estimates are provided for arbitrary connectivity signals but they are beyond the scopes of this paper.

In view of Theorem 3.2 and the linearity of the problem one can write the solution as $\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}^0$ with $\Phi(t, s)$ the transition matrix satisfying

$$(3.4) \qquad \big|\big|(\Phi(t, t_0) - \mathbf{1}\mathbf{1}^T\mathbf{c})\mathbf{x}^0\big|\big|_2 \leq Ke^{-\theta(t-t_0)}||\mathbf{x}^0||_2$$

for some $K, \theta > 0$ that depend on system parameters and the 2-norm and $\mathbf{c} \geq 0$ with $\sum_i c_i = 1$.

3.2 The probabilistic case

In the first probabilistic type of consensus systems. The event that there is a connection from agent j to i is no longer of deterministic nature. More specifically, for any t the stochastic matrix $P(t)$ is generated by a steering force that obeys independent probability rules. These rules take place in a product space $(\mathbb{X}, \mathcal{B}, \mu) = \prod_{n\geq 0}(\mathbb{Y}, 2^{\mathbb{Y}}, m)$ for some measure function m and the induced product measure as it was discussed in §2.2. The shift $T : \mathbb{X} \to \mathbb{X}$ is a measure preserving transformation since for any $A \in \mathcal{J}$, $\mu(T^{-1}A) = \mu(A)$. We understand $\chi_i \in \mathbb{X}$ by an $N \times N$ matrix with elements from $\{0, 1\}$ so tha all diagonals are 0 and the off-diagonals are 1 if there is a connection from the agent of the column to the agent of the row, otherwise they attain the zero value. Let the family of functions $\{a_{ij}(n)\}_{i\neq j} \in [f(n), \bar{a})$ where f is a fixed non-increasing function that vanishes as $n \to \infty$. We define the stochastic matrix

$$(3.5) \qquad P(n) = \phi\big(\{T^n\chi\}\big)$$

where $\phi : \mathbb{X} \to \mathcal{S}$ is a measurable function defined as follows: For $i \neq j$

$$\big[\phi\big(\{T^n\chi\}\big)_{ij}\big] = \begin{cases} \frac{a_{ij}(n)}{\varepsilon \sum_{l:[\{T^n\chi\}_{il}]=1} \bar{a}}, & [\{T^t\chi\}_{ij}] = 1 \\ 0, & [\{T^t\chi\}_{ij}] = 0 \end{cases}$$

for some fixed $\varepsilon > 1$ and

$$\big[\phi\big(\{T^n\chi\}\big)_{ii}\big] = 1 - \sum_j \big[\phi\big(\{T^n\chi\}\big)_{ij}\big]$$

Because of the uncertainty on the connection status we chose this slightly different version on the weights. Now, it is easy to see that for any $\chi_i \in \mathbb{X}$, $P(n)$ is always stochastic. Next we consider the set

$$Q_B := \big\{\chi \in \mathbb{X} : \mathbb{G}_{P_{n,B}} \in \mathcal{T}\ \forall\ n \geq 0\big\} \in \mathcal{B}.$$

4 Results on probabilistic consensus networks

The setting makes $\mathbf{x}$ a stochastic process defined on a probability space closely related to $(\mathbb{X}, \mathcal{B}, \mu)$. Consensus may be achieved if μ assigns a positive value to Q_B for some B such that $\sum_n f^\sigma(n\sigma) = \infty$ for $\sigma = \sigma(B)$ of course. A measure theoretic analogue of Theorem 3.1 is the following:

Theorem 4.1. *Let the initial value problem (3.2) with $P(n)$ defined as in (3.5). Let Assumptions of Theorem 3.1 hold with $M = 1$. The solution $\mathbf{x}$ converges to Δ with probability $\mu(P_B)$ if $\sum_n f^\sigma(n\sigma) = \infty$.*

Theorem 4.1 illustrates the interdependence between the non-uniform lower bound f and the induced statistical regularity and it is only of theoretical interest. Almost sure convergence is ensured if the event $\bigcup_{B\geq 1} Q_B$ 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. It is exactly this case where there is no difference between the existence of connection and its weight for simple convergence to Δ.

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

Corollary 4.1. *Let $T : \mathbb{X} \to \mathbb{X}$ be an ergodic shift on the product space $(\mathbb{X}, \mathcal{B}, \mu)$, $P(n)$ with the form of* (3.5) *and $a_{ij}(n) \in \{0\} \cup (0,1)$ for $i \neq j$. Then the solution $\mathbf{x}$ of* (3.2) *converges to consensus with probability one if $\mu(Q_B) > 0$ for some $B \geq 0$.*

Proof. At first we show that the set $W = \bigcup_B Q_B$ is T-invariant. Fix $B > 0$. Then for $\chi \in Q_B$ we have $T^{-1}\chi \in Q_{B+1}$.

$$T^{-1}W = T^{-1}\bigcup_B Q_B = \bigcup_B T^{-1}Q_B \subset \bigcup_B Q_{B+1} \subset W$$

Also, $\chi \in Q_B \Rightarrow T\chi \in Q_B$ so that $Q_B \subset T^{-1}Q_B$ and this is true over the union for all $B \geq 0$. Consequently

$$W \subset T^{-1}W$$

and we conclude that W is T-invariant. 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 $\mu(W) \geq \mu(Q_B) > 0$. Then the only realization of 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 $P(n)$ is not a stationary process. By construction the measure μ does not concern the weights $a_{ij}(n)$. The stationarity property can be observed in $\mathbb{G}_{P(n)}$ which, as we mentioned above, is the only key feature for the stability analysis.

4.1 Examples

Example. [Stationary Ergodic processes [18]] The problem of consensus over stationary ergodic processes assumes that the matrix $P(n)$ 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 [10]. Given a stationary ergodic process that produces stochastic matrices $P(n)$ process one can easily verify whether this particular shift is ergodic after applying Birkhoffs ergodic theorem: If for some $B > 0$

$$\lim_{\bar{n}} \frac{1}{\bar{n}} \sum_{n=0}^{\bar{n}-1} \mathbf{1}_{\mathcal{T}}(P_{n,B}) > 0$$

where $\mathbf{1}_A(s)$ is a dual function that takes value 1 if $s \in A$ and 0 otherwise, $\mathcal{T}$ is the set of routed-out branching graphs, then the corresponding shift that is ergodic and consensus is proved in the almost sure sense. Corollary 4.1 reproduces the results of [18] 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.

Example. [IID processes [7, 17]] One of the first works on the topic of probabilistic consensus in [7], formulated (3.2), as a stochastic linear equation with symmetric connectivity weights ($a_{ij} = a_{ji}$) to randomly take values at each time $n \in \mathbb{N}$. The partition of interest would be $a_{ij}(n) \neq 0$ with probability p and $a_{ij}(n) = 0$ with probability $1-p$, independently of the rest of the connections and times.

Let us digress for a moment and see $P(n) = P_n(y)$ as a random process defined on a probability space $(\mathbb{Y}, 2^{\mathbb{Y}}, \mathbb{P})$. Then $P(n)$ takes values in the space of stochastic matrices with positive diagonals and uniformly bounded weights. Then the backward product $P_{n,B}(y)$ 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 $\mathbb{G}_{P_{n,B}}$ 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 $q \in [1, N(N-1))$ denote the minimal number of edges so that each additional edge will keep $\mathbb{G}$ routed-out branching. Then for Q_B as defined before

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

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

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 $N(N-1)$ 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 [19], T is measure preserving. It is a standard exercise to show that T is ergodic [19]. It is only left to show that for some $B > 0$, $\mu(Q_B) > 0$, a calculation very similar to the one carried before and Corollary 4.1 applies.

Example. [Markov processes [11]] The authors considered (**??**) 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 [19]. Then the event of connectivity over a B-interval of times is dictated by the invariant measure to be of positive measure and Corollary 4.1 applies.

5 Results on noisy flocking networks

Let us now turn to the study of Eq. (2.1) subject to initial data $\mathbf{X}^0, \mathbf{U}^0$. In the absence of noise the model reduces to Eq. (3.3). We begin with a useful form of its solution.

Proposition 5.1. *The solution* $\big(\mathbf{X}_t, \mathbf{U}_t\big)$ *of* (2.1) *satisfies*

$$\mathbf{X}_t = \mathbf{X}^0 + \int_{t_0}^{t} \mathbf{U}_s\, ds$$

$$\mathbf{U}_t = \Phi(t, t_0)\mathbf{U}^0 + \int_{t_0}^{t} \Phi(t, s)G(s)\, d\mathbf{B}_s$$

where $\Phi(t, s)$ *is defined in 3.4 and* $t \in [t_0, T]$.

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

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

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

$$\begin{aligned} d\big(\Phi(t, t_0)\mathbf{V}_t\big) &= G(t)d\mathbf{B}_t - L\Phi(t, t_0)\mathbf{V}_t dt \\ &= -L\mathbf{U}_t dt + G(t)d\mathbf{B}_t. \end{aligned}$$

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

Theorem 5.1. *Let Assumptions 3.3 and 3.4 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*

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

then 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 sense of Def. 2.1.

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 [10]. Next from the definition of k and Proposition 5.1

$$\begin{aligned} \mathbf{U}_t - \mathbf{1}k = \tilde{\Phi}(t, t_0)\mathbf{U}^0 &+ \int_{t_0}^{t} \tilde{\Phi}(t, s)G(s)d\mathbf{B}_s \\ &+ \mathbf{1c}^T \int_{t}^{\infty} G(s)d\mathbf{B}_s \end{aligned}$$

Also by the properties of Itô's integral and the Cauchy-

Schwarz inequality we obtain:

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

By assumption $g_{ij}^2(t)$ vanishes. Now, $\mathbb{E}\big[||\mathbf{U}_t - \mathbf{1}k||_2^2\big]$ is bounded from above by three terms, each of which converges to zero as $t \to \infty$: the first because, by Assumption 3.4, $\theta > 0$, 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, $\varepsilon > 0$

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

it is an easy exercise to show that all of the terms that bound $\mathbb{E}\big[||\mathbf{U}_t - \mathbf{1}k||_2^2\big]$ from above in Eq. (5) 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 $\mathbf{1}k \in \Delta$ follows (see Theorem 4(c) of §7.2 in [4]). Finally,

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

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 [4]).

6 Discussion

Recent results in consensus systems allowed us to reconsider fundamental topics of the stochastic consensus problem in a broader framework.

On the one hand, we discussed the discrete time consensus problem, and by invoking a general measure theoretic framework we outlined the crucial connection between lower bounds of the connectivity weights and the switching connectivity. Whenever the switching signal is probabilistic and the connectivity weights are not bounded from below, consensus can be achieved only with some positive probability. If we strengthen our results to uniformly bounded weights then our setting reproduced the results from several, seemingly distant, types of stochastic formulation of the problem, in an elegant and unified way.

It should be noted here that $P(n)$ as defined in (3.5) is *not* a stationary process as by construction the measure μ does not concern the weights p_{ij}^n. The stationarity property can be observed in the quantity $\mathbb{G}_{P(n)}$ which as we mentioned above is the only key feature for the stability analysis when the weights are uniformly lower bounded away from zero.

The approach above can 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, [9]. Then one can read these dynamics as stochastic and consider the consensus problem with communication topology driven by chaotic signals.

Finally, we note that all the aforementioned results can be delivered in a continuous time setting, which due to space limitation was omitted and the reader is kindly referred to [15].

Next we switched to the study a standard type of stochastic perturbations of a nominal general linear consensus system in continuous time. The new convergence results allow for a more generalized study of flocking models in the literature [5, 1, 3] as we did not need any symmetry assumption on the connectivity weights as well as the stability in variation argument we implemented allowed for time varying diffusion coefficients as well as a closed expression of the consensus point.

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 $\sum_{i,j} c_i^2 \int_{t_0}^{\infty} g_{ij}^2(s)ds$.

In [15] we also provide a more elaborate extension of (2.1) with state dependent diffusion coefficients. It is in our immediate research interest to extend our analysis to purely state-dependent connectivity weights and analyze our (non-linear) model by deriving sufficient conditions among the initial data for asymptotic stochastic flocking in the same spirit as the deterministic alternative already proved in [15].

References

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

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

[3] F. Cucker and E. Mordecki, *Flocking in noisy environments*, J. Math. Pures Appl., 89 (2008), pp. 278–296.

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

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

[6] D. J. Hartfiel, *Markov Set-Chains*, vol. 1695, Springer, 1998.

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

[8] A. Jadbabaie, J. Lin, and A. S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Transactions on Automatic Control, 48 (2003), pp. 988–1001.

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

[10] L. B. Koralov and Y. G. Sinai, *Theory of Probability and Random Processes*, Springer, 2 ed., 2007.

[11] I. Matei, J. S. Baras, and C. Somarakis, *Convergence results for the linear consensus problem under markovian random graphs*, SIAM J. Control Optim., (2013), pp. 1574–1591.

[12] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, Princeton University Press, 2010.

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

[14] M. Porfiri and D. L. Stilwell, *Consensus seeking over random weighted directed graphs*, IEEE Transactions on Automatic Control, 52 (2007), pp. 1767–1773.

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

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

[17] A. Tahbaz-Salehi and A. Jadbabaie, *A necessary and sufficient condition for consensus over random networks*, IEEE Transactions on Automatic Control, 53 (2008), pp. 791–795.

[18] ——, *Consensus over ergodic stationary graph processes*, IEEE Transactions on Automatic Control, 55 (2010), pp. 225–230.

[19] P. Walters, *An introduction to Ergodic Theory*, vol. 79, Springer, 1982.