THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2012-03

Linear Processes Under Vanishing Communications

The Consensus Algorithm

Christoforos Somarakis

Linear Processes Under Vanishing Communications
The Consensus Algorithm
TECHNICAL REPORT

Christoforos Somarakis

April 7, 2012

Abstract

In this report, we revisit the classical multi-agent distributed consensus problem under the dropping of the general assumption that the existence of a connection between agent implies weights uniformly bounded away from zero. We reformulate and study the problem by establishing global convergence results both in discrete and continuous time, under fixed, switching and random topologies.

1 Introduction

Consensus problems arise in many instances of collaborative control of multi-agent complex systems; where it is important for the agents to act in a distributed whereas coordinated manner, [15, 7, 11], to name a few. However, in the vast majority of all relevant works there is a fundamental assumption: The exchange of information among any two communicating nodes occurs under established connection with a time varying weight that is, however, uniformly bounded away from zero. This ensures the applicability of a large number of analytical tools from linear algebra, algebraic graph theory, probability theory etc. When running a distributed algorithm on the network the main consensus result suggests that an assumption of strong connectivity in principal ensures convergence to a common value through a repeated convex averaging of states.

In this report, we revisit the classical problem from the elementary static case, to switching signals as well as random failures versions, where we study and establish results for global convergence when the weights are allowed to vanish. These results suggest that the rate at which weights can vanish should not be faster than a critical value.

2 Notations and Definitions

In this section we will discuss the general theoretical framework in which we will establish our results. In the following $[k] = \{1, \ldots, k\}$, $\mathbb{Z}$ is the set of integers, $\mathbb{Z}$ is the set of naturals, $\mathbb{R}$ the set of real numbers. By $|| \cdot ||_p$ we denote the p-vector norm and $| \cdot |_\infty$ the infinity norm in $\mathbb{R}^k$, $k \in \mathbb{N}$. All vectors lie are considered column, unless otherwise stated. By $\mathbb{1}$ we understand the $k-$dimensional vector with all entries equal to 1. The *agreement* or *consensus* space $\mathcal{C}$ is defined as the subset of $\mathbb{R}^n$ such as

$$\mathcal{C} = \{x \in \mathbb{R}^n : x_1 = x_2 = \cdots = x_n\}$$

A *rank*-1 matrix A is such that it has identical rows. Obviously $Ax \in \mathcal{C}\ \forall x \in \mathbb{R}^n$.

2.1 Graph theory

By a *weighted undirected graph* $\mathbb{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ where $\mathcal{V} = [k]$ is the set of vertices, $\mathcal{E} = \{(i,j) : i,j \in \mathcal{V}\}$ the set of edges and $\mathcal{W} = \{w_{ij} : (i,j) \in \mathcal{E}\}$ the set of weights on edges. The *degree* $l = l_i$ of a node i is the number of adjacent edges to i. The graph $\mathbb{G}$ is connected if for any two vertices i,j there is a path of edges $(l_k, l_{k-1})|_{k=0}^{m}$ such that $l_0 = i$ and $l_m = j$. A graph $\mathbb{G}$ is called *simple* if it has no self-loops. The classes of graphs of most interest for this work are: the *complete* graph, K_n, the *path* graph, P_n, the *star* graph, S_n and the *cycle* graph, C_n.

2.1.1 Matrix representation of graphs

The *adjacency matrix* $A = [a_{ij}]_1^k$ of $\mathbb{G}$ is a matrix with $a_{ij} = w_{ij}$ if $(i,j) \in \mathcal{E}$ and 0 otherwise. The *incidence matrix* $\underline{A}$ is when $a_{ij} = \pm 1$ for $(i \leftarrow j / i \rightarrow j) \in \mathcal{E}$. The *degree matrix* $D := \mathrm{Diag} \sum_{j=1 \dots l_i} w_{ij}$. The Laplacian of $\mathbb{G}$ is the matrix defined as $L := D - A$. The spectrum (i.e. the set of eigenvalues) of L yields valuable information about the properties of $\mathbb{G}$, most of them are summarized in the following proposition.

Proposition 1. *The Laplacian of a graph $\mathbb{G}$, L has the following properties*

1. *L is symmetric.*
2. *L is positive semi-definite.*
3. *$\lambda_1 = 0$ is an eigenvalue of L associated with the left eigenvector $\mathbb{1}$*
4. *The spectrum of L can be ordered as $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$*
5. *The number of connected components of $\mathbb{G}$ equals the number of zero eigenvalues of L.*

Proof. All these are well known result from algebraic graph theory. See for example [3, 2, 10] □

2.2 Non-negative matrices

Since a graph can be represented as a matrix the machinery from linear algebra will be used. More specifically since these matrices have non-negative entries the powerful theory of non-negative, stochastic matrices with the fundamental Perron-Frobenius theorem will come at hand. Let's review some results on the theory of non-negative matrices.

A *non-negative (positive)* matrix $A = [a_{ij}]$ is such that $a_{ij} \geq 0 (> 0)$.

A *stochastic* matrix is a non-negative matrix $A = [a_{ij}]$ such that $\sum_j a_{ij} = 1\ \forall i$, so $\lambda = 1$ is an eigenvalues and $\mathbb{1}$ is the corresponding left eigenvector.

A *primitive* matrix A is such that A^m is positive for some $m \in \mathbb{N}$.

An *ergodic* matrix A is such that $\lim_{m \to \infty} A^m$ exists and is a rank-1 matrix, (i.e. matrix with identical rows).

It is easy to see that for any non-negative matrix with rows sum up to zero, say W, the matrix $G = I - \epsilon W$ is stochastic for some $\epsilon > 0$ small enough.

A fundamental result of the theory of non-negative matrices is the theorems of Perron-Frobenius (for example see [13]).

2.2.1 The Coefficient Of Ergodicity

The main analytical tool for studying products of matrices is the *coefficient of ergodicity.*

Definition 1. *For a stochastic matrix* $A = [a_{ij}]$*, the quantity*

$$\tau(A) := \max_{||z||_1=1, z'\mathbb{1}=0} ||A'z|| = \frac{1}{2}\max_{i,j}\sum_k |a_{ik} - a_{jk}| = 1 - \min_{i,j}\sum_k \min\{a_{ik}, a_{jk}\} \tag{1}$$

is called the coefficient of ergodicity of A*.*

Note that in (1) we are using dummy variables. Unless otherwise stated, the sum variables will always be dummy and should not be confused with explicitly defined global variables.

One can think of τ either as a vector norm maximized over $\mathbb{R}^n \backslash \mathcal{C}$ or a eigenvalue bounds expressed in terms of a deflated matrix, with the deflation approximating the dominant spectral projector. There are many different definitions and names for τ all in applicable for different spaces. For a recent thorough review on the subject see [6]. The following proposition introduces some basic properties of τ.

Proposition 2. *For* A *a stochastic matrix we have the following properties:*

1. $0 \leq \tau(A) \leq 1$
2. $\tau(A) < 1$ *if and only if every pair of rows* α, β *of* A *have a common position* k *such that* $a_{\alpha k}a_{\beta k} > 0$.
3. $\tau(A) = 0$ *if and only if* A *is a rank* 1 *matrix.*

Proof. 1. It follows from the definition and the assumption on A.

2. ($\Rightarrow$) From the definition

$$\min_{i,j}\sum_k \min\{a_{ik}, a_{jk}\} > 0$$

and this implies

$$\sum_k \min\{a_{ik}, a_{jk}\} > 0$$

for all i, j so that $\min\{a_{ik}, a_{jk}\} > 0$ for some k or equivalently $a_{ik}a_{jk} > 0$.

($\Leftarrow$) By contraposition. If there exist i, j rows such that $a_{ik}a_{ik} = 0$ for all k then take this path of zeros and achieve a sum equal to zero. By assumption on A this is a minimum and of course $\tau(A) = 1$.

3. ($\Rightarrow$) the minimum over every pair of rows i, j gives a sum equal to one. By the stochastic form of the matrix the result follows.
($\Leftarrow$) rank 1 matrix implies rows to be equal elementwise. By definition of τ and the stochasticity of the matrix, the result follows.

□

Another important property is illustrated in the next proposition

Proposition 3. *For A and B stochastic matrices*

$$\tau(AB) \leq \tau(A)\tau(B) \tag{2}$$

Proof. See [13, 4] □

So the coefficient of ergodicity is a norm that measures the amount of contraction of an operator matrix over the disagreement space and recognizes a matrix with identical rows. The family of operators that are recognizable by τ play an important role.

2.2.2 On Scrambling Matrices

From the discussion of [13] [§1.2, 4.3, 4.4] we draw the following useful definitions:

Definition 2. *The stochastic matrix A such that $\tau(A) < 1$, is called scrambling.*

Equivalently, A is scrambling if any two of it's rows have at least one positive element in a coincident position.

Definition 3. *The indices of a stochastic matrix which communicate with each other are called essential and form equivalence classes, called essential classes. An $n \times n$ stochastic matrix is said to be regular if its essential (i.e. persistent) indices form a single essential class which is aperiodic.*

We denote this set by G_1. Note that all the stochastic matrices we are dealing with, here, are aperiodic as they have strictly positive diagonals.

Definition 4. *An $n \times n$ stochastic matrix P is called a Markov matrix if at least one column of P is entirely positive.*

We denote this set by M.

Definition 5. *An $n \times n$ stochastic matrix $P \in G_2$ if*

- $P \in G_1$
- $QP \in G_1, \ \forall \ Q \in G_1$

Definition 6. *A stochastic matrix P belongs to G_3 if $\tau(P) < 1$ i.e. if given two rows i, j there is at least one column k such that $p_{ik} > 0$ and $p_{jk} > 0$. Such a matrix is called scrambling.*

Another equivalent definition of scrambling matrices is this:

Definition 7. *An $n \times n$ stochastic matrix P is called scrambling if no two rows are orthogonal.*

Theorem 1. $M \subset G_3 \subset G_2 \subset G_1$

Proof. See [13]. □

Corollary 1. *For any stochastic matrix Q, QP is scrambling, for fixed scrambling P.*

Proof. See [13]. □

Lemma 1. *If $P \in G_3$ ($P \in M$) then $PQ, \ QP \in G_3$ ($PQ, \ QP \in M$) for any stochastic matrix Q.*

Proof. See [13]. □

Matrices in G_3 thus have two special properties

1. It is easy to verify whether or not a matrix $\in G_3$
2. if all P_k are scrambling then

$$T_{p,k} = P_{p+1}P_{p+2}\cdots P_{p+k} \ \in \ G_1, \quad \forall\, p \geq 0, k \geq 1$$

Lemma 2. $P, Q \in G_2 \Rightarrow PQ,\ QP \in G_2$.

Proof. See [13]. □

Thus G_2 is closed under multiplication but G_1 is not.

The following theorem comes at hand when A is scrambling.

Theorem 2. *Let w be a non-negative vector and A a stochastic matrix. If $z = Aw$ then*

$$\max_i z_i - \min_i z_i \leq \tau(A)\big(\max_i w_i - \min_i w_i\big) \tag{3}$$

Proof. See [4]. □

3 The discrete time problem

Consider k agents lying in a d-dimensional Euclidean space. At each time each agent will update it's status $z_i(t) \in \mathbb{E}^d$ according to

$$z_i(t+1) - z_i(t) = \sum_{j=1}^{k} a_{ij}(t)(z_j(t) - z_i(t)) \qquad i = 1, \ldots, k \tag{4}$$

Assumption 1. *The connectivity functions as functions of time are defined as follows $a_{ij}(t) : \mathbb{Z}_+ \to \mathbb{R}_+$ such that*

$$a_{ij}(t) < \tfrac{1}{k-1} \tag{5}$$

$$a_{ij}(t) \in \omega(t^{-\alpha}) \tag{6}$$

where $\alpha \geq 0$, t is time and k the number of the agents.

Remark 1. *Assumption* (6) *says that the weights dominate an appropriate function of time that is the form $t^{-\alpha}$ (i.e. a function of time that does not violate* (5)*). This in general implies that ψ is free to vanish asymptotically, but not faster than $t^{-\alpha}$.*

Remark 2 (Definition of $\omega(\cdot)$)**.** *It is important to clarify this notation, for the sake of the analysis that will follow.*

Definition 8. *We say that a function $f(t)$ is asymptotically dominant to $g(t)$ as $t \to a$ and write $f(t) \in \omega(g(t))$ if $\lim_{t\to a} f(t)/g(t) = \infty$.*

We understand that in our case $a = \infty$, and the limit says that $\forall\ , i, j\ \exists \mathcal{T} : a_{ij} \geq ct^{-\alpha}\ \forall\ t \geq \mathcal{T}$ and in our case we also demand $c = c(k) < \mathcal{T}^{\alpha}/(k-1)$ so that Assumption 1(i) is satisfied.

With these in mind we may conclude that for appropriately chosen $\mathcal{T}$ for any constant r we have $\omega(rg(t)) = r\omega(g(t)) = \omega(g(t))$.

Remark 3. *Note that Assumption 1 generalizes the classic consensus problem in the sense that $\alpha = 0$ we are exactly back there.*

We make the following observations regarding (4). For any i we rewrite as

$$z_i(t+1) = \big(1 - \sum_{j=1}^{k} a_{ij}(t)\big) z_i(t) + \sum_{j=1}^{k} a_{ij}(t) z_j(t)$$

or in the matrix form

$$z(t+1) = (G(t) \otimes I_d) z(t) \tag{7}$$

where $G(t) := I_k - L_x$, I_k is the k-dimensional identity matrix L_x is the Laplacian

$$L_x = D_x - A_x \tag{8}$$

Also, $\otimes$ is the Kronecker product. Since for A, B, C, D matrices $(A \otimes B)(C \otimes D) = AC \otimes DB$ we conclude that we can work in the one-dimensional space without loss of generality. So from now own, unless otherwise stated, we take $d = 1$. The following lemma is trivial and helps us clarify a silent assumption.

Lemma 3. *Given any vector w and a stochastic matrix A, Aw can be written as the sum of a nonnegative vector and a nonpositive constant vector.*

Proof. For any column vector w trivially write $w = (w + \min_i\{w_i\}\mathbb{1}) - \min_i\{w_i\}\mathbb{1} := u + z$. Obviously u is nonnegative and since the effect of any such z is A-invariant the result follows. . □

Lets rewrite (4) in the k-dimensional form

$$z(t+1) = G(t) z(t) \tag{9}$$

where

$$G(t) := \begin{pmatrix} 1 - \sum_{j\neq 1} a_{1j} & a_{12} & \cdots & a_{1k} \\ a_{21} & 1 - \sum_{j\neq 2} a_{2j} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & 1 - \sum_{j\neq k} a_{kj} \end{pmatrix}_{a_{ij} = a_{ij}(t)} \tag{10}$$

Given the initial vector $z(0)$, the general solution at time t of (9) is of course

$$z(t) = G(t-1)G(t-2)\cdots G(0)z(0)$$

Definition 9. *We say that the system (9) converges to unconditional consensus if for any initial vector $z(0)$ such that $\inf_{\bar{z} \in \mathcal{C}} ||z(t) - \bar{z}||_p \to 0$ as $t \to \infty$*

So the root to convergence passes through the dynamics of backward products of non-homogeneous matrices with weights that are not uniformly bounded away from zero. The handling of products of matrices is a rather intricate subject. Seneta in [13] discusses the different notions of ergodicity that appear depending on the the products, i.e. whether they are forward or backward. The *weak ergodicity* refers to the case where given a product $T_{p,k} = \{t^{(p,r)}\}_{i,j}$ of matrices we have $t^{p,r}_{i,s} - t^{p,r}_{j,s} \to 0$ as $r \to \infty$ for any i, j, s, p. On the other hand, *strong ergodicity* is such that $lim_r t^{(r,p)}_{i,j}$ exists and is independent of i. In case of forward products of matrices, these notions are distinct and there would be no hope for our case to deduce a strong ergodicity result as it is indeed the uniform boundedness of the positive weights a necessary condition for forward products to obtain strong ergodicity. In case of backward products however (our case) the two notions of ergodicity coincide and there is no restriction to bounded weights and hence we are consistent with Definition 9.

We therefore simply exploit the structure of $G(t)$ in order to get estimates of $\tau(G(t))$ as t grows large. In the following elementary example we will show that the role of exponent α is critical.

Example Consider the 2-D dynamical system

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} 1-f(t) & f(t) \\ g(t) & 1-g(t) \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \tag{11}$$

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

$$|x(t+1) - y(t+1)| = (1 - f(t) - g(t))(x(t) - y(t)) = \delta \prod_{i=0}^{t} (1 - f(i)) \to C \sin(\pi\sqrt{K_f + K_g}) \qquad t \to \infty \tag{12}$$

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

3.1 Upper bounds of $\tau(G(t))$

The definition of τ requires to find the two rows i, j that minimize the sum

$$S(i,j) = \sum_{n=1}^{k} \min\{a_{in}, a_{jn}\} \tag{13}$$

the structure of $G(t)$ implies that for every i, j the sum must have three forms depending on the number of diagonal elements that will be included in the sum (i.e. 0,1,2). So we have to take three cases:

- 1 diagonal element:

  Assume, without loss of generality, that the diagonal element is from the i^{th} row. So let also i^{th} row to contribute with $l-1$ (non-diagonal) elements and j^{th} row to contribute with the remaining $k-l$, for some $2 \leq l \leq k-1$ ($l = 0, k$ implies $\tau = 0$ and $l = 1$ means that we must have two diagonal elements). Then

$$S(i,j) = 1 - \sum_{\{k-l+1\}\text{terms of } i^{th} \text{ row}} a_{is} + \sum_{\{k-l\}\text{terms of } j^{th} \text{ row}} a_{js} \tag{14}$$

Note that under no condition can the sums cancel each other. Moreover the first sum is in absolute value greater than the second sum. In view of the assumptions of $a_{i,j}$ it follows in this case that

$$\begin{aligned} \tau(G(t)) &= 1 - \min_{i,j} S(i,j) \\ &= \sum_{\{k-l+1\}\text{terms of } i^{th} \text{ row}} a_{is} - \sum_{\{k-l\}\text{terms of } j^{th} \text{ row}} a_{js} \\ &< \frac{k-l+1}{k-1} - (k-l)\omega(t^{-\alpha}) \end{aligned} \tag{15}$$

which is strictly less than one.

- 2 diagonal elements:

  Using the same arguments we calculate $S(i,j)$. We understand that if i^{th} row contributes with $l \geq 1$ elements then 1 must be the diagonal and $l-1 \leq k-2$ the remaining ones. Then the j^{th} must contribute with $k-l$ elements 1 of which is diagonal and $k-l-1$ the remaining ones. Then

  $$S(i,j) = 2 - \sum_{\{k-l+1\}\text{terms of } i^{th} \text{ row}} a_{is} - \sum_{\{l+1\}\text{terms of } j^{th} \text{ row}} a_{js} \tag{16}$$

  so that:

  $$\tau(G(t)) = \sum_{k+2 \text{ terms}} a_{i.js} - 1 < \frac{3}{k-1} \tag{17}$$

  the cases $k=2$ and $k=3$ are pathological and will be considered below.

- 0 diagonal elements:

  In this case

  $$\tau(G_q(t)) = 1 - \sum_{k\text{terms}} a_{i,js} \leq 1 - k\omega(t^{-\alpha}) \tag{18}$$

The two pathological cases are:

1. $k=2$: The matrix is of the form

   $$G(t) := \begin{pmatrix} 1-a_{12} & a_{12} \\ a_{21} & 1-a_{21} \end{pmatrix} \tag{19}$$

   so that $\tau(G) = a_{12} + a_{21} - 1 < 1$ by Assumption 1 or $\tau(G) = 1 - a_{12} - a_{21} < 1 - 2\omega(t^{-\alpha})$.

2. $k=3$: The matrix is of the form

   $$G(t) := \begin{pmatrix} 1-a_{12}-a_{13} & a_{12} & a_{13} \\ a_{21} & 1-a_{21}-a_{23} & a_{23} \\ a_{31} & a_{32} & 1-a_{31}-a_{32} \end{pmatrix} \tag{20}$$

   and in the case of two diagonal elements $\tau < \frac{1}{2}$.

This concludes our calculation.

Remark 4. *From the discussion above we considered some arbitrary by finite* $t \in \mathbb{Z}_+$*.*

In view of Remark 2 we are ready to state the following important proposition:

Lemma 4. *Assume* $k \geq 2$ *agents, which update their speed according to* (9)*. Then the coefficient of ergodicity satisfies*

$$\tau(G(t)) \leq 1 - \omega(t^{-\alpha}) \tag{21}$$

for t *large enough.*

3.2 A first convergence result

Theorem 3. *Under Assumption 1, the system* (9) *converges to consensus* $\forall \alpha \in [0, 1]$

Proof. It is enough to show that the maximum coordinate of $z(t)$ approaches the minimum as $t \to \infty$. Set $m(t) := \max_i z_i(t) - \min_j z_j(t)$ and apply Theorem 1 to get

$$m(t+1) \leq \tau(G(t))m(t) \leq \prod_{s=0}^{t} \tau(G(s))m(0)$$

as well as that since $m(t) \geq 0$, the limit exists $\lim_t m(t)$ exists. Finally, in view of proposition 3 we easily calculate:

$$\begin{aligned} \lim_{t\to\infty} m(t) &\leq \lim_{t\to\infty} \prod_{s=0}^{t} \tau(G(s))m(0) \\ &\leq C \lim_{t\to\infty} \prod_{s=\mathcal{T}}^{t} (1 - \omega(s^{-\alpha}))m(0) \\ &\leq C \lim_{t\to\infty} \prod_{s=\mathcal{T}}^{t} e^{-\omega(s^{-\alpha})} m(0) \\ &\leq C \lim_{t\to\infty} e^{-\sum_{s=\mathcal{T}}^{t} \omega(s^{-\alpha})} m(0) \\ &= 0 \end{aligned} \tag{22}$$

for $0 \leq \alpha < 1$ and $\mathcal{T}$ goes according to the definition of ω notation and is independent of t.

□

In subsequent sections we will generalize the case of switching topologies to families that include more connectivity topologies and derive sufficient conditions for consensus. For the moment let's review an important application.

3.3 Static Topologies

In the analysis above we assumed that essentially $G(t)$ is an adjacency matrix of K_k with weights that never vanish. However Lemma 4 still holds if simply $G(t) \in G_3$ $\forall t$. This allows one to actually digress from the complete graph to a graph that corresponds to a scrambling matrix.

The coefficient of ergodicity allows a number of "failures" (i.e. no connectivity , zero weights), such that the general upper bound in Lemma 3 is not violated. Is it possible to "remove" connections while preserving this estimate? The answer is affirmative and comes from Proposition 2.

Proposition 4. *For any $t < \infty$ assume the matrix $G(t)$ with non-negative weights such that the only zero elements are those where there is broken connection. Then $\tau(G(t)) = 1$ if and only if there exist two nodes with broken connection and with no common neighbours.*

Proof. By Proposition 2, $\tau(G(t)) = 1$ if and only if there exists two rows, say i, j with a path of zeros from 1 to k. This is equivalent to say that for every $s :\ a_{is} > 0 \Rightarrow a_{js} = 0$ and vice versa. Equivalently, i, j do not talk to each other and have no common neighbours. □

From now on we will write $\tau < 1$ to describe the connectivity graph which fulfils Proposition 3.

One can easily inspect the families of graphs that are compatible with proposition 4. They should have high connectivity or several triangle loops. There is however a family of minimum number of edges that admit $\tau < 1$.

Proposition 5. *The star graph S_k is the one topology with the minimum number of connections that ensures $\tau < 1$.*

Proof. A star graph satisfies $\tau < 1$. It suffices to prove that it just demands the minimum number of connections. The node i must at be least adjacent to itself and another node (it is the necessary condition for the speed consensus process- connectedness). This establishes a lower bound on the number of connections. If there is another node j, not adjacent to i and with no common adjacent node then i needs another connection to service j. So the graph cannot be a star. □

3.4 Switching Topologies

This can be also generalized in case that a switching signal, $\sigma(t) : \mathbb{Z}_+ \to [m]$ that controls the connectivity of graph $\mathcal{G}_{\sigma(t)} \in \{\mathcal{G}_i\}_{i=1}^m$, so that the family is in G_3. What is actually need to be determined is the existence of a $\mathcal{T}$ such that in any case the analysis in the proof of lemma 4 can be carried out. Such appropriate $\mathcal{T}$ always exists since the cardinality of the set $\{\mathcal{G}_i\}$ is finite.

In this section we will review some variations of the original problem that yield the same results. More specifically we will show that the assumption of persistent scramblingness is too strong since convergence results can also be obtained under weaker assumptions.

3.4.1 Scrambling recurrence

In this section we will discuss the possibility to relax the scrambling property of the connection matrix corresponding to $G(t)$ for every $t \in \mathbb{Z}_+$. Consider the graph $\mathbb{G}(t) = (V, E(t))$ and the corresponding $\mathbb{G}(t) \in G_1$ where connections vary with time. From the discussion so far one would come up with the following assumption

Assumption 2. *$G(t) \in G_3$ for infinitely many t.*

However this is not enough and here is why

Proposition 6. *[15] Under Assumption 2 there a sequence $\{t_i\}_i \geq 1$ with $\lim_i t_i = \infty$ such that solutions of* (DS) *that do not converge.*

Assumption 3. *The connectivity functions as functions of time are defined as follows: $a_{ij}(t) : \mathbb{Z}_+ \to \mathbb{R}_+ \cup \{0\}$*

$$\begin{aligned} a_{ij}(t) &< \frac{1}{k-1} \\ a_{ij}(t) &\in \omega(t^{-\alpha}) \cup \{0\} \end{aligned} \tag{23}$$

Assumption 4. *And there is some* $B > 0$ *such that for any solution* $z(t)$ *of* (9) *and all* t *the matrix product*

$$G(t)G(t+1)\cdots G(t+B-1)$$

contains at least a scrambling matrix.

Under Assumption 3 we can prove unconditional consensus using the same techniques as above. The trick here is to observe that there is a strictly increasing subsequence $\{t_i\}_i$ with $t_1 > 0$, $\lim_i t_i \to \infty$ and $|t_i - t_{i+1}| \leq B$. Then of course $t_i \leq (i-1)B + t_1$ for $i \geq 2$ and

$$\sum_i \frac{1}{t_i^\alpha} \geq \sum_i \frac{1}{((i-1)B + t_1)^\alpha} = \infty \quad \forall \alpha \in [0, 1] \tag{24}$$

So

$$m(t) \leq \prod_i \tau(G(t_i))m(0) \leq e^{-\sum_i \omega(t_i^{-\alpha})}m(0) = 0 \tag{25}$$

So as a conclusion to theorem 3

Theorem 4. *Under Assumptions 3 and 4, unconditional consensus can be achieved for* (9)

3.5 Beyond Recurrent Scramblingness

The main results of the consensus dynamics are based on two main assumptions: The assumption of strong connectivity and the assumption of uniform lower bounds for the connection weights (when they are established). In this section we take a step further investigating consensus under the assumption of plain strong connectivity. It would be very interesting to establish results without the scrambling matrix assumption. Let's review some well known results.

Definition 10. *A matrix* A *is said to be SIA if* A *is stochastic, indecomposable*[1] *and aperiodic.*

Lemma 5. *Let* $\mathbb{M}$ *be a set of SIA matrices. There exists an integer* $\gamma > 0$*, known as the scrambling index, such that any* n*-length matrix product of* $n \geq \gamma$ *members picked from* $\mathbb{M}$ *is scrambling.*

Proof. See [13] □

The following proposition gives an upper bound on γ.

Proposition 7. *For* $\mathbb{G} = (V, E)$ *with* $|V| = k$*, we have that*

$$\gamma \leq \begin{cases} \frac{k}{2} & k, \textit{even} \\ \frac{k-1}{2} & k, \textit{odd} \end{cases}$$

Proof. Take the worst case scenario of a strongly connected graph, that is, the path graph, P_k. Then the first product of scrambling transition matrix is that of a Markov matrix after $k/2$ of $(k-1)/2$ steps. □

Finally we connect scrambling matrices with graphs containing a spanning tree.

[1] Instead of "indecomposable" another commonly used term is this of "irreducible".

Lemma 6. *If the graph corresponding to a stochastic matrix A has a spanning tree and a self-link at one of its root vertices, then A is SIA.*

Proof. See [13] $\square$

Using the above lemma's with the assumptions above the standard results obtain product of matrices that constitute scrambling matrices with uniform lower bounds. The combination of the uniform lower bounds and the SIA nature of the set of transition matrices yields to the existence of a scrambling index, γ. This enables τ to identify a global uniform upper bound which is strictly less than 1, so that exponential convergence to $\mathcal{C}$ is guaranteed. This is not the case though when the weights are not bounded from below.

It is desirable to establish a bound for an arbitrary product of SIA matrices which constitute a scrambling one. Let's consider an example, at first.

Example Consider a network of 4 agents with the following $G(t)$

$$G(t) = \begin{pmatrix} 1 - a_{12}(t) & a_{12}(t) & 0 & 0 \\ a_{21}(t) & 1 - a_{21}(t) - a_{23}(t) & a_{23}(t) & 0 \\ 0 & a_{32}(t) & 1 - a_{32(t)} - a_{34}(t) & a_{34}(t) \\ 0 & 0 & a_{43}(t) & 1 - a_{43}(t) \end{pmatrix} \tag{26}$$

It is easy to check that $\tau(G(t)) = 1$ for all t. However any product of such two matrices is scrambling. In particular for all t the matrix $G(t+1)G(t)$ is scrambling with the first and fourth row to have nonzero entries that sum up to

$$a_{12}(t+1)a_{23}(t) + a_{32}(t)a_{43}(t) \geq \omega((t+1)^{-2\alpha}) \tag{27}$$

In fact I can arrange the weights so that

$$\tau(G(t+1)G(t)) = 1 - a_{12}(t+1)a_{23}(t) - a_{32}(t)a_{43}(t) \leq 1 - \omega((t+1)^{-2\alpha}) \tag{28}$$

It can be easily seen that

$$\lim_t \tau(\prod_{i=1}^{t} G(i)) \leq \lim_t Ce^{-\sum_j^t j^{-2\alpha}} \tag{29}$$

which converges to zero for $0 \leq \alpha \leq 1/2$.

3.5.1 A more general result

Assume a static network with graph $\mathbb{G} = (V, E)$ that is strongly connected, so that it has a scrambling index γ. This means that any product of $G(t)G(t+1)\cdots G(t+\gamma-1) \in G_3$. (Be reminded that if A, B are SIA of the same degree then AB is SIA too).

Consider an element of the matrix product

Theorem 5. *Assume a static network with graph $\mathbb{G} = (V, E)$ that is strongly connected, so that it has a scrambling index γ. Then* (9) *associated with $\mathbb{G}$ will reach consensus under Assumption 3 for $\alpha \in [0, 1/\gamma]$.*

Proof. At first note that for any t_1, t_2, $G(t_1)G(t_2)$ is of the same form as (10). From the proof of Lemma 4 it suffices to check the case where the minimum path of elements contains no diagonal element. From this, one may use the observation that $a_{ij}(t_1)a_{lm}(t_2) \geq \omega(t_2^{2\alpha})$ for $t_1 < t_2$ to conclude by induction, to

$$\tau(G(t)G(t+1)\cdots G(t+\gamma-1)) \leq 1-\omega\big((t+\gamma-1)^{-\gamma\alpha}\big) \qquad t >> 1 \tag{30}$$

then the result follows as in Theorem 3. □

3.6 Random Graphs

In this section we wil study an important generalization of the consensus problem which is under random graphs. Relevant literature for random graphs is significantly large (see [5, 12, 9, 14]).

Consider a probability space $(\Omega_0, \mathcal{B}, \mathbb{P}_0)$, where Ω_0 is the set of $n\times n$ incident $(0-1)$ matrices with positive diagonals, $\mathcal{B}$ is the Borel $\sigma-$ algebra of Ω_0 and $\mathbb{P}_0$ a probability measure on Ω_0. Define the product probability space $(\Omega, \mathcal{F}, \mathbb{P}) = \prod_k(\Omega_0, \mathcal{B}, \mathbb{P}_0)$ and by Kolmogorov extension theorem there exists a measure $\mathbb{P}$ that makes its coordinates stochastically independent while preserving the marginal distributions. The elements of the product space have the following forms:

$$\begin{aligned} &\Omega = \{(\omega_1,\ldots,\omega_k,\ldots) : \omega_k \in \Omega_0\} \\ &\mathcal{F} = \mathcal{B}\times\mathcal{B}\times\cdots \\ &\mathbb{P} = \mathbb{P}_0\times\mathbb{P}_0\times\cdots \end{aligned} \tag{31}$$

It follows that the mapping $W_k : \Omega \to \Omega_0$ be the k^{th} coordinate function which for all $\omega \in \Omega$ is defined as

$$W_k(\omega) = \omega_k \tag{32}$$

and guarantees that $W_k(\omega)$ $k \geq 1$ are independent random $0-1$ matrices with positive diagonals and common distribution $\mathbb{P}$. Define now

$$G(W_t, t) : \Omega \times \mathbb{Z}_+ \to G_1 \tag{33}$$

to be the stochastic matrix of (DS) with $0-1$ connectivity between i and j and connectivity weight a_{ij}. In the following $G(W_t, t) := G(t)$ we take it as random variable of the product space such that $G_t,\ G_s$ are independent for $s \neq t$.

Define

$$\begin{aligned} &m(t) = \max_i x_i(t) - \min_j x_j(t) \\ &\mu(t) = \mathbb{E}[m(t)] \end{aligned} \tag{34}$$

The first is for every t a random variable and the second is it's expected value. We are interested in the probability of the event

$$A_t := \left\{ \sum_{s\geq t} m(s) \quad \text{converges} \right\} \tag{35}$$

Obviously, $A_1 \supset A_2 \supset \cdots$ and of course the occurence of A_t implies that the occurence of $\{\limsup_t m(t) = 0\}$. Such events are events of the tail σ-field.

Definition 11. *If* $\mathcal{F}'_k = \sigma(W_k, W_{k+1}, \dots)$ *then the* σ*-field defined as*

$$\mathcal{F}_\infty = \bigcap_k \mathcal{F}'_k$$

is the tail σ*-field of the sequence* $\{W_k\}$

Kolmogorov's $0-1$ law assures that under the independence assumption if $A \in \mathcal{F}_\infty$ then $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$. This means that we can derive sufficient conditions for almost sure consensus if we derive conditions for the summability of $m(t)$.

The following theorem is a trivial sufficient condition:

Theorem 6. *The stochastic version of* (9) *exhibits unconditional consensus if* $\mathbb{E}[G(t)] \in G_3$

Proof. We understand that

$$\begin{aligned} x(t+1) = G(t)x(t) \Rightarrow \\ \mathbb{E}[x(t+1)] = \mathbb{E}[G(t)x(t)] = \mathbb{E}[G(t)]\mathbb{E}[x(t)] \end{aligned} \tag{36}$$

by independence. So $\mu(t+1) \leq \tau(\mathbb{E}[G(t)])\mu(t)$ and the result follows as in the static case. We will give a note on the speed consensus: Given any filtration of $m(k)|_0^t$ we see that

$$\mathbb{E}[m(t+1)|m(0), \dots, m(t)] \leq \tau(\mathbb{E}[G(t)])m(t) \leq (1 - \omega(t^{-\alpha}))m(t) \tag{37}$$

since $\sum_k \omega(k^{-\alpha}) = \infty$ for $0 \leq leq < 1$ we have that $m(t) \to 0$ with probability 1, as a direct application of the super-martingale convergence theorem. □

Note that we cannot really give a necessary condition in terms of the coefficient of ergodicity, exactly because the non-summability of $m(t)$ does not imply non-flocking. It follows that a condition of the form $\tau(\mathbb{E}[G(t)]) = 1$ for all t implies inconclusiveness.

A necessary (but rather unhelpful) condition is this of the connectivity of $\mathbb{E}[G(t)]$. If $|\lambda_2(\mathbb{E}(G(t))| = 1$ then speed consensus is an event that happens almost never [14].

3.7 Application in flocking dynamics

Consider a population of k birds leaving a d-dimensional Euclidean space $\mathbb{E}^d$. At each time $t \in \mathbb{Z}_+$ every bird, i, has a vector of state $x_i(t)$ and a vector of velocity $v_i(t)$ in $\mathbb{E}^d$. Consider also the standard Euclidean norm, $||\cdot||$, as well as the infinity norm $|\cdot|_\infty$.

Definition 12. *Consider the population of k birds with positions and velocities as defined above. We say that we have* ***unconditional flocking*** *if* $\forall\ i, j \in \{1, \dots, k\}$ *and all initial positions and velocities, both the following two conditions hold*

$$\begin{aligned} &(i) \quad \lim_{t\to\infty} ||v_i(t) - v_j(t)|| = 0 \\ &(ii) \quad \sup_{0\leq t<\infty} ||x_i(t) - x_j(t)|| < \infty \qquad \textit{uniformly in } t \end{aligned}$$

Every bird adjusts it's velocity by adding to it a weighted average of the differences of it's velocity with those of the other birds.

$$v_i(t+1) - v_i(t) = \sum_{j=1}^{k} a_{ij}(v_j(t) - v_i(t)) \tag{38}$$

where $a_{ij} = \psi(||x_i - x_j||)$ is a positive strictly decreasing function. Assuming $x(t+h) = x(t) + hv(t)$, and taking $h = 1$ for convenience we obtain the dynamical system

$$\begin{aligned} x_i(t+1) &= x_i(t) + v_i(t) \\ v_i(t+1) - v_i(t) &= \sum_{j=1}^{k} a_{ij}(t)(v_j(t) - v_i(t)) \end{aligned} \tag{DS}$$

Given the Assumption 1. We make the following observations regarding (DS). For any i we rewrite as

$$v_i(t+1) = \big(1 - \sum_{j=1}^{k} a_{ij}(t)\big)v_i(t) + \sum_{j=1}^{k} a_{ij}(t)v_j(t)$$

or in the matrix form

$$v(t+1) = (G(t) \otimes I_d)v(t) \tag{39}$$

where $G(t) := I_k - L_x$, I_k is the k-dimensional identity matrix L_x is the Laplacian of the adjacency $k \times k$ matrix $A = [a_{ij}]$,

$$L_x = D_x - A_x \tag{40}$$

Also, $\otimes$ is the standard Kronecker product. Since for A, B, C, D matrices $(A \otimes B)(C \otimes D) = AC \otimes DB$ we conclude that we can work in the one-dimensional space without loss of generality. So from now own, unless otherwise stated, we take $d = 1$. The following lemma is trivial and helps us clarify a silent assumption.

Lemma 7. *Given any vector w and a stochastic matrix A, Aw can be written as the sum of a nonnegative vector and a nonpositive constant vector.*

Proof. For any column vector w trivially write $w = (w + \min_i\{w_i\}\mathbb{1}) - \min_i\{w_i\}\mathbb{1} := u + z$. Obviously u is nonnegative and since the effect of any such z is A-invariant the result follows. . □

Lets rewrite (39) in the one dimensional form $v(t+1) = G(x(t))v(t)$ where

$$G(x(t)) := \begin{pmatrix} 1 - \sum_{j\neq 1} a_{1j} & a_{12} & \cdots & a_{1k} \\ a_{21} & 1 - \sum_{j\neq 2} a_{2j} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & 1 - \sum_{j\neq k} a_{kj} \end{pmatrix}_{x(t)} \tag{41}$$

From the Assumption 1 and the structure of G in (41) we conclude that G is a doubly stochastic matrix for every $x(t)$. Hence Lemma 1 allows us to use non-negative values for the velocities and thus non-negative values of the positions of the birds.

Also, it follows from (41) that the update of the velocity is the convex combination of the velocities of the birds. Such an averaging effect decreases the maximum value and increases the minimum value of $v(t) = (v_1(t), \cdots, v_k(t))$. It follows that as time evolves $|v(t)|_\infty \leq |v(t-1)|_\infty$ is bounded. The next lemma is also trivial

Lemma 8. *Under assumption 1, the dynamics of* (DS) *admit a bounded velocity $v(t)$ and hence a sub-linear growth of $x(t)$.*

Proof. From the observation above $||v(t)||_2 \leq \sqrt{k}|v(t)|_\infty \leq \sqrt{k}|v(0)|_\infty$ and trivially $\forall\ i$

$$||x_i(t)|| = ||x_i(0) + \sum_{s=0}^{t-1} v_i(s)|| \leq ||x_i(0)|| + \sqrt{k}|v(0)|_\infty t \tag{42}$$

□

Lemma 2 helps us understand the second part of Assumption 1. The sub-linear growth of $x_i(t)$ implies the sub-linear growth of $||x_i(t) - x_j(t)||$ with time for all i, j.

Following Cucker-Smale we may assume that the connectivity weight between i and j birds can decrease as their distance, but not faster than linear. This assumption is mild from the point of view that in the classical consensus dynamics problem we require the connectivity weights to be bounded away from 0 uniformly, however this assumption should not be relaxed too much.

3.7.1 Convergence Results for the static problem

Flocking is about speed consensus Def.(i) with sufficiently fast rate so that the flock is not destroyed Def.(ii). Combination of these two requirement yields that that for every i, j birds

$$||x_i(t) - x_j(t)|| \leq ||x_i(0) - x_j(0)|| + \sum_{s=0}^{t-1} ||v_i(s) - v_j(s)|| \tag{43}$$

so it suffices for the speed differences between any two birds to be summable.

For any t we define $m(t) := \max_i v_i(t) - \min_j v_j(t)$ by Theorem 1 we know it satisfies

$$m(t+1) \leq \tau(G(t))m(t) \leq \prod_{s=0}^{t} \tau(G(s))m(0)$$

as well as that since $m(t) \geq 0$, the limit exists $\lim_t m(t)$ exists. Finally, in view of proposition 3 we easily calculate:

$$\begin{aligned} \lim_{t\to\infty} m(t) &\leq \lim_{t\to\infty} \prod_{s=0}^{t} \tau(G(s))m(0) \\ &\leq C \lim_{t\to\infty} \prod_{s=\mathcal{T}}^{t} (1 - \omega(s^{-\alpha}))m(0) \\ &\leq C \lim_{t\to\infty} \prod_{s=\mathcal{T}}^{t} e^{-\omega(s^{-\alpha})} m(0) \\ &\leq C \lim_{t\to\infty} e^{-\sum_{s=\mathcal{T}}^{t} \omega(s^{-\alpha})} m(0) \\ &= 0 \end{aligned} \tag{44}$$

for $0 \le \alpha < 1$ and $\mathcal{T}$ goes according to the definition of ω notation and is independent of t. It only remains to fulfil Def(ii). Note that for any $t > 0$

$$\begin{aligned}
\sum_{s=0}^{t} m(s) &< \sum_{s=0}^{\infty} m(s) \\
&\le \sum_{s=0}^{\infty} \tau(G(s))m(0) \\
&\le B(\mathcal{T}, m(0)) + m(0) \sum_{s=\mathcal{T}}^{\infty} \exp\{-\sum_{j=\mathcal{T}}^{s} \omega(j^{-\alpha})\}
\end{aligned} \tag{45}$$

The double sum converges as follows:

$$\begin{aligned}
\sum_{s=\mathcal{T}}^{\infty} \exp\{-\sum_{j=\mathcal{T}}^{s} \omega(j^{-\alpha})\} &\le \sum_{s=\mathcal{T}}^{\infty} \exp\{-\omega(s^{1-\alpha}) + (\mathcal{T}-1)\omega(s^{-\alpha})\} \\
&\le \exp\{(\mathcal{T}-1)\omega(\mathcal{T}^{-\alpha})\} \sum_{s=\mathcal{T}}^{\infty} \exp\{-\omega(s^{1-\alpha})\} \\
&\le D(\mathcal{T}) \sum_{s=\mathcal{T}}^{\infty} (e^{-c})^{s^{1-\alpha}}
\end{aligned}$$

The last sum is treated as follows:

$$\begin{aligned}
\sum_{s=\mathcal{T}}^{\infty} (e^{-c})^{s^{1-\alpha}} &\le \sum_{s=0}^{\infty} (e^{-c})^{s^{1-\alpha}} \\
&\le \int_0^{\infty} e^{-cs^{1-\alpha}} ds \\
&= c^{\frac{\alpha}{1-\alpha}} \int_0^{\infty} t^{-\frac{\alpha}{1-\alpha}} e^{-t} dt \\
&= c^{\frac{\alpha}{1-\alpha}} \Gamma\left(\frac{1}{1-\alpha}\right) < \infty, \qquad 0 \le \alpha < 1
\end{aligned} \tag{46}$$

It follows that the range of the speed of the flock is summable uniformly in time since it just depends on the initial velocities and the weight function such that $\mathcal{T}$ is independent of time. This means that any difference of the flock speed So we proved the following theorem

Theorem 7. *Given the system* (DS) *with weight functions according to Assumption 1, we have unconditional flocking for all* $0 \le \alpha < 1$.

Remark 5. *It would be interesting to ask what happens at* $\alpha = 1$. *All relevant studies note the need of additional assumptions, in the initial conditions. This of course means that the flocking is not unconditional. Indeed in such case the the range of the speed differences cannot admit such bounds as above. In any case the scope of this study is on the study of unconditional flocking.*

Remark 6. *The analysis so far is not based on the symmetry of* a_{ij}. *Although assumed symmetric by Assumption 1, this argument of symmetry was not used.*

Remark 7. *There is great resemblance with the Cucker-Smale flocking model. In the present work we neither assume a specific weight function nor symmetry.*

3.7.2 Random Failures

In this section we discuss the random failure version model of our system.

Flocking is about speed consensus Def.(i) with sufficiently fast rate so that the flock is not destroyed Def.(ii). Combination of these two requirement yields that that for every i, j birds

$$||x_i(t) - x_j(t)|| \leq ||x_i(0) - x_j(0)|| + \sum_{s=0}^{t-1} ||v_i(s) - v_j(s)|| \tag{47}$$

so it suffice for the speed differences of any bird to be summable. In a probabilistic environment we require a certain sequence of random variables to be summable.

Proposition 8. *$\sum_t m(t)$ converges with probability one if and only if $\sum_t \mu(t)$ converges to $\mathbb{R}_+$*

Proof. ($\Rightarrow$) $\sum_t m(t) < \infty$ w.p. 1. This implies $\sum_{t=1}^T m(t) < \infty$ and $\mathbb{E}[\lim_T \sum_{t=1}^T m(t)] = \lim_T \mathbb{E}[\sum_{t=1}^T m(t)] = \lim_T \sum_{t=1}^T \mathbb{E}[m(t)]$ by the Monotone Convergence Theorem.

($\Leftarrow$) It follows from the Chebyshev inequality and the fact that W_k is a finite variance process (W_k can only take finitely many difference values). The event of a diverging sum of random variables is a zero probability event. □

Remark 8. *The summability of $\mu(t)$ implies the following:*

- $\mu(t) \to 0$*: Convergence in probability by the Markov inequality and since $|\mu(t)|$ is bounded for every t we have convergence in r^{th} mean for every r.*
- *The observation above implies the summability of $\mathbb{P}(|m(t)| > \varepsilon)$ for arbitrary $\varepsilon > 0$, hence a.s. convergence.*

Given the population of k birds the number of possible graphs is $2^{\binom{k}{2}}$. From this family we separate the ones that satisfy $\tau < 1$ from the remaining graphs (all that satisfy $\tau = 1$). All of them, however, must have positive diagonals. It is reminded that G_3 is the family of the good (scrambling) graphs and that also if $P \in G_3$ then $QP \in G_3$ for any stochastic matrix Q.

Assume a renewal process $\{T_t\}_{t\in\mathbb{Z}}$ such that every $G(T_t) \in G_3$. If $X_t := T_t - T_{t-1}$ is the inter-renewal process then we assume that

Assumption 5. *There exist constants C and B independent of time such that*

$$0 < C < \mathbb{E}[X_t] < B < \infty \quad \forall\, t \tag{48}$$

All in all, we do not require neither stationarity nor independence of the renewal process. The above regularity assumption will be enough for our objectives which is to prove unconditional flocking under those random perturbations, that guarantee infinitely many good connections.

Solution Strategy It is all about the process $\mu(t)$. From the discussion above we understand that this limit exists a.s. since for any $T > 0$ and any given filtration $\mathcal{F}_T$, $c_t := \{\mu(t)|\mathcal{F}_t\}$ is a non-increasing, non-negative sequence. The velocities have no-where to go but to *collapse*. Every bird *talks* to some other bird infinitely often and more or less it suffices to prove that any connectivity regime takes a subset of the flock and suppresses it's velocities. This step is non-trivial and of most significance.

Note that although the time interval process may not necessarily have a constant mean, it is however well bounded. Obviously,

$$\limsup_{n} \mathbb{E}[X_t] \leq B \tag{49}$$

and for $\mathcal{N}(t)$ the number of renewals by time t, the standard renewal theorem gives the bound

$$\limsup_{t} \frac{\mathbb{E}[\mathcal{N}(t)]}{t} \leq B^{-1} \tag{50}$$

Applying Theorem 1 we calculate the asymptotic behaviour of $\mu(t)$. By recalling that for any given filtration the limit of $\mu(t)$ exists we work as follows:

$$\begin{aligned}
\lim_{t\to\infty} \mu(t) &= \limsup_{t} \mu(\mathbb{E}\mathcal{N}(t)) \\
&\leq \limsup_{t} \prod_{i=1}^{\mathbb{E}\mathcal{N}(t)} \tau(G(i))\mu(0) \\
&\leq D \limsup_{t} \prod_{i=1:\ T_i \geq \mathcal{T}}^{\mathbb{E}\mathcal{N}(t)} (1 - \omega(T_i^{-\alpha})) \\
&\leq D \limsup_{t} e^{-\mathbb{E}[\mathcal{N}(t)]\omega(T_{N(t)}^{-\alpha})} \\
&\leq D \limsup_{t} e^{-cB^{-1}t^{1-\alpha}} \\
&= 0
\end{aligned} \tag{51}$$

where $0 \leq \alpha < 1$, $D = D(\mathcal{T}, \mu(0))$, B is the upper bound of the inter-renewal time and c is the constant of the ω notation.

It only remains to fulfil Def(ii). Consider the renewal times $\{T_k\}$

$$\begin{aligned}
\lim_{t\to\infty} \sum_{s=0}^{t} \mu(s) &\leq D(\mathcal{T}, \mu(0)) + \limsup_{t\to\infty} \sum_{s\geq\mathcal{T}}^{\mathbb{E}\mathcal{N}(t)} \mu(s) \\
&\leq D(\mathcal{T}, \mu(0)) + \limsup_{t\to\infty} \sum_{s\geq\mathcal{T}}^{\mathbb{E}\mathcal{N}(t)} \prod_{i=0}^{s} \tau(G_i)\mu(0) \\
&\leq D(\mathcal{T}, \mu(0)) + \limsup_{t\to\infty} B \sum_{s\geq\mathcal{T}}^{\mathbb{E}[\mathcal{N}(t)]} e^{-cB^{-1}s^{1-\alpha}} \mu(0) < \infty
\end{aligned} \tag{52}$$

The last step is due to the fact that the sum can be handled as in the static case to lead to a Gamma function.

Theorem 8. *Given the system* (DS) *with connectivity functions according to Assumption 1 and random perturbation according to Assumption 2, we have unconditional flocking with probability 1.*

4 The Continuous Problem

In this section we will discuss the continuous problem.

4.1 Some preliminaries

Let us begin with some notations and definitions for continuous time systems.

Definition 13. *The upper Dini derivative of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined by*

$$D^+ f(x)(v) = \limsup_{\alpha \downarrow 0,\ v' \to v} \frac{f(x + \alpha v') - f(x)}{\alpha} \tag{53}$$

Given the Euclidean space $\mathbb{E}^k$ the consensus subspace is defined as

$$\mathcal{C} = \{x \in \mathbb{E}^k : x_1 = \cdots = x_k\} \tag{54}$$

In the continuous case, the dynamics are:

$$\dot{x}(t) = -L_t x(t) \tag{55}$$

where $x, v \in \mathbb{E}^k$ the positions and velocities of the k agents respectively. Again we assume the live and move on the line, without loss of generality. Also $L_t = D(t) - A(t)$ is the Laplacian matrix:

$$L(t) := \begin{pmatrix} \sum_{j\neq 1} a_{1j} & -a_{12} & \cdots & -a_{1k} \\ -a_{21} & \sum_{j\neq 2} a_{2j} & \cdots & -a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{k1} & -a_{k2} & \cdots & \sum_{j\neq k} a_{kj} \end{pmatrix} \tag{56}$$

where $a_{ij} \geq 0,\ \sum_{j\neq l} a_{lj} > 0$ they satisfy the second part of Assumption 1.

4.2 The static problem. Lyapunov Stability

Since in the continuous time case there is no such tool as the coefficient of ergodicity it is at some point needed to be constructed. Note, however, that

$$m(t) = \max_i x_i - \min_i x_i$$

is a candidate Lyapunov function that is positive definite with respect to $\mathcal{C}$.

4.2.1 Convergence of $m(t)$

Consider the continuous problem (55) and the corresponding connectivity graph $G = (V, E)$ where $E = E(t)$ may change with time. The following theorem is a verbal result with mild assumptions on the connectivity regime. Note that $m(t)$ is the sum of two functions $V_1(t) = \max_i\{x_i(t)\}$ and $V_2(t) = -\min_i\{x_i(t)\}$ so that none of these two equations are smooth throughout the solution. So we use the Dini derivative to show convergence. Take:

$$D^+ V_1(t) = \max_i\{\dot{x}_i(t)\} = \dot{x}_m(t) = \sum_j a_{mj}(t)(x_j(t) - x_m(t)) \leq 0 \tag{57}$$

where we assumed that the m node has the highest acceleration of all other nodes at time t. If there are more than one pick the one that has the highest $\ddot{x}_i(t)$. If it's still more than one birds, pick among those at random and stick to this until at least a change in the roaster of $x's$ occurs. So we see that the Dini derivative decreases along a solution and is identically zero if $\dot{x}_m(t) \equiv 0$ if and only if $x_m(t) := \max_i x_i(t) = x_1(t) = \ldots x_k(t)$, assuming a complete graph ($a_{ij} > 0$). The same analysis holds for $V_2(t)$. So the invariance set $\{\dot{V}_1 = 0\}$ and the invariance set $\{\dot{V}_2 = 0\}$ are subsets of the consensus space and are both asymptotically attracting the solutions of the system, by La Salle's invariance principle and trivially by connectivity we conclude that, given a solution, the limit sets are identical in absolute value, so that $m(t) \to 0$.

$$D^+V(x(t)) \leq -\theta(t)V(x(t)) \tag{58}$$

for some particular positive function $\theta(t) \geq \omega(t^{-\alpha})$.

Complete Graph Topology The complete connectivity graph is the standard of the classic problem. Given $t > 0$ there is again a roaster in the speed of birds. Let m be the fastest and n the slowest (if there are more than one fastest (slowest) pick the one with the maximum acceleration as above). Then from the system dynamics we have

$$\begin{aligned} D^+x_m(t) &= \sum_j a_{mj}(x_j - x_m) \leq a_{mn}(x_n - x_m) = -a_{mn}V(x(t)) \\ D^+x_n(t) &= \sum_j a_{nj}(x_j - x_m) \geq a_{nm}(x_m - x_n) = a_{nm}V(x(t)) \\ &\Rightarrow \\ D^+V(x(t)) &= D^+x_m(t) - D^+x_n(t) \leq -(a_{mn} + a_{nm})V(x(t)) \end{aligned} \tag{59}$$

Star Graph Topology A similar result can be established if we pass from the complete connectivity to a star-like graph connectivity. Indeed in such case for any given $t > 0$ there exists a κ sending signal to every other else. Observe that if κ is in fact the fastest and/or the slowest then we can handle this situation as above. In every other case, take $a(t) := \min\{a_{m\kappa}, a_{n\kappa}\}$

$$\begin{aligned} D^+x_m(t) &= \sum_j a_{mj}(x_j - x_m) \leq a_{m\kappa}(x_\kappa - x_m) \leq -a(t)(x_m - x_\kappa) \\ D^+x_n(t) &= \sum_j a_{nj}(x_j - x_m) \geq a_{n\kappa}(x_\kappa - x_n) \geq a(t)(x_\kappa - x_n) \\ &\Rightarrow \\ D^+V(x(t)) &= D^+x_m(t) - D^+x_n(t) \leq -a(t)V(x(t)) \end{aligned} \tag{60}$$

where $a(t) \in \omega(t^{-\alpha})$ by assumption.

Remark 9. *Note that no other topology can establish equations like* (58) *without further assumptions. This establishes a bridge between the scrambling and Markov matrices in the discrete problem.*

So we have proved the following theorem

Theorem 9. *Consider the system* (55) *and an arbitrary solution of* $(x(t))$*. Under a complete or star graph topology, unconditional consensus is established for* $\alpha \in [0, 1]$.

4.3 The General Case

The general case is under the assumption of strong connectivity of the communication graph. The non homogeneous continuous Markov chain (NHCTMC) results will come in hand.

4.3.1 Non Homogeneous Continuous Markov Chains

The main observation is that $-L_t$ could be considered as an intensity matrix with some $a_{ij} = 0$ so that the graph representation $G = (V, E)$ constitutes an irreducible finite state markov chain. Be reminded that the scrambling index, γ introduced before is in fact $\frac{diam(G)-1}{2}$. In other words for any two nodes from where two random walks are generated, γ is the minimum amount of time that, with positive probability, the two walkers will find a common neighbour (or see each other). The problem of ergodicity in NHCTMC is discussed in papers like [16, 8] under the assumption of uniform boundedness in the following sense:

$$\int_{nT}^{(n+1)T} a_{i,j}(s)ds \geq c > 0 \tag{61}$$

for some $T > 0$ and all i, j, n.

Set $a(t) = \min_{i,j} a_{ij}(t)$ and consider three connected agents i and j and l

$$\begin{aligned}
\dot{x}_i &= \sum_j a_{ij}(x_j - x_i) \geq -\frac{1}{k-1}x_i \Rightarrow x_i(t_1) \geq x_i(t_0)e^{-\frac{t_1-t_0}{k-1}} \\
&\geq \min_i x_i(t_0)e^{-\frac{t_1-t_0}{k-1}} := \min_i x_i(t_0)b(t_1, t_0) = \min_i x_i(t_0)b \\
\dot{x}_j &\geq -\frac{1}{k-1}x_j + a_{ji}x_i \geq -\frac{1}{k-1}x_j + ba_{ij} \Rightarrow \\
x_j(t_1) &\geq x_i(t_0)b^2 \int_{t_0}^{t_1} a_{ij}(u)du \geq x_i(t_0)b^2 \int_{t_0}^{t_1} a(u)du \\
\dot{x}_l &\geq -\frac{1}{k-1}x_l + a_{lj}x_j \Rightarrow x_l(t_1) \geq x_i(t_0)b^3 \int_{t_0}^{t_1} a(s) \int_{t_0}^{s} a(k)dk
\end{aligned}$$

Note that $a(t) > 0$ since there is a connection between i, j, l and that it is a C^0 function of time. In the general case take any i, j nodes and find the shortest path between them. The length of this path is at most $2\gamma + 1$. Now, let m be the node that is a common neighbour of i, j after γ steps. From (55) the fundamental matrix solution is for some appropriate $T > 0$ and $t >> 1$

$$x(t) = \Phi(t, t_0)x(t_0) = \Phi(t, nT)\Phi(nT, (n-1)T)\Phi((n-1)T, (n-2)T)\cdots\Phi(T, t_0)x(t_0)$$

This manipulation above can be used for some $T > 0$ so that

$$[\Phi(nT, (n+1)T)]_{ij} \geq \int_{nT}^{(n+1)T} a_{ii_1}(s_1) \int_{nT}^{s_1} a_{i_1 i_2}(s_2)\cdots \int_{nT}^{s_\gamma} a_{i_{\gamma-1}j}(s_\gamma)ds_1\cdots ds_\gamma \geq C\omega((n+1)^{-\gamma\alpha}) \tag{62}$$

for $T > 0$ fixed, independent of time and $n >> 1$ and using the standard mean value theorem for integrals. Furthermore,

$$\begin{aligned}
||x_m((n+1)T) - x_j((n+1)T)|| &\leq ||(x_m(nT) - x_j(nT))\Phi((n+1)T, nT)|| \\
&\leq [1 - \omega((n+1)^{-\gamma\alpha})]||x_m(nT) - x_j(nT)|| \\
&\leq e^{-\sum_{k=1}^{n}\omega((k+1)^{-\alpha\gamma})}||x_m(0) - x_j(0)||
\end{aligned} \tag{63}$$

where $||\cdot|| = ||\cdot||_2$. Since there is no restriction on the agents i, j, m note that for different n they are free to change and the convergence is still guaranteed when $\alpha \in [0, \frac{1}{\gamma}]$, in view of the fact that Φ is always a bounded matrix.

4.4 Switching Topologies

All the results from the discrete case discussion directly apply. The results of the previous section indicate that the continuous time problem can be equivalently handled in a discrete variant.

4.5 Random Failure Setup

In this section we establish a random failure regime for the flocking dynamics system models. We define the quantities:

$$\begin{aligned} \Lambda(t) :&= \frac{1}{2}\sum_{i\neq j} ||v_i(t) - v_j(t)||^2 = v^T Q v \\ \Gamma(t) :&= \frac{1}{2}\sum_{i\neq j} ||x_i(t) - x_j(t)||^2 = x^T Q x \end{aligned} \tag{64}$$

where Q is the Laplacian of the complete $k \times k$ graph with unit weights. Then of course, Λ and Γ are stochastic processes.

4.5.1 Poisson Counters

Consider for every bird i a Poisson counter N_{ij} which controls the instantaneous connectivity with the bird j. This means that the moment N_{ij} fires, the bird i loses connection with bird j. We assume that this is a Poisson process with rate λ_{ij}. Consider the system

$$\begin{aligned} dx(t) &= v(t)dt \\ dv(t) &= -L_x v(t)dt \end{aligned} \tag{65}$$

where

$$L(x(t)) := \begin{pmatrix} \sum_{j\neq 1} \tilde{a}_{1j} & -\tilde{a}_{12} & \cdots & -\tilde{a}_{1k} \\ -\tilde{a}_{21} & \sum_{j\neq 2} \tilde{a}_{2j} & \cdots & -\tilde{a}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ -\tilde{a}_{k1} & -\tilde{a}_{k2} & \cdots & \sum_{j\neq k} \tilde{a}_{kj} \end{pmatrix}_{x(t)} \tag{66}$$

and $\tilde{a}_{ij} := z_{ij} a_{ij}$ with z_{ij} a Poisson driven variable that takes values in $\{0, 1\}$ and governed by

$$dz_{ij} = (1 - 2z_{ij})dN_{ij} \tag{67}$$

So the speed of the i^{th} bird at time t is governed by the system of equations

$$\begin{aligned} dx_i &= v_i dt \\ dv_i(t) &= \sum_j z_{ij} a_{ij}(t)(v_j(t) - v_i(t))dt \\ dz_{ij} &= (1 - 2z_{ij})dN_{ij} \end{aligned} \tag{68}$$

By standard arguments we have the solutions of $\mathbb{E}[z_{ij}(t)]$

$$\mathbb{E}[z_{ij}(t)] = e^{-2\lambda_{ij}t}[z_{ij}(0) - 1/2] + \frac{1}{2} \tag{69}$$

We understand that z_i are random variables taking values in $\{0,1\}$. Consequently, x_i and v_i are r.v.'s too such that $z_i^{\prime}s$ are independent among themselves and also z_i are independent of x_i and v_i. However, x_i and v_i are not independent. Of course, a_{ij} can be considered as a random variable that is a function of x_i, x_j. The following assumption will be used

Assumption 6. $\forall t: \quad a_{ij}(t) \geq h(t)$ *a.s. where* $h(t)$ *is a nonnegative function of time with* $h(t) \in \omega(t^{-\alpha})$ *for* $0 \leq \alpha < 1$

Flocking of two birds Let's consider a warm up example in view of (68) when $k = 2$. In such case each bird is equipped with a single Poisson counter that characterizes the on/off periods of connectivity with the other bird. Lets z_1 and z_2 be the corresponding counters. Then if we consider the column vector $y = (x_1, x_2, v_1, v_2, z_1, z_2)$ we can write it in the following form

$$dy = f(y)dt + g_1(y)dN_1 + g_2(y)dN_2 \tag{70}$$

where

$$f(y) = \begin{pmatrix} v_1 \\ v_2 \\ z_1 a_{12}(v_2 - v_1) \\ z_2 a_{21}(v_1 - v_2) \\ 0 \\ 0 \end{pmatrix} \quad , \quad g_1(y) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 - 2z_1 \\ 0 \end{pmatrix} \quad , \quad g_2(y) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 - 2z_2 \end{pmatrix}$$

The quantities of interest are Λ, Γ. The Ito rule [1] yields

$$d\Lambda = -2(z_1 a_{12} + z_2 a_{21})\Lambda dt$$

taking the expectation,

$$\frac{d}{dt}\mathbb{E}[\Lambda] = -2\mathbb{E}\big[(z_1 a_{12} + z_2 a_{21})\psi\big] \leq -2[(E[z_1(t)] + \mathbb{E}[z_2(t)])f(t)\big]\mathbb{E}[\Lambda] \tag{71}$$

Furthermore, for Γ

$$d\Gamma = 2(x_1 - x_2)(v_1 - v_2)dt$$

and taking expectations we obtain

$$\frac{d}{dt}\mathbb{E}[\Gamma] = 2\mathbb{E}[(x_1 - x_2)(v_1 - v_2)] \leq 2\mathbb{E}[\Delta] \tag{72}$$

where $\Delta(t) := |(x_1 - x_2)(v_1 - v_2)|$, and we need to compute the expected value of Δ. In order to do this in a useful way we must take cases for the relative signs of x_1, x_2, v_1, v_2. As it turns out

$$d\Delta = \pm(v_1 - v_2)^2 - (z_1 a_{12} + z_2 a_{21})|(x_1 - x_2)(v_1 - v_2)| \leq \Gamma - (z_1 a_{12} + z_2 a_{21})\Delta dt$$

It follows that

$$\frac{d}{dt}\mathbb{E}[\Delta] = \mathbb{E}[\Lambda] - \mathbb{E}[(z_1 a_{12} + z_2 a_{21})F] \leq \mathbb{E}[\Lambda] - [(\mathbb{E}[z_1(t)] + \mathbb{E}[z_2(t)])h(t)]\mathbb{E}[\Delta] \tag{73}$$

Note that the expression converges as $t \to \infty$

$$-2\int_0^t (K_1 e^{-2\lambda_1 s} + K_2 e^{-2\lambda_2 s} + 1)h(s)ds \le C - \omega(t^{1-\alpha}) - \omega(t^{2-\alpha})$$

where we used the formula $e^{-s} \ge 1 - s \; \forall s \in \mathbb{R}$.

Then (abusing the constant symbols) for t large enough:

$$\mathbb{E}[\Lambda(t)] \le Ce^{-\omega(t^{1-\alpha})-\omega(t^{2-\alpha})} \to 0$$

$$\mathbb{E}[\Delta(t)] \le C[e^{-\omega(t^{1-\alpha})-\omega(t^{2-\alpha})} + te^{-\omega(t^{1-\alpha})-\omega(t^{2-\alpha})}]$$

From (57) it follows that $\mathbb{E}(\Gamma)$ is uniformly bounded concluding the example □

4.6 The general case.

In the general case we assume that each bird has $k-1$ independent Poisson counters. This means that there exist $k(k-1)$ independent Poisson counters in the whole flock. So here y is a $k(k+1)\times 1$ vector and (55) looks the same with the general case only with

$$f(y) = \begin{pmatrix} v \\ -\tilde{L}v \\ \mathbf{0} \end{pmatrix} \tag{74}$$

where $v^T = \{v_1, \dots, v_k\}$, $\tilde{L}$ is the Laplacian of the complete graph with weights $z_{ij}a_{ij}$ and $\mathbf{0}$ is a $k(k-1)\times 1$ zero vector. Since there are $k(k-1)$ Poisson counters there are equally many g functions such that

$$dy = f(y)dt + \sum_{i=1}^{k}\sum_{j=1, j\neq i}^{k} g_{ij}(y)dN_{ij} \tag{75}$$

where $g_{ij}(y)$ is a zero column vector apart from the i^{th} entry where one has $(1-2z_{ij})$ (i.e. the instant connection status of the i^{th} to the j^{th}). The first quantity we have to calculate is $\Lambda := v^T Q v$. Since,

$$\frac{\partial \Lambda}{\partial y}^T = \left(\mathbf{0}, \frac{\partial \Lambda}{\partial v}, \mathbf{0}\right) = \left(\mathbf{0}, 2v^T Q, \mathbf{0}\right) \tag{76}$$

from the Ito rule again we have:

$$\frac{d}{dt}\mathbb{E}[\Lambda(t)] = \mathbb{E}[\frac{\partial \Lambda}{\partial v}\frac{dv}{dt}] = -2\mathbb{E}[v^T Q L_x v] \tag{77}$$

In the symmetric case (i.e. $\tilde{a}_{ij} = \tilde{a}_{ij}$) we have $QL_x = kL_x$. Using the identity [REF]

$$\langle v, L_x v\rangle = \frac{1}{2}\sum_{i,j}\tilde{a}_{ij}(v_i - v_j)^2 \tag{78}$$

(77) gives

$$\frac{d}{dt}\mathbb{E}[\Lambda(t)] = -2k\mathbb{E}[v^T L_x v] = -k\sum_{i,j}\mathbb{E}[\tilde{a}_{ij}(v_i - v_j)^2] \tag{79}$$

Note that, by independence

$$\mathbb{E}[\tilde{a}_{ij}(t)(v_i - v_j)^2] = \mathbb{E}[z_{ij}]\mathbb{E}[a_{ij}(v_i - v_j)^2] = [Ce^{-2\lambda_{ij}t} + \frac{1}{2}]\mathbb{E}[a_{ij}(v_i - v_j)^2] := \gamma_{ij}(t)\mathbb{E}[a_{ij}(v_i - v_j)^2] \tag{80}$$

For $a_{ij}(t) \geq \omega(t^{-\alpha})$ as $t \to \infty$: Although both $a_{ij}(t)$ and $v_i(t), x_i(t)$ are random variables considering that $a_{ij}(t)$ is eventually bounded from below by $\omega(t^{-\alpha})$ we have

$$\frac{d}{dt}\mathbb{E}[\Lambda(t)] \leq -\hat{C}\omega(t^{-\alpha})\mathbb{E}[\Lambda(t)] \tag{81}$$

for $t \geq \mathcal{T}$, where $\mathcal{T}$ does not depend on time or on any initial condition and for $\hat{C} > 0$ an appropriate constant.

Remark 10. *We silently assumed that $a_{ij} = a_{ji}$. This assumption is crucial for a population greater or equal to 3.*

Remark 11. *This result tells us that no matter what the failure rate will be, as long as it is a given constant, unconditional flocking will occur in the second mean for $0 \leq \alpha < 1$. This should come as no surprise though. It would be interesting to investigate what happens when the rate is time varying. For this we must modify the formulas.*

References

[1] R Brockett, *Stochastic control*, Control **16** (2008), no. 5, 702–714.

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

[3] Chris Godsil and Gordon Royle, *Algebraic graph theory*, vol. 207, Springer, 2001.

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

[5] Y Hatano and M Mesbahi, *Agreement over random networks*, IEEE Transactions on Automatic Control **50** (2005), no. 11, 1867–1872.

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

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

[8] Dean L. Isaacson Mark Scott, Barry C. Arnold, *Strong ergodicity for continous time non-homogeneous markov chains*, J. Appl. Prob. **19** (1982), 692–694.

[9] Ion Matei, Nuno Martins, and John S Baras, *Almost sure convergence to consensus in markovian random graphs*, 2008 47th IEEE Conference on Decision and Control (2008), 3535–3540.

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

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

[12] Maurizio Porfiri and Daniel L Stilwell, *Consensus seeking over random weighted directed graphs*, Automatic Control IEEE Transactions on **52** (2007), no. 9, 1767–1773.

[13] E Seneta, *Non-negative matrices and markov chains*, vol. 37, Springer, 2006.

[14] Alireza Tahbaz-Salehi and Ali Jadbabaie, *On consensus over random networks*, vol. 53, pp. 1315–1321, 2006.

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

[16] A. I. Zeifman, *Quasi-ergodicity for non-homogeneous continuous-time markov chains*, J. Appl. Prob. **26** (1989), 643–648.