# Consensus dynamics over networks

Fabio Fagnani
Dipartimento di Scienze Matematiche 'G.L Lagrange'
Politecnico di Torino
fabio.fagnani@polito.it

January 29, 2014

**Abstract**

Consensus dynamics is one of the most popular multi-agent dynamical systems. It shows up in socio-economic contexts as a model for consensus formation in a society of individuals. In the engineering world, it has been considered as a basic algorithm to be employed in networks of agents (sensors, vehicles, etc.) for efficient distributive computation of global functions. In this survey we review the basic theory of consensus dynamics presenting a number of classical examples. Particular emphasis is devoted to random consensus dynamics and to models considering the presence of stubborn agents in the network.

## 1 Introduction

Multi-agent systems constitute one of the fundamental paradigms of the science and technology of present century [12, 37, 41, 42]. The main idea is that of creating complex dynamical evolutions from the interactions of many simple units. Indeed such collective behaviors are quite evident in biological and social systems and were indeed considered in earlier times [45]. More recently, the digital revolution and the miniaturization in electronics have made possible the creation of man-made complex architectures of interconnected simple devices (computers, sensors, cameras). Moreover, the creation of internet has opened totally new form of social end economic aggregation. This has strongly pushed towards a systematic and deep study of multi-agent dynamical systems. Mathematically they typically consist of a graph where each node possesses a state variable; states are coupled at the dynamical level through dependences determined by the edges in the graph. One of the challenging problem in the field of multi-agent systems, is to analyze the merging of complex collective phenomena from the interactions of the units which are typically quite simple. Complexity is typically the outcome of the topology and the nature of interconnections which are often of stochastic type. From the applicative point of view such dynamical systems can represent algorithms on some infrastructure network (e.g sensor network) or rather be a model for biological or socio-economic behaviors.

Consensus dynamics (also known as average dynamics) [44, 4, 7, 35, 39] is one of the most popular and simplest multi-agent dynamics. One convenient way to introduce it is with the language of social sciences. Imagine that a number of independent units possess an information represented by a real number, for instance such number can represent an opinion on a given fact. Units interact and change their opinion by averaging with the opinions of other units. Under certain assumptions this will lead the all community to converge to a consensus opinion which takes into consideration all the initial opinion of the agents. In social sciences, empiric evidences [26, 46] have shown how such aggregate opinion may give avery good estimation of unknown quantities: such phenomenon has been proposed in the literature as wisdom of crowds [43].

Below, we present a brief outline of the content of this survey paper. In Chapter 2, we formally present consensus dynamical systems based on the evolution of a stochastic matrix . We recall basic notions of the Perron-Frobenius theory and discuss conditions for convergence and for the estimation of its speed. Chapter 3 is devoted to the presentation of some basic examples and we investigate the scaling of the convergence time as a function of the number of node which is a fundamental index in the applications. In Chapter 4 we deepen our analysis focusing on the important family of time-reversible stochastic matrices. We present the classical Cheeger bound which relates the convergence time to the geometry of the underlying graph. Chapter 5 considers asynchronous random consensus dynamics presenting some key examples (e.g. gossip models) and the basic elements of the mean square theory to study consensus in a probabilistic sense. Chapter 6 contains some more advanced topics. Here we deal with consensus dynamics in the presence of stubborn agents never changing their state. Analysis of such systems naturally leads to study electrical networks and they classical relation with time-reversible matrices. We conclude presenting a result which studies the relation between polarization of the opinions in the social network and the geometry of the underlying graph.

Literature on consensus has witnessed an exponential growth in the last decade. This survey paper does not pretend to cover all the fundamental aspects of consensus. Our choice has privileged topics for which a reasonably complete mathematical theory is available. Important topics not included in this survey are time-varying dynamics [35], second order models [40, 11, 14] and bounded confidence models.[32, 16].

## 1.1 Notation

All notation conventions will be recalled when introduced in the paper. For the reader's convenience, we here gather some of the most used notation. We will typically consider vector spaces of type $\mathbb{R}^{\mathcal{V}}$ where $\mathcal{V}$ is a finite set so that vector components are indicized by $v \in \mathcal{V}$. The symbol $\mathbb{1} \in \mathbb{R}^{\mathcal{V}}$ always denote the vector whose components are all equal to 1. $e_v \in \mathbb{R}^{\mathcal{V}}$ is the vector with all components equal to 0 except the $v$-th component which is equal to 1. Given a real matrix $A$, the adjoint matrix will be denoted by $A^*$. If $\mathcal{V}$ is a finite set, $|\mathcal{V}|$ denotes its cardinality.

## 2 Consensus dynamics, graphs and stochastic matrices

Mathematically, consensus dynamics are special linear dynamical systems of type

$$x(t+1) = Px(t) \tag{1}$$

where $x(t) \in \mathbb{R}^{\mathcal{V}}$ and $P \in \mathbb{R}^{\mathcal{V}\times\mathcal{V}}$ is a *stochastic matrix* (e.g. a matrix with non-negative elements such that every row sums to 1). $\mathcal{V}$ represents the finite set of units (agents) in the network and $x(t)_v$ is to be interpreted has the state (opinion) of agent $v$ at time $t$. Typically, we put $N := |\mathcal{V}|$. Update equation (1) implies that state of agents at time $t+1$ are convex combinations of the components of $x(t)$: this motivates the term averaging dynamics.

The network structure is hidden in the non-zero pattern of $P$. Indeed we can associate to $P$ a graph: $\mathcal{G}_P = (\mathcal{V}, \mathcal{E}_P)$ where $\mathcal{V}$ is the set of nodes and where the set of edges is given by $\mathcal{E}_P := \{(u,v) \in \mathcal{V}\times\mathcal{V} \mid P_{uv} > 0\}$. Elements in $\mathcal{E}_P$ represent the communication edges among the units; specifically, in our setting, the existence of the edge $(u,v)$ has to be interpreted in the sense that unit $u$ has access to the state of unit $v$ (or, if we prefer, that $v$ can transmit information to $u$). Denote by $\mathbb{1} \in \mathbb{R}^{\mathcal{V}}$ the all 1's vector. Notice that $P\mathbb{1} = \mathbb{1}$: this shows that once the states of units are at consensus, they will no longer change. The crucial point to understand is if dynamics will always converge to a consensus point.

Remarkably, some of the key properties of $P$ responsible for the transient and asymptotic behavior of the linear system (1) are determined by the connectivity properties of the underlying graph $\mathcal{G}_P$. We recall that, given two vertices $u, v \in \mathcal{V}$, a *path* (of length $l$) from $u$ to $v$ in $\mathcal{G}_P$ is any sequence of vertices $u = u_1, u_2, \ldots, u_{l+1} = v$ such that $(u_i, u_{i+1}) \in \mathcal{E}_P$ for every $i = 1, \ldots, s$. $\mathcal{G}_P$ is said to be *strongly connected* if for any pair of vertices $u \neq v$ in $\mathcal{V}$ there is a path in $\mathcal{G}_P$ connecting $u$ to $v$. The *period* of a node $u$ is defined as the greatest common divisor of the length of all closed paths from $u$ to $u$. In strongly connected graph all nodes have the same period, and the graph is called *aperiodic* if such a period is 1. A stochastic matrix $P$ is said to be *irreducible* if $\mathcal{G}_P$ is strongly connected, and *primitive* if $\mathcal{G}_P$ is strongly connected and aperiodic. The following classical result holds true [27]:

**Theorem 1. (Perron-Frobenius)** *Assume that $P \in \mathbb{R}^{\mathcal{V}\times\mathcal{V}}$ is a primitive stochastic matrix. Then,*

1. $1$ *is an algebraically simple eigenvalue of $P$.*
2. *There exists a (unique) probability vector $\pi \in \mathbb{R}^{\mathcal{V}}$ ($\pi_v > 0$ for all $v$ and $\sum_v \pi_v = 1$) which is a left eigenvector for $P$, namely $\pi^* P = \pi^*$.*
3. *All the remaining eigenvalues of $P$ are of modulus strictly less than* $1$.

A straightforward consequence of this result is that $P^t \to \mathbb{1}\pi^*$ for $t \to +\infty$. This yields

$$\lim_{t\to+\infty} x(t) = \lim_{t\to+\infty} P^t x(0) = \mathbb{1}(\pi^* x(0)) \tag{2}$$

In other terms dynamics leads asymptotically to a consensus: all agents' state converging to the common value $\pi^* x(0)$, called consensus point which is a convex combination of the initial states with weights given by the invariant probability components.

If $\pi$ is the uniform vector (i.e. $\pi_i = |\mathcal{V}|^{-1}$ for all $i$), the common asymptotic value is simply the arithmetic mean of the initial states In this case all agents equally contribute to the final common state (we often refer to this case as average consensus). This uniformity condition amounts to assume that $\mathbb{1}^* P = \mathbb{1}^*$, namely that also columns of $P$ sum to 1. Such matrices are called *doubly stochastic*, a sufficient condition for this being that $P$ is symmetric. In many applications this uniformity condition is necessary and is enforced in the model for instance assuming that $P$ is symmetric. Indeed, the distributed computation of the arithmetic mean is an important step to solve estimation problems for sensor networks [47]. As a specific example consider the situation where there are $N$ sensor deployed in a certain area and each of them makes a noisy measurement of a physical quantity $x$. Let $y_v = x + \omega_v$ be the measure obtained by sensor $v$, where $\omega_v$ is a zero mean Gaussian noise. It is well known that if noises are independent and identically distributed, the optimal mean square estimator of the quantity $x$ given the entire set of measurements $\{y_v\}$ is exactly given by $\hat{x} = N^{-1} \sum_v y_v$. The use of consensus in more sophisticated estimation problems where the quantity to be estimate is time-varying or where sensors may have different performances, have been considered in the literature [8, 13, 21]. Other fields of application is in computer load balancing [36], and in the control of cooperative autonomous vehicles [23, 29]. Finally, consensus dynamics in socio-economic settings have appeared in [28, 30, 1, 2].

## 2.1 The rate of convergence

Basic linear algebra allows to study the rate of convergence to consensus: it will be clearly dictated by the largest in modulo among the eigenvalues of $P$ except 1; precisely,

**Proposition 2.** *Let $P \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$ be a primitive stochastic matrix. Consider all its eigenvalues $\mu_i$ but 1 and put $\rho_2 = \max\{|\mu_i| < 1\}$. Then, for every $\epsilon > 0$ there exists a constant $C_\epsilon$ such that*

$$||(P^t - \mathbb{1}\pi^*)x_0||_2 \leq C_\epsilon (\rho_2 + \epsilon)^t ||x_0||_2 \quad \textit{for all } t\,. \tag{3}$$

The parameter $\rho_2$, introduced in the statement of the proposition above, is also called the *second eigenvalue* of $P$, and the difference $1 - \rho_2$ the *spectral gap* of $P$. The above result essentially says that convergence to consensus happens exponentially fast as $\rho_2^t$; this is not exactly true because of the $\epsilon > 0$ we have to fix, but since $\epsilon$ can be chosen arbitrarily small, morally, this is true. The $\epsilon$ is needed because of the possibility that the algebraic multiplicity relative to the second eigenvalue is strictly larger than the geometric multiplicity.

Stochastic matrices owe their name to their use in probability. Indeed, given a stochastic matrix $P \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$, the term $P_{vw}$ can be interpreted as the probability of making a transition from state $v$ to state $w$: you can imagine to be sitting at state $v$ and to walk along one of the available outgoing edges from $v$ according to the various probabilities $P_{vw}$. In this way you construct what is called a random walk on the graph $\mathcal{G}$. In this probabilistic setting, it is interesting to consider the dual dynamics induced by $P$. Indeed, if $\zeta \in \mathbb{R}^V$ is a probability vector where $\zeta_v$ indicates the probability that at the initial instant the state is equal to $v$, then $(\zeta^* P)_v$ indicates the probability of finding the process in state $v$ at the next time. Therefore, $\zeta^* P^t$ represent the marginal probability distribution of the random walk at time $t$. Theorem 1 establishes that if $P$ is primitive such marginals converge to the probability $\pi$ disregarding of the initial probability vector $\zeta$.

An important class of stochastic matrices, particularly used in applications, are the *time-reversible* stochastic matrices. They are defined through a symmetric non-negative valued matrix $C \in \mathbb{R}^{\mathcal{V}\times\mathcal{V}}$ (called *conductance matrix*) such that $C_v = \sum_{v'} C_{vv'} > 0$ for all $v \in \mathcal{V}$, by putting

$$P_{vv'} = \frac{C_{vv'}}{C_v} \tag{4}$$

It is immediate to check that $\pi_v = C_v / \sum_{v'} C_{v'}$ is an invariant probability for $P$. If $P$ is primitive this is the unique invariant probability.

In order to present a special important example, we first review few concepts from graph theory. Consider a strongly connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The *adjacency matrix* of $\mathcal{G}$ is a square matrix $A_{\mathcal{G}} \in \{0,1\}^{\mathcal{V}\times\mathcal{V}}$ such that $(u,v) \in \mathcal{E}$ iff $(A_{\mathcal{G}})_{uv} = 1$. $\mathcal{G}$ is said to be symmetric if $A_{\mathcal{G}}$ is symmetric (e.g. $(u,v) \in \mathcal{E}$ iff $(v,u) \in \mathcal{E}$). The time-reversible stochastic matrix $P$ induced by the conductance matrix $A_{\mathcal{G}}$ is called the *simple random walk (SRW)* on $\mathcal{G}$. Notice that if $(u,v) \in \mathcal{E}$, $P_{uv} = d_u^{-1}$ where $d_u = \sum_v (A_{\mathcal{G}})_{uv}$ is the *degree* of node $u$. Notice that $\pi_v = d_v/|\mathcal{E}|$. The corresponding random walk thus put uniform probability on all the edges outgoing any given node in the graph. Notice that since $\mathcal{G}_P = \mathcal{G}$, $P$ is irreducible. If $P$ is also primitive, than, we have convergence to the consensus point

$$\pi^* x(0) = \frac{\sum_v d_v x(0)_v}{|\mathcal{E}|}$$

Each node contributes with its initial state to this consensus with a weight which is proportional to the degree of the node. Primitivity is not an important issue in general since the SRW can always be modified adding some weight on the diagonal terms by considering the so called *lazy SRW*, formally defined as

$$P_\lambda = (1-\lambda)P + \lambda I$$

where $\lambda \in (0,1)$. $P_\lambda$ has the same invariant probability than $P$ and is always primitive. A quite popular choice in the literature is to pick $\lambda = 1/2$. This has the advantage that, in case the eigenvalues of $P$ are real in $[-1,1]$ (this happens for instance in the symmetric case), then $P_{1/2}$ has all positive eigenvalues.

Notice how the SRW $P$ is in general not symmetric, and neither doubly stochastic, in spite of the symmetricity of the underlying graph. Indeed, it is immediate to verify that $P$ is symmetric if and only if $\mathcal{G}$ is $d$-regular which is when all nodes have the same degree $d$. In applications where symmetricity is a requirement, one can consider a variation of the SRW known as *Metropolis random walk*. It is formally defined by putting, for every $(u,v)\in\mathcal{E}$ with $u\neq v$

$$P_{uv}=\min\left\{\frac{1}{d_u},\frac{1}{d_v}\right\}$$

and adjusting accordingly the weights on the diagonal part of the matrix to enforce stochasticity of the matrix.

The SRW as defined above, actually works fine on any graph, not necessarily symmetric: the degree $d_u$ is more precisely the out-degree of node $u$, namely the number of outgoing edges from $u$. A remarkable difference with respect to the symmetric case is that, in general, there is no simple expression for the invariant probability $\pi$.

A key fact about time-reversible matrices is the fact that they are diagonalizable as discussed below. Let $P\in\mathbb{R}^{\mathcal{V}\times\mathcal{V}}$ be a primitive time-reversible stochastic matrix and let $\pi\in\mathbb{R}^{\mathcal{V}}$ be its invariant probability measure. Consider $D_\pi$ the diagonal matrix such that $(D_\pi)_{vv}=\pi_v$ and define $A=D_\pi^{1/2}PD_\pi^{-1/2}$. Time-reversibility implies that $A$ is symmetric. Let $\phi_j$'s, for $j=1,\dots,n$, be an orthonormal basis of eigenvectors for $A$ with correspondent real eigenvalues $\mu_j$. Order them in such a way that $1=\mu_1>\mu_2\geq\cdots\geq\mu_N$. It is immediate to check that $\pi^{1/2}$ is indeed an eigenvector with eigenvalue 1. Therefore, we will assume that $\phi_1=\pi^{1/2}$. Using the usual orthonormal splitting expression for a symmetric matrix we can write

$$A^t=\pi^{1/2}(\pi^{1/2})^*+\sum_{j\geq 2}\mu_j^t\phi_j\phi_j^*,$$

from which we can derive the following useful representation for $P^t$

$$P^t=\mathbb{1}\pi^*+D_\pi^{-1/2}\sum_{j\geq 2}\mu_j^t\phi_j\phi_j^*D_\pi^{1/2}.$$

From this expression, straightforward steps allow to obtain the following estimation

$$||P^tx(0)-\mathbb{1}\pi^*x(0)||_2\leq\frac{\max_v\pi_v^{1/2}}{\min_v\pi_v^{1/2}}||x(0)||_2\rho_2^t\quad\forall t\in\mathbb{N},\tag{5}$$

where in this case $\rho_2=\max\{\mu_2,-\mu_N\}$. In the special case when $P$ is symmetric, the first term in the right hand side of (5) is equal to 1 and the estimation takes a particularly simple form, to be compared with the general estimation (3).

# 3 Examples and large scale analysis

In this chapter we present a number of classical examples based on families of graphs with larger and larger number of nodes $N$. In this setting, particularly relevant is to understand the behavior of the second eigenvalue $\rho_2$ as a function of $N$. Typically, one consider $\epsilon > 0$ fixed and solves the equation $\rho_2^t = \epsilon$. The solution $\tau = (\ln \rho_2^{-1})^{-1} \ln \epsilon^{-1}$ will be called the *convergence time*: it essentially represents the time needed to shrink of a factor $\epsilon$ the distance to consensus. Dependence of $\rho_2$ on $N$ will also yield that $\tau$ will be a function of $N$. In the sequel we will investigate such dependence for SRW's on certain classical families of graphs.

**Example 1** (SRW on a Complete graph)**.** *Consider the complete graph on the set $\mathcal{V}$: $K_{\mathcal{V}} := (\mathcal{V}, \mathcal{V} \times \mathcal{V})$ (also self loops are present). The SRW on $K_{\mathcal{V}}$ is simply given by $P = N^{-1}\mathbb{1}\mathbb{1}^*$ where $N = |\mathcal{V}|$. Clearly, $\pi = N^{-1}\mathbb{1}$. Eigenvalues of $P$ are $1$ with multiplicity $1$ and $0$ with multiplicity $N-1$. Therefore, $\rho_2 = 0$. Consensus in this case is achieved in just one step: $x(t) = N^{-1}\mathbb{1}\mathbb{1}^* x(0)$ for all $t \geq 1$.*

**Example 2** (SRW on circulant graphs)**.** *Circulant graphs are graphs having the node set equal to the cyclic group $\mathbb{Z}_N = \{0, 1, \dots, N-1\}$ and the set of edges invariant by translation. More precisely, the graph $\mathcal{G} = (\mathbb{Z}_N, \mathcal{E})$ is circulant iff*

$$(u, v) \in \mathcal{E}, \Rightarrow (u + s, v + s) \in \mathcal{E}, \ \forall s$$

*(where sum in the formula above, has to be interpreted in mod $N$ sense). The corresponding adjacency matrix $A_{\mathcal{G}}$ is also called circulant: it possesses the property that its rows can be obtained from the first one, simply, by cyclic permutations: $(A_{\mathcal{G}})_{uv} = (A_{\mathcal{G}})_{0(v-u)}$. Let $d = \sum_u (A_{\mathcal{G}})_{0u}$. Clearly, each node $u$ has $d$ outgoing and $d$ incoming arcs. As a consequence $P = d^{-1}A_{\mathcal{G}}$ is a doubly stochastic matrix representing the SRW on $\mathcal{G}$ (even if $\mathcal{G}$ may as well be a non symmetric graph). Spectral analysis of circulant matrices turns out to be very simple using Fourier analysis. Indeed, if we consider $\theta$ the first row of $P$, eigenvalues are given by*

$$\lambda_k = \sum_{\ell=0}^{N-1} \theta_\ell e^{i\frac{2\pi k l}{N}}, \quad k = 0, 1, \dots, N-1$$

*with corresponding eigenvectors*

$$\phi_k = \left[1, e^{i\frac{2\pi k}{N}}, \dots, e^{i\frac{2\pi k(N-1)}{n}}\right]^* .$$

*As a special case consider the symmetric cycle graph $C_N$ which is the circulant graph with $N$ nodes whose adjacency matrix $A_{C_N}$ has as first row $\theta = (0, 1, 0 \dots, 0, 1)$ (node $0$ is connected to $1$ and to $N-1$). The corresponding SRW is given by $2^{-1}A_{C_N}$. Eigenvalues can be computed through the general formula above*

$$\lambda_k = \sum_{\ell=0}^{n-1} \theta_\ell e^{i\frac{2\pi k l}{n}} = \frac{e^{i\frac{2\pi k}{N}} + e^{i\frac{-2\pi k}{N}}}{2} = \cos\frac{2\pi k}{N}$$

*If $N$ is even, we have that $-1$ is an eigenvalue: indeed in this case, $P$ is not primitive (all nodes have period equal to 2). To avoid this type of problems, it is always convenient to consider the lazy SRW version $P_{1/2} = 1/2I + 1/2P$ having eigenvalues*

$$\frac{1}{2}(1+\lambda_k) = \frac{1}{2} + \frac{1}{2}\cos\frac{2\pi k}{N}\,.$$

*Therefore,*

$$\rho_2 = \frac{1}{2} + \frac{1}{2}\cos\frac{2\pi}{N} = 1 - \pi^2\frac{1}{N^2} + o(N^{-2}) \quad \text{for } N \to +\infty \tag{6}$$

*Therefore the corresponding convergence time is given by*

$$\tau = (\ln \rho_2^{-1})^{-1} \ln \epsilon^{-1} \asymp N^2 \quad \text{for } N \to +\infty$$

**Example 3** (SRW on toroidal grids). *Given two loopless graphs $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$ one can consider the product graph $\mathcal{G}_1 \times \mathcal{G}_2 = (\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{E})$ characterized by the adjacency matrix $A_{\mathcal{G}_1 \times \mathcal{G}_2}$ which is defined as*

$$(A_{\mathcal{G}_1\times\mathcal{G}_2})_{(u_1,u_2)(v_1,v_2)} := (A_{\mathcal{G}_1})_{u_1,v_1}\mathbb{1}_{\{u_2=v_2\}} + (A_{\mathcal{G}_2})_{u_2,v_2}\mathbb{1}_{\{u_1=v_1\}}$$

*A straightforward but useful property is that if $A_{\mathcal{G}_1}x_1 = \lambda_1 x_1$ and $A_{\mathcal{G}_2}x_2 = \lambda_2 x_2$, then, $A_{\mathcal{G}_1\times\mathcal{G}_2}x_1x_2^* = (\lambda_1+\lambda_2)x_1x_2^*$. This implies that the eigenvalues of $A_{\mathcal{G}_1\times\mathcal{G}_2}$ can be obtained as sum of one eigenvalue of $A_{\mathcal{G}_1}$ and one of $A_{\mathcal{G}_2}$. It is immediate to check that if $\mathcal{G}_1$ and $\mathcal{G}_2$ are symmetric, also $\mathcal{G}_1 \times \mathcal{G}_2$ is symmetric. Moreover if $\mathcal{G}_1$ and $\mathcal{G}_2$ are, respectively, $d_1$-regular and $d_2$-regular, then $\mathcal{G}_1 \times \mathcal{G}_2$ is $d_1+d_2$-regular. Considering the cycle graph $C_n$, we have that $C_n^2 = C_n \times C_n$ is the so-called toroidal 2-grid, and, more generally, $C_n^d$, the product of $d$ copies of $C_n$, represents the toroidal $d$-grid. $C_n^d$ is a symmetric $2d$-regular graph and the SRW on $C_n^d$ is given by $P = (2d)^{-1}A_{\mathcal{G}_1\times\mathcal{G}_2\times\cdots\times\mathcal{G}_d}$. Eigenvalues of $P_{1/2} = 1/2(I+P)$ are thus*

$$\frac{1}{2} + \frac{1}{2d}\sum_{j=1}^{d}\cos\frac{2\pi k_j}{n}\,, \quad (k_1,\ldots,k_d) \in \mathbb{Z}_n^d\,.$$

*The second eigenvalue is thus given by*

$$\rho_2 = \frac{1}{2} + \frac{1}{2d}\left(d - 1 + \cos\frac{2\pi}{n}\right) = 1 - \frac{\pi^2}{dn^2} + o(n^{-2})$$

*Apparently nothing has changed with respect to the cycle (see formula (6). However notice that, in the toroidal $d$-grid the number of nodes is $N = n^d$. This yields*

$$\tau \asymp N^{2/d} \quad \text{for } N \to +\infty$$

*This shows that the convergence time exhibits a slower growth in $N$ as the dimension $d$ of the grid increases: this is intuitive since the increase in $d$ determines a better connectivity of the graph and a consequently faster diffusion of information.*

We complete this chapter with a general comment. In all those situations where we have a sequence of stochastic matrices as in the examples above with second eigenvalues satisfying $1 - \rho_2 \asymp \alpha(N)$ for $N \to +\infty$ where $\alpha(N)$ is infinitesimal in $N$, then,

$$\tau = (\ln \rho_2^{-1})^{-1} \ln \epsilon^{-1} \asymp \alpha(N)^{-1} \asymp (1-\rho_2)^{-1}\,, \quad N \to +\infty$$

Namely, the inverse of the spectral gap dictates the scaling in $N$ of the convergence time.

# 4 The spectral gap of time-reversible matrices

In the previous chapter we have presented a number of classical examples for which it is possible to explicitly compute the spectral gap and thus the convergence time. In general, however, there is not a closed formula for such index. Indeed, computation of eigenvalues is a formidable task for large scale matrices even from a numerical point of view. It is therefore important to develop tools for efficient estimation of the spectral gap. To this aim, in this chapter, we deepen our analysis of time-reversible chains and we review a basic technique to estimate the second eigenvalue on the basis of the geometry of the underlying graph.

Consider a primitive time-reversible stochastic matrix $P$ associated with a conductance matrix $C$. Compactly we can write $P = D_{C\mathbb{1}}^{-1} C$ where $D_{C1}$ is the diagonal matrix such that $(D_{C1})_{uu} = (C\mathbb{1})_u$. The invariant probability is given by $\pi = (\mathbb{1} * C\mathbb{1})^{-1} C\mathbb{1}$. A simple verification shows that $P$ and $\pi$ satisfy the relation

$$\pi_u P_{uv} = \pi_v P_{vu} \tag{7}$$

for every pair of nodes $u$ and $v$. (7) is known as the *detailed balance condition* and is not difficult to see that it is actually a sufficient condition for time-reversibility. In the probabilistic framework, considering $P$ as the transition matrix of a random walk $X_t$ having initial probability vector $\pi$, the left and right terms of (7) can be interpreted, respectively as $\mathbb{P}(X_t = u, X_{t+1} = v)$ and $\mathbb{P}(X_t = v, X_{t+1} = u)$. This motivates the name time-reversible.

To investigate the spectral properties of $P$, it is convenient to introduce the so-called *Laplacian* of $P$ defined as $L(P) = I - P$. Clearly, $L(P)$ has all real non-negative eigenvalues and 0 has multiplicity 1. Define the inner product $\langle x, y\rangle_\pi :=< x, D_\pi y >= \sum_v \pi_v x_v y_v$. The following standard linear algebra results hold (see for instance [33]).

**Proposition 3.** *Assume that $P$ is a time-reversible stochastic matrix with invariant probability measure $\pi$. For every $x \in \mathbb{R}^V$, it holds*

$$\langle x, L(P)x\rangle_\pi = \frac{1}{2}\sum_{v,w} P_{vw}\pi_v(x_v - x_w)^2. \tag{8}$$

**Proposition 4.** *Assume that $P$ is a time-reversible stochastic matrix with invariant probability measure $\pi$. Let $\lambda_2$ be the second smallest eigenvalue of $L(P)$. It holds*

$$\lambda_2 = \min_{x\neq 0, \langle x, \mathbb{1}\rangle_\pi = 0} \frac{\langle x, L(P)x\rangle_\pi}{\langle x, x\rangle_\pi}. \tag{9}$$

A useful technique, based on the above results, to upper bound the spectral gap of a time-reversible stochastic matrix $P$ is through the so called *bottleneck ratio*, a sort of index measuring how well the "flow" represented by the matrix is spreading along the underlying graph. Suppose $\pi$ is the usual invariant probability measure of $P$, and for every $S \subset \mathcal{V}$, define $\pi(S) = \sum_{v\in S}\pi_v$ and

$$Q(S, S^c) = \sum_{v\in S, w\notin S} \pi_v P_{vw}\,.$$

Then, we define

$$\Phi(S) := \frac{Q(S, S^c)}{\pi(S)}$$

and the *bottleneck ratio* of $P$ as

$$\Phi_* := \min_{S:\pi(S)\leq\frac{1}{2}} \Phi(S).$$

In the probabilistic interpretation $Q(S, S^c)$ represents the probability that the random walk at the invariant regime $\pi$ is in $S$ at time $t$, and outside of $S$ at the next time $t+1$. $\Phi(S)$ is instead the probability that the random walk is outside of $S$ at time $t+1$ conditioned to be in $S$ at time $t$. We have the following result:

**Proposition 5** (Cheeger bound)**.** *Let $\mu_2$ be the second largest eigenvalue of a time-reversible matrix $P$, and let $\Phi_*$ be the bottleneck ratio of $P$. Then,*

$$1 - \mu_2 \leq 2\Phi_*. \tag{10}$$

**Proof** Given $S \subseteq V$, consider the vector $\phi \in \mathbb{R}^V$ defined by $\phi_v = \pi(S^c)$ if $v \in S$, and $\phi_v = -\pi(S)$ if $v \in S^c$. Then, from Proposition 3 and the detailed balance condition (7), it follows that

$$\begin{aligned}\langle \phi, L(P)\phi\rangle_\pi &= \tfrac{1}{2}\textstyle\sum_{v,w} \pi_v P_{vw}(\phi_v - \phi_w)^2\\ &= \textstyle\sum_{v\in S, w\notin S} \pi_v P_{vw}(\phi_v - \phi_w)^2\\ &= \textstyle\sum_{v\in S, w\notin S} \pi_v P_{vw}(\pi(S) + \pi(S^c))^2\\ &= \textstyle\sum_{v\in S, w\notin S} \pi_v P_{vw} = Q(S, S^c)\,.\end{aligned}$$

On the other hand,

$$\langle \phi, \phi\rangle_\pi = \sum_v \pi_v\phi_v^2 = \sum_{v\in S}\pi_v\pi(S^c)^2 + \sum_{w\notin S}\pi_w\pi(S)^2 = \pi(S)\pi(S^c).$$

From the variational characterization of Proposition 4, and assuming $\pi(S) \leq 1/2$ we thus conclude

$$1 - \mu_2 = \lambda_2 \leq \tfrac{\langle\phi, L(P)\phi\rangle_\pi}{\langle\phi,\phi\rangle_\pi} = \tfrac{Q(S,S^c)}{\pi(S)\pi(S^c)} \leq 2\Phi(S).$$

Since this inequality holds for all $S$ such that $\pi(S) \leq \frac{1}{2}$, the upper bound is proved. ■

Notice that, since $\rho_2 \geq \mu_2$, we also have a bound for the spectral gap

$$1 - \rho_2 \leq 2\Phi_* \,.$$

In the case when $P$ is the SRW on a symmetric graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the bottleneck ratio takes a peculiar form which is convenient to work out:

$$\Phi(S) = \frac{\displaystyle\sum_{v \in S, w \in S^c} \frac{d_v}{|\mathcal{E}|} (A_{\mathcal{G}})_{vw} \frac{1}{d_v}}{\displaystyle\sum_{v \in S} \frac{d_v}{|\mathcal{E}|}} = \frac{\sum_{v \in S, w \in S^c} (A_{\mathcal{G}})_{vw}}{\sum_{v \in S} d_v} \tag{11}$$

This says that $\Phi(S)$ equals the fraction of those edges which start inside $S$ and end outside $S$. The corresponding $\phi_*$ in this case is also called the bottleneck of the graph $\mathcal{G}$.

**Example 4** (Graphs with a bottleneck)**.** *Consider two symmetric strongly connected graphs $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$ with $|\mathcal{V}_1| = |\mathcal{V}_2| = n$. Fix two points $v_1 \in \mathcal{V}_1$ and $v_2 \in \mathcal{V}_2$ and consider the graph*

$$\mathcal{G} = (\mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E}_1 \cup \mathcal{E}_2 \cup \{(v_1, v_2), (v_2, v_1)\}) \,.$$

*If we take $S = \mathcal{V}_1$, we obtain*

$$\Phi(S) = \frac{1}{|\mathcal{E}_1| + 1}$$

*Taking into consideration that (since $\mathcal{G}_1$ is connected) $|\mathcal{E}_1| \geq n - 1$, we obtain the following universal upper bound on the spectral gap of the SRW on $\mathcal{G}$:*

$$1 - \rho_2 \leq \frac{2}{n}$$

*This implies that the convergence time is at least of the order of $n$. Even if the SRW on the two graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ converge in a much faster way, the interconnected graph can not beat the bound due to the presence of the bottleneck created in the interconnected graph. An extreme case is represented by the case when $\mathcal{G}_1$ and $\mathcal{G}_2$ are two copies of the complete graph over $n$ vertices. In that case things are even worse since*

$$\Phi(S) = \frac{1}{|\mathcal{E}_1| + 1} = \frac{1}{n^2 + 1}$$

*so that the convergence time is of the order of $n^2$ in spite of the fact that in the two complete graphs convergence would have happened in finite time! This final example is known as the barbell graph and will be considered again in later chapters.*

A sort of lower bound involving the bottleneck ratio can also be obtained (see [33, Theorem 13.14].)

**Proposition 6.** *Let $\mu_2$ be the second largest eigenvalue of a time-reversible matrix $P$, and let $\Phi_*$ be the bottleneck ratio of $P$. Then,*

$$1 - \mu_2 \geq \frac{{\Phi_*}^2}{2}\,. \tag{12}$$

Apparently, this seems not to be sufficient to establish a lower bound for the spectral gap since one only knows that the second eigenvalue satisfy the inequality $\rho_2 \geq \mu_2$. However, in many cases it is actually an equality. An important case is when $P$ is a SRW and we consider the lazy SRW $P_{1/2} = 1/2(I + P)$.

Unfortunately, the two bounds (10) and (12) do not close: in all cases when the bottleneck ratio goes to 0 as $\alpha(N)$ (where $N$ is the number of nodes), the two estimations yield bounds on the convergence time of type $C_1\alpha(N)^{-1} < \tau < C_2\alpha(N)^{-2}$, and could not conclude, by this road, the exact scaling of the convergence time.

An important application of Proposition 6 is for constructing examples of high performance consensus dynamics. A sequence of symmetric graphs $\mathcal{G}_N$ over $N$ nodes, is called an *expander graph* if the corresponding bottleneck ratios $\Phi_*^{(N)}$ are bounded away from 0. Thanks to Proposition 6, SRW on expander graphs have the spectral gap bounded away from 0 and therefore their convergence time turns out to be constant in $N$. One might guess that such expander graphs guarantee such good properties at the price of increasing the degree of the nodes, but, surprisingly enough, there are examples of expander graphs where the degree of all nodes remain bounded with respect to the size $N$. Construction of such graphs is typically done by random techniques [3] and their analysis is beyond the scope of this paper.

Many other useful techniques to bound the spectral gap of time reversible matrices, for instance through comparing arguments with respect to some other time reversible matrix, can be found in [33].

# 5 Random models

The dynamical systems studied so far are based on the assumption that units share a common clock and update their state in a synchronous fashion. This is of course not a feasible assumption in many contexts, if we reflect on the fact that the problem of synchronizing clocks in a network of sensors [9] is of (at least) comparable complexity than implementing a consensus algorithm. Moreover, in the opinion dynamics modelling, it is not realistic to assume that all interactions happen at the same time: agents are embedded in a physical continuous time and interactions can be imagined to take place at different times, for instance in a pairwise fashion.

This motivates the study of different asynchronous models where typically some random process determines the possible interactions among the agents. Of course, randomness is not strictly necessary in the definition of asynchronous models, but is, nevertheless, a useful and, in many cases, realistic assumption. Randomness is akin to a nice mathematical analysis and this is the reason why such models have obtained such a great popularity in the last ten years [4, 19].

## 5.1 Random consensus dynamics

We now describe a basic model for random consensus dynamics subsuming most of the models of interest in the applications. Given a set of nodes $\mathcal{V}$ of finite cardinality $N$, we consider a sequence of random independent and identically distributed stochastic matrices $P(t) \in \mathbb{R}^{\mathcal{V}\times\mathcal{V}}$. They induce the random dynamics

$$x(t+1) = P(t)x(t) \qquad t \in \mathbb{N}\,. \tag{13}$$

In these notes we think of $x(0) \in \mathbb{R}^{\mathcal{V}}$as a fixed (not random) initial condition. $x(t)$ is thus a stochastic process where all the randomness is in the choice of the matrices $P(t)$'s.

We now present some basic examples

**Example 5** (The symmetric gossip model)**.** *This is probably the most famous random consensus model. Fix a real number $q \in (0,1)$ and a symmetric graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. At every time instant $t$ an edge $(u,v) \in \mathcal{E}$ is activated with uniform probability $|\mathcal{E}|^{-1}$ and nodes $u$ and $v$ exchange their states and produce a new state according to the equations*

$$\begin{aligned} x_u(t+1) &= (1-q)x_u(t) + qx_v(t) \\ x_v(t+1) &= qx_u(t) + (1-q)x_v(t) \end{aligned}$$

*The states of the other units remain unchanged. More formally for every $(u,v) \in \mathcal{E}$, we let*

$$R^{uv} = I - q(e_u - e_v)(e_u - e_v)^*\,,$$

*(we recall that $e_u$ is the vector with $1$ in position $u$ and all other components equal to $0$). Then, $P(t)$ is concentrated on these matrices and*

$$\mathbb{P}[P(t) = R^{uv}] = |\mathcal{E}|^{-1}$$

**Example 6** (The asymmetric-gossip model)**.** *In this case we start from a real number $q \in (0,1)$ and a fixed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. At every time instant $t$ an agent $u$ is activated with uniform probability $|\mathcal{V}|^{-1}$ and, subsequently, an out neighbor $v$ of $u$ is chosen with probability $1/d_u^+$ ($d_u^+$ is the out-degree of node $u$, namely the number of outgoing edges from $u$). Then, node $v$ sends its state to $u$ and $u$ produces a new state according to the equation*

$$x_u(t+1) = (1-q)x_u(t) + qx_v(t)$$

*Formally, define, for every $(u,v) \in \mathcal{E}$, $R^{uv} = I - qe_u(e_u - e_v)^*$ and let*

$$\mathbb{P}[P(t) = R^{uv}] = |\mathcal{V}|^{-1}(d_v^+)^{-1}$$

*This model can also be considered when the parameter $q = 1$. In this case agent $u$ simply copies the state of agent $v$: it is called the 'voter model' in opinion dynamics.*

We may as well consider non-gossip models:

**Example 7** (The broadcasting model)**.** *We start from any graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. We assume that at every time instant node $u$ is chosen with uniform probability $|\mathcal{V}|^{-1}$. This node $u$ then broadcasts its state to all its out-neighbors which then average their states with the received state. In this case $P(t)$ concentrates on the $N$ matrices*

$$R^u = I - q \sum_{v \in N_u^+} (e_v e_v^* - e_u e_v^*)$$

*and we let $\mathbb{P}[P(t) = R^u] = |\mathcal{V}|^{-1}$. Further considerations on this model can be found in [24, 22]*

Finally, randomness can also be due environmental effects. In the next example we describe some possible models for such situations.

**Example 8** (The packet drop model)**.** *We start from a fixed stochastic matrix $P$ such that $P_{uu} > 0$ for all $u$ and such that the graph $\mathcal{G}_P$ is strongly connected. Then we know that the algorithm*

$$x(t+1) = Px(t) \tag{14}$$

*yields consensus. In some situations there might be data loss in the communication between nodes. Here we model this by assuming that the data transmission over an edge $(v, u)$ of $\mathcal{G}$ from the node $v$ to the node $u$ can occur with some probability $p$. More precisely, consider the family of independent binary random variables $L_{uv}(t)$, $t \in \mathbb{N}$, $u, v = 1, \dots, N$, $u \neq v$, such that*

$$\mathbb{P}[L_{uv}(t) = 1] = p \quad \text{and} \quad \mathbb{P}[L_{uv}(t) = 0] = 1 - p \quad \textit{if } u \neq v$$

*We emphasize the fact that independence is assumed among all $L_{uv}(t)$ as $u, v$ and $t$ vary. Consider the random matrix $\tilde{A}(t)$ defined by $\tilde{A}_{uv}(t) = (A_{\mathcal{G}})_{uv} L_{uv}(t)$. Clearly, $\tilde{A}(t)$ is the adjacency matrix of a random graph $\tilde{\mathcal{G}}(t)$ obtained from $\mathcal{G}$ by deleting the edge $(u, v)$ when $L_{uv}(t) = 0$. In general our initial matrix $P$ is not going to be compatible with the graph $\bar{\mathcal{G}}(t)$ and, as a consequence, the consensus algorithm has to be modified to consider this fact. There are several ways to adapt $P$ in order to take into account the missing data. One possibility is to consider*

$$\begin{aligned} x_i(t+1) &= \Big(P_{uu} + \textstyle\sum_{\substack{v=1 \\ v \neq u}}^{N} P_{uv}(1 - L_{uv}(t))\Big) x_u(t) \\ &+ \textstyle\sum_{\substack{v=1 \\ v \neq u}}^{N} P_{uv} L_{uv}(t) x_v(t) \end{aligned}$$

*Roughly speaking, according to this method the weights $P_{uv}$ are kept constant if $u \neq v$ while they are varied if $u = v$ in order to keep the stochasticity of $P(t)$. More details and different models can be found in [20].*

We now comments on the examples introduced above

**Remark:** Notice how, in spite of our original intention to de-synchronize the dynamics, in the proposed setting and in all the examples above we have maintained a global integer time. Moreover activation of nodes or edges is governed by a random but centralized mechanism. These apparently strong drawbacks are actually not really crucial. Indeed, one could consider the scenario where nodes or edges are equipped with clocks whose ticking are modeled by independent Poisson random variables. As simultaneous clicking is a zero measure event, one can easily get back to the proposed model by interpreting time as a logical one which simply counts the global number of clock ticks in the network. Notice, moreover, that the assumption of uniformity in the sample probability can be generalized to different probabilistic models.

## 5.2 Probabilistic consensus

As the above example very well illustrate the requirement that the dynamics (13) always converges to a consensus would be too strong and typically not verified (e.g. in the symmetric gossip model when the same edge is sampled at every instant of time). More realistic is to require that $x(t)$ converges to a consensus with probability one. In order to investigate this situation, it is convenient to consider the mean dynamics $e(t+1) = \bar{P}e(t)$ where $\bar{P} := \mathbb{E}[P(t)]$ is the mean stochastic matrix and where $e(t) = \mathbb{E}[x(t)]$. Clearly, if $x(t)$ converges to a consensus with probability one, then, being bounded, also $e(t)$ will converge to a consensus. Therefore a natural (indeed necessary if $\mathcal{G}_{\bar{P}}$ is symmetric ) assumption to be made for $x(t)$ to achieve consensus with probability one is $\bar{P}$ to be primitive. While this is not sufficient to guarantee convergence of the process, a slightly stronger assumption does, as is exactly stated in the following result [19]:

**Theorem 7.** *Consider the dynamics (13). Assume that*

1. $\mathcal{G}_{\overline{P}}$ *is strongly connected,*
2. *for any* $u \in V$ *we have that* $P(t)_{uu} > 0$ *almost surely.*

*Then,* $x(t)$ *achieves consensus with probability one.*

It is immediate to check that all the examples presented above satisfy the assumptions of Theorem 7 as long as the underlying graph is strongly connected. Regarding Example 6 on the asymmetric gossip model when $q = 1$ (the voter model), notice instead that assumption 2. in Theorem 7 is not verified. For this model we can indeed prove convergence to a consensus but with a completely different argument. The basic fact is that in the voter model, state of the agents at every time belong to a fixed finite set determined by the initial states; evolution can thus be modeled as a finite Markov chain possessing absorbing states (corresponding to all agents sharing the same state). Since from every possible initial configuration state, if the graph is connected, there is a possible path leading to an absorbing states, classical results of Markov chains imply that consensus will be reached in finite time with probability one. Further details on the voter model and related reference can be found in [12]

There are two basic issues which are not addressed by previous result. First, the consensus point is in principle a random variable in general not directly computable. Can we compute its mean and its variance? Second, can we characterize the speed of convergence to consensus as in the deterministic case?

Let us start with the first point. Put $Q(t) = P(t)P(t-1)\cdots P(0)$. The fact that, for $t \to +\infty$, $Q(t)x(0) \to \bar{x}\mathbb{1}$ with probability 1, easily implies that, for $t \to +\infty$, $Q(t) \to \mathbb{1}\rho^*$ with probability 1 where $\rho \in \mathbb{R}^{\mathcal{V}}$ is necessarily a random probability vector. We thus have that the consensus point can be written as $\bar{x} = \rho^* x(0)$. The mean $\mathbb{E}[\rho]$ can immediate be characterized by the mean matrix $\bar{P}$: indeed $\mathbb{E}[\rho]$ is the invariant probability of $\bar{P}$. In order to analyze the second moment of $\rho$, is necessary to pass from the mean dynamics governed by $\bar{P}$ to the covariance dynamics of the process. Notice that we can write

$$\mathbb{E}[x^*(t)x(t)] = x^*(0)\Delta(t)x(0)$$

where $\Delta(t) := \mathbb{E}[Q(t-1)^*Q(t-1)]$ if $t \geq 1$ and where $\Delta(0) := I$. A simple recursive argument shows that

$$\Delta(t+1) = \mathbb{E}[P(0)^*\Delta(t)P(0)]$$

(we are using the fact that the $P(t)$'s are identically distributed). This shows that $\Delta(t)$ is the evolution of a linear dynamical system which can be written in the form

$$\Delta(t+1) = \mathcal{L}(\Delta(t))$$

where $\mathcal{L} : \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}$ is given by

$$\mathcal{L}(M) = \mathbb{E}[P(0)^*MP(0)]$$

If we consider on $\mathbb{R}^{N\times N}$ the classical inner product $< A, B >= \mathrm{Trace}(B^*A)$, we easily obtain that the adjoint operator $\mathcal{L}^*$ is given by $\mathcal{L}^*(M) = \mathbb{E}[P(0)MP(0)^*]$ which turns out to be a stochastic operator ($\mathcal{L}^*(\mathbb{1}\mathbb{1}^*) = \mathbb{1}\mathbb{1}^*$). On the other hand, it holds

$$\mathcal{L}^t(M) = \mathbb{E}[Q(t-1)^*MQ(t-1)] \to \mathbb{E}[\rho\mathbb{1}^*M\mathbb{1}\rho^*] = \mathbb{E}[\rho\rho^*]\mathbb{1}^*M\mathbb{1}$$

This implies that $\mathcal{L}^*$ is primitive and the unique invariant probability is given by $\mathbb{E}[\rho\rho^*]$. Therefore the first two moments of the random invariant probability $\rho$ can be characterized in terms of the invariant probabilities of the two deterministic operators $\bar{P}$ and $\mathcal{L}$.

The operator $\mathcal{L}$ also plays a crucial role in the analysis of the speed of convergence to consensus. Indeed, if we consider the following mean square distance from the consensus

$$d(t) = \mathbb{E}||x(t) - N^{-1}\mathbb{1}\mathbb{1}^*x(t)||^2\,,$$

it is immediate to check that $d(t) = x(0)^* \mathcal{L}^t(\Omega) x(0)$ where $\Omega = I - N^{-1}\mathbb{1}\mathbb{1}^*$. In [19] it is shown that the rate of convergence to 0 of $d(t)$ is dictated by the spectral radius of the operator $\mathcal{L}$ when restricted to the space of symmetric matrices $M$ such that $M\mathbb{1} = 0$. A useful upper bound can be obtained by the following argument.

Notice that

$$\begin{aligned} x^* \mathbb{E}[P^*(0)\Omega P(0)]x &= \mathbb{E}[x^* \Omega P^*(0)\Omega P(0)\Omega x] \\ &\leq ||\mathbb{E}[P^*(0)\Omega P(0)]|| x^* \Omega x \end{aligned}$$

This shows that $||\mathcal{L}(\Omega)||\Omega \geq \mathcal{L}(\Omega)$ (where order relation is the one of positive definitiveness) Iterating this inequality we find that

$$\mathcal{L}^t(\Omega) \leq ||\mathcal{L}(\Omega)||^t \Omega$$

This yields the bound

$$d(t) \leq \gamma^t ||\Omega x(0)||^2 \tag{15}$$

where $\gamma$ is the spectral radius of the matrix $\mathcal{L}(\Omega)$.

In general the analysis of specific examples need a considerable amount of computation and efforts. Below we shortly present a couple of examples and we refer to [4, 19, 20, 22, 18, 25] for a detailed analysis of these and of many more examples. Related models are also considered in [6, 10]

**Example 9** (Performance of the symmetric gossip)**.** *In the case of the symmetric gossip algorithm, all matrices $P(t)$ are indeed symmetric. As a consequence $\rho = N^{-1}\mathbb{1}$: consensus point is always (deterministically) the arithmetic mean of the initial states. Regarding the rate of convergence, notice that in this case*

$$\mathcal{L}(\Omega) = \mathbb{E}[P^*(t)\Omega P(t)] = \mathbb{E}[P^2(t)] - \frac{1}{N}\mathbb{1}\mathbb{1}^*$$

*In the special case where $q = 1/2$, we have that $P^2(t) = P(t)$. Therefore $\gamma$ coincides with the send eigenvalue of the matrix $\bar{P}$. It can be shown that the bound (15) is tight in this case [4, 19]. In many examples, the computation of the second eigenvalue of $\bar{P}$ is an affordable problem. If we further specialize to the case when $\mathcal{G}$ is the complete graph, we obtain*

$$\bar{P} = \frac{1}{N^2}\sum_{u,v}\left[I - \frac{1}{2}(e_u - e_v)(e_u - e_v)^*\right] = \left(1 - \frac{1}{N}\right) I + \frac{1}{N} N^{-1}\mathbb{1}\mathbb{1}^*$$

*Therefore, for the complete graph, $\gamma = 1 - N^{-1}$. This implies that the convergence time (in mean square sense) scales as $N$. This deterioration of performance with respect to the synchronous case where the convergence time is constant, is only due to the chosen time scale: in one unit of time here only two nodes update their value, while, in the synchronous case, all nodes update. If we were using the Poisson clock model in continuous time where each edge is activated at rate 1, then this discrepancy would disappear.*

**Example 10** (Performance of the asymmetric gossip)**.** *We assume the underlying graph $\mathcal{G}$ to be symmetric. Notice that the matrices $P(t)$ are in any case no longer symmetric. The mean stochastic matrix is given by*

$$\bar{P} = \frac{1}{N^2}\sum_{u,v}(I - q(e_u(e_u - e_v)^*) = I - \frac{q}{N^2}D_{A_{\mathcal{G}}\mathbb{1}} + \frac{q}{N^2}A_{\mathcal{G}}$$

*which is symmetric. This implies that the random probability vector $\rho$ determining the consensus point has its mean $\mathbb{E}[\rho] = N^{-1}\mathbb{1}$ even if $\rho$ will not in general be equal to the uniform probability. Indeed a lengthy, but not difficult, computation shows that the second moment is given by*

$$\mathbb{E}[\rho\rho^*] = \frac{1}{qN + (1-q)N^2}[qI + (1-q)\mathbb{1}\mathbb{1}^*]$$

*Remarkably, the second moment does not depend on the topology of the underlying graph. From this, we can for instance compute the mean displacement of the consensus point from the arithmetic mean:*

$$\mathbb{E}|\rho^*x(0) - N^{-1}\mathbb{1}^*x(0)|^2 = x(0)^*[\mathbb{E}[\rho\rho^*] - N^{-2}\mathbb{1}\mathbb{1}^*]x(0) = \frac{q}{qN + (1-q)N^2}||\Omega x(0)||^2$$

*Notice how the mean displacement goes to $0$ as $N \to +\infty$. This says that, in spite of the fact that the matrices $P(t)$ are not doubly stochastic, the consensus point concentrate around the arithmetic mean in large scale graphs. For more details on this example including detailed analysis of the speed of convergence, we refer the reader to [18].*

# 6 Stubborn agents and electrical networks

## 6.1 Consensus dynamics with stubborn agents

In this chapter, we investigate consensus dynamics models where some of the agents do not modify their own state (stubborn agents). These models are of interest in socio-economic models [2] and also in vehicle rendezvous problems where certain vehicles want to remain fixed and make the other gather around them [31]

Consider a symmetric connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. We assume a splitting $\mathcal{V} = \mathcal{S} \cup \mathcal{R}$ with the understanding that agents in $S$ are *stubborn* agents not changing their state while those in $\mathcal{R}$ are *regular* agents whose state modifies in time according to the dynamics described below. Let $P \in \mathbb{R}^{\mathcal{V}\times\mathcal{V}}$ be a stochastic matrix such that, for $u \neq v$

$$P_{uv} = 0 \;\Leftrightarrow \{u, v\} \notin \mathcal{E} \;\text{or}\; i \in S \tag{16}$$

We will say that $P$ is adapted to the pair $(\mathcal{G}, \mathcal{S})$ when it satisfies the constraint above. Dynamics of opinions is described by the relation $x(t+1) = Px(t)$. If we order elements in $\mathcal{V}$ in such a way that elements in $\mathcal{R}$ come first, the matrix $P$ will exhibit the block structure:

$$P = \begin{bmatrix} Q^{11} & Q^{12} \\ 0 & I \end{bmatrix} \tag{17}$$

Splitting accordingly the state vector $x(t) = (x^{\mathcal{R}}(t), x^{\mathcal{S}}(t)) \in \mathbb{R}^{\mathcal{V}}$ we thus have dynamics

$$\begin{array}{rcl} x^{\mathcal{R}}(t+1) & = & Q^{11}x^{\mathcal{R}}(t) + Q^{12}x^{\mathcal{S}}(t) \\ x^{\mathcal{S}}(t+1) & = & x^{\mathcal{S}}(t) \end{array} \tag{18}$$

Notice that $Q^{11}$ is a sub stochastic matrix, namely all row sums are $\leq 1$. Moreover, thanks to the adaptivity assumption (16), there is at least one row whose sum is strictly less than one (the row corresponding to a regular agent connected to a stubborn one). Using the connectivity of the graph, this easily implies that there exists $t > 0$ such that $(Q^{11})^t$ has the property that all its rows have sum strictly less than one. This immediately yields that the matrix is asymptotically stable (e.g. spectral radius $< 1$). Henceforth, $x^{\mathcal{R}}(t) \to x^{\mathcal{R}}(\infty)$ for $t \to +\infty$ with the limit opinions satisfying the relation

$$x^{\mathcal{R}}(\infty) = Q^{11}x^{\mathcal{R}}(\infty) + Q^{12}x^{\mathcal{S}}(0) \tag{19}$$

which is equivalent to

$$x^{\mathcal{R}}(\infty) = (I - Q^{11})^{-1}Q^{12}x^{\mathcal{S}}(0) \tag{20}$$

Put $\Xi := (I - Q^{11})^{-1}Q^{12}$ and notice that $\Xi_{us} = \sum_n [(Q^{11})^n Q^{12}]_{us}$ is always non negative and is not equal to 0 if and only if there exists a path in $\mathcal{G}$ connecting the regular agent $u$ to the stubborn agent $s$ and not touching other stubborn agents. Moreover, the fact that $P$ is stochastic easily implies that $\sum_s \Xi_{us} = 1$ for all $u \in R$: asymptotic opinions of regular agents are thus convex combinations of the opinions of stubborn agents. The above analysis shows, in particular, that if all stubborn agents are in the same state $x$ (for instance this happens if there is just one stubborn agent), then, consensus is reached by all agents in the point $x$. However, typically, consensus is not reached in such a system: we will discuss few examples later on.

## 6.2 Electrical Networks

In order to deepen our study of the asymptotic states of a consensus dynamics with stubborn agents, it is convenient to make use of the so called electrical network interpretation of a graph [17, 33]: this leads to the appealing interpretation of the asymptotic states as voltages when the state of the stubborn agents are fixed voltage. Below we describe this interpretation in some detail closely following the presentation in [34].

Given a symmetric strongly connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , denote by $\bar{\mathcal{E}}$ the set of *undirected* edges of $\mathcal{G}$: namely $\bar{\mathcal{E}}$ consists of those subsets $\{u, v\}$ of cardinality 2 such that $(u, v) \in \mathcal{E}$ (possible self-loops present in $\mathcal{G}$ are disregarded in the construction of $\bar{\mathcal{E}}$). An *incidence matrix* on $\mathcal{G}$ is any matrix $B \in \{0, +1, -1\}^{\bar{\mathcal{E}} \times \mathcal{V}}$ such that $B\mathbb{1} = 0$ and $B_{eu} \neq 0$ iff $u \in e$. It is immediate to see that given $e = \{u, v\}$, the $e$-th row of $B$ has all zeroes except $B_{eu}$ and $B_{ev}$: necessarily one of them will be $+1$ and the other one $-1$ and this will be interpreted as choosing a direction in $e$ from the node corresponding to $+1$ to the one corresponding to $-1$. Consider now a symmetric conductance matrix $C \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$ such that $C_{uv} \geq 0$ for every $u, v \in \mathcal{V}$ and $C_{uv} > 0$ iff $\{u, v\} \in \mathcal{E}$ and define $D_C \in \mathbb{R}^{\bar{\mathcal{E}} \times \bar{\mathcal{E}}}$ to be the diagonal matrix such that $(D_C)_{ee} = C_{uv} = C_{vu}$ if $e = \{u, v\}$.

Interpret now $\mathcal{G}$ as an electrical circuit where an edge $\{u, v\}$ has electrical conductance $C_{uv} = C_{vu} > 0$ (and thus resistance $R_{uv} = C_{uv}^{-1}$) while $C_{uv} = 0$ for all pairs $(u, v) \notin \mathcal{E}$. The pair $(\mathcal{G}, C)$ will be called an *electrical network* from now on. Consider moreover a vector $\eta \in \mathbb{R}^V$ such that $\eta^* \mathbb{1} = 0$: we interpret $\eta_v$ as the input current injected at node $v$ (if negative being an outgoing current). Given $C$ and $\eta$, we can define the voltage $W \in \mathbb{R}^{\mathcal{V}}$ and the current flow $\Phi \in \mathbb{R}^{\bar{\mathcal{E}}}$ in such a way that the usual Kirchoff and Ohm's law are satisfied on the network. Compactly, they can be expressed as

$$\begin{cases} B^* \Phi = \eta \\ D_C B W = \Phi \end{cases} \tag{21}$$

(Notice that $\phi_e$ is the current flowing on edge $e$ and sign is positive iff flow is along the conventional direction individuated by $B$ on edge $e$). Coupling the two equations we obtain the following equation for $W$:

$$B^* D_C B W = \eta \tag{22}$$

A straightforward calculation shows that $B^* D_C B = D_{C\mathbb{1}}(I - D_{C\mathbb{1}}^{-1} C)$. Notice that $Q = D_{C\mathbb{1}}^{-1} C$ is the time-reversible Markov chain associated with $C$ and that the part in parenthesis is thus the so called Laplacian of $Q$ which we have previously denoted $L(Q)$. Relation (22) can thus be rewritten as

$$L(Q) W = D_{C\mathbb{1}}^{-1} \eta \tag{23}$$

It follows from Proposition 3 that, since $\mathcal{G}$ is connected, $L(Q)$ has rank $N - 1$ with $L(Q)\mathbb{1} = 0$. This shows that (23) determines $W$ up to translations.

It is often possible to replace an electrical network with a simplified one without changing certain quantities of interest. An useful operation is *gluing:* if we merge vertices having the same voltage into a single one, while keeping all voltages and currents unchanged, because current never flows between vertices with the same voltage. Another useful operation is replacing a portion of the electrical network connecting two nodes $u, v$ by an *equivalent resistance*, a single resistance denoted as $R_{u,v}^{eff}$ which keeps the difference of voltage $W(u) - W(v)$ unchanged. There are two basic laws to compute equivalent resistances. One is the so called Series law: if $u \in V$ is a node of degree 2 with neighbors $v$ and $w$ and $\eta_u = 0$, then edges $\{u, v\}$ and $\{u, w\}$ can be replaced by a single edge $\{v, w\}$ of resistance $R_{v,w}^{eff} = R_{uv} + R_{uw}$. All voltages and currents in the new network remain the same as in the original one, while the current that flows from $v$ to

$w$ equals $\Phi_{\{v,w\}} = \Phi_{\{u,v\}} = \Phi_{\{u,w\}}$ and the voltage difference between $v$ and $w$ is unchanged. The other one is the Parallel law: suppose edges $e_1$ and $e_2$, with conductances $c_1$ and $c_2$, respectively, share vertices $u$ and $v$ as endpoints. Then both edges can be replaced with a single edge $e$ of conductance $c_1 + c_2$ without affecting the rest of the network. All voltages and currents in $\bar{\mathcal{E}} \backslash \{e_1, e_2\}$ are unchanged and the current $\Phi_e$ equals $\Phi_{e_1} + \Phi_{e_2}$. In terms of resistances we have a single edge with equivalent resistance of $(c_1 + c_2)^{-1}$.

## 6.3 Opinions as voltages

We are now ready to state the relationship between electrical networks and the consensus dynamics with stubborn agents. Consider a symmetric strongly connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a subset of stubborn agents $\mathcal{S} \subseteq \mathcal{V}$, and a conductance matrix $C$ on $\mathcal{G}$. Assume that the corresponding time-reversible matrix $Q = D_{C\mathbb{1}}^{-1} C$ is primitive. Let $P$ be the stochastic matrix coinciding with $Q$ on regular agents and such that $P_{ss} = 1$ for all $s \in \mathcal{S}$. Notice that relation (19) can be written as

$$L(Q) \begin{pmatrix} x^{\mathcal{R}}(\infty) \\ x^{\mathcal{S}}(0) \end{pmatrix} = \begin{pmatrix} 0 \\ \theta \end{pmatrix} \tag{24}$$

for some vector $\theta \in \mathbb{R}^S$. Confronting with (23), this implies that $x^{\mathcal{R}}(\infty)$ can be interpreted, in the electrical network $(\mathcal{G}, C)$, as the vector of voltages of the regular agents when stubborn agents have fixed voltage $x^{\mathcal{S}}(0)$ or, equivalently, when input currents $\eta = D_{C\mathbb{1}}\theta$ are applied to the stubborn agents. Notice that the translation ambiguity of the voltages implicit in relation (23), is here completely solved by the fact that stubborn agents have fixed voltage.

Thanks to the electrical analogy we can compute the asymptotic opinion of the agents by computing the voltages in the graph seen as an electrical network. Consider the following simple example.

**Example 11** (Stubborn agents in a Line graph)**.** *Consider the line graph $L_N = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, 2, \ldots, N\}$ and where $\mathcal{E} = \{(u, u+1), (u+1, u) \,|\, u = 1, \ldots, N-1\}$. Assume that $\mathcal{S} = \{1, N\}$ and $\mathcal{R} = \mathcal{V} \setminus \mathcal{S}$. Consider the stochastic matrix $P$ corresponding to the SRW on regular nodes in $\mathcal{R}$, while $P_{11} = P_{NN} = 1$. Following the considerations above, it follows that the asymptotic states of the regular agents $x^{\mathcal{R}}(\infty)$ correspond to the voltages of the regular nodes when nodes 1 and $N$ are kept at fixed voltage, respectively, $x_1^{\mathcal{S}}(0)$ and $x_N^{\mathcal{S}}(0)$ in the electrical network $(L_N, A_{L_N})$ (all edges have conductance equal to 1). By repeated application of the Series law, the line can be replaced by a single edge connecting 1 and $N$ having equivalent resistance $N-1$. Therefore, by Ohm's law, the current flowing along the replaced edge from 1 to $N$ (and thus, by Kirchoff's law, along all edges in the original network) is equal to $\Phi = (N-1)^{-1}[x_N^{\mathcal{S}}(0) - x_1^{\mathcal{S}}(0)]$. If we now fix an arbitrary node $v \in \mathcal{V}$ and applying again the same arguments in the part of graph from 1 to $v$, we obtain that the voltage at $v$, $x_v^{\mathcal{R}}(\infty)$ satisfies the relation $x_v^{\mathcal{R}}(\infty) - x_1^{\mathcal{S}}(0) = \Phi(v-1)$. We thus obtain*

$$x_v^{\mathcal{R}}(\infty) = x_1^{\mathcal{S}}(0) + \frac{v-1}{N-1}[x_N^{\mathcal{S}}(0) - x_1^{\mathcal{S}}(0)]\,.$$

## References

[1] D. Acemoglu, A. Ozdaglar, "Opinion dynamics and learning in social networks", *Dynamic Games and Applications*, vol. 1(1), pp. 3–49, 2011.

[2] D. Acemoglu, G. Como, F. Fagnani, A. Ozdaglar, "Opinion fluctuations and disagreement in social networks", *Mathematics of Operation Research*, Vol. 38 (1), pp.1-27, 2013.

[3] N. Alon and Y. Roichman, "Random Cayley graphs and expanders", *Random Structures & Algorithms*, vol. 5(2), pp. 271–284, 1994.

[4] S. Boyd, A. Ghosh, B. Prabhakar, D. Shah, "Randomized gossip algorithms", *IEEE Transactions on Information Theory*, vol. 52(6), pp. 2508–2530, 2006.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers", *Foundations and Trends in Machine Learning*, vol. 3(1), pp. 1124, 2011.

[6] M. Cao, D. A. Spielman, E. M. Yeh, "Accelerated gossip algorithms for distributed computation", in Proceedings of the 44th Annual Allerton Conference on Communication, Control, and Computation, Monticello, IL, USA, Sep. 2006.

[7] R. Carli, F. Fagnani, A. Speranzon, S. Zampieri, "Communication constraints in the average consensus problem", *Automatica*, vol. 44(3), pp. 671–684, 2008.

[8] R. Carli, A. Chiuso, L. Schenato, S. Zampieri, "Distributed Kalman filtering based on consensus strategies", *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 622–633, 2008.

[9] R. Carli, S. Zampieri, "Networked clock synchronization based on second order linear consensus algorithms", IEEE Conference on Decision and Control, 2010.

[10] R. Carli, F. Fagnani, P. Frasca, S. Zampieri, "Gossip consensus algorithms via quantized communication", *Automatica*, vol. 46, pp. 70-80, 2010.

[11] R. Carli, A. Chiuso, L. Schenato, S. Zampieri, "Optimal Synchronization for Networks of Noisy Double Integrators", *IEEE Transactions on Automatic Control*, vol. 56(5), pp. 1146-1152, 2011.

[12] C. Castellano, S. Fortunato, V. Loreto, "Statistical physics of social dynamics", *Review of Modern Physics*, vol. 81, pp. 591–646, 2009.

[13] A. Chiuso, F. Fagnani, L. Schenato, S. Zampieri (2011) "Gossip algorithms for simultaneous distributed estimation and classification in sensor networks", *IEEE J. Selected Topics in Signal Processing*, vol. 5 n. 4, pp. 691-706, 2011.

[14] F. Cucker, S. Smale, "Emergent behavior in flocks" *IEEE Transactions on Automatic Control*, vol. 52(5), pp. 852–862, 2007.

[15] S. Currarini, M. O. Jackson, P. Paolo, "An economic model of friendship: homophily, minorities, and segregation", *Econometrica*, vol. 77(4), pp. 1003–1045, 2009.

[16] G. Deffuant, D. Neau, F. Amblard, G. Weisbuch, "Mixing beliefs among interacting agents", *Advances in Complex Systems*, vol. 3, pp. 87–98, 2000.

[17] P.G. Doyle and J.L. Snell, "Random Walks and Electric Networks", Carus Monographs. Mathematical Association of America, 1984.

[18] F. Fagnani, S. Zampieri, "Asymmetric randomized gossip algorithms for consensus", *IFAC World Conference*, pp. 9052-9056, 2008.

[19] F. Fagnani, S. Zampieri, "Randomized consensus algorithms over large scale networks", *IEEE Journal on Selected Areas of Communications*, vol. 26, pp. 634–649, 2008.

[20] F. Fagnani, S. Zampieri, ", Average consensus with packet drop communication", *SIAM J. Control Optim.*, vol. 48, pp. 102-133, 2009.

[21] F. Fagnani, S. Fosson, C. Ravazzi, "Input driven consensus algorithm for distributed estimation and classification in sensor networks", *IEEE Conference on Decision and Control and European Control Conference*, Orlando, Florida, Dec. 12-15, 2011. pp. 6654-6659, 2011.

[22] F. Fagnani, P. Frasca, "Broadcast gossip averaging: interference and unbiasedness in large Abelian Cayley networks", *IEEE J. Selected Topics in Signal Processing*, vol. 5 n. 4, pp. 866-875, 2011.

[23] J. A. Fax, R. M. Murray, "Information flow and cooperative control of vehicle formations", *IEEE Transaction on Automatic Control*, vol. 49(9), pp. 1465–1476, 2004.

[24] M. Franceschelli, A. Giua, C. Seatzu, "Consensus on the average on arbitrary strongly connected digraphs based on broadcast gossip algorithms",1st IFAC Workshop on Estimation and Control of Networked Systems, Venice, Italy, Sett. 2009.

[25] P. Frasca and J. M. Hendrickx, "On the mean square error of randomized averaging algorithms", *Automatica*, to appear.

[26] F. Galton, "Vox populi", *Nature*, vol. 75, pp. 450–451, 1907.

[27] F. R. Gantmacher, *The theory of matrices*, Chelsea Publishers, New York, 1959.

[28] B. Golub, M. O. Jackson, "Naïve learning in social networks and the wisdom of crowds", *American Economic Journal: Microeconomics*, vol. 2(1), pp. 112–149, 2010.

[29] A. Jadbabaie, J. Lin, A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules" *IEEE Transactions on Automatic Control*, vol. 48(6), pp. 988–1001, 2003.

[30] M. O. Jackson, *Social and economic networks*, Princeton University Press, 2008.

[31] M. Ji, G. Ferrari-Trecate, M. Egerstedt, A. Buffa, "Containment control in mobile networks" *IEEE Transactions on Automatic Control*, vol. 53(8), pp. 1972–1975, 2008.

[32] U. Krause, "A discrete nonlinear and non-autonomous model of consensus formation", *Communications in Difference Equations*, pp. 227–236, S. Elaydi, G. Ladas, J. Popenda, and J. Rakowski editors, Gordon and Breach, Amsterdam, 2000.

[33] D. A. Levin, Y. Peres, E. L. Wilmer *Markov chains and mixing times.* AMS, 2008.

[34] E. Lovisari, S. Zampieri. "Performance metrics in the average consensus problem: a tutorial", *Annual Reviews in Control*, 2012

[35] L. Moreau, "Stability of multiagent systems with time-dependent communication links", *IEEE Transactions on Automatic Control*, vol. 50, pp. 169–182, 2005.

[36] S. Muthukrishnan, B. Ghosh, M.H. Schultz, "First- and second-order diffusive methods for rapid, coarse, distributed load balancing", *Theory of Computing Systems*, vol. 31, pp. 331–354, 1998.

[37] M. E. J. Newman, *Networks: an introduction*, Oxford University Press, 2010.

[38] P. Niyogi, *The Computational nature of language, learning, and evolution*, MIT Press, 2006.

[39] R. Olfati-Saber, J.A. Fax, R.M. Murray, "Consensus and cooperation in networked multi-agent systems", *Proceedings of the IEEE*, vol. 95(1), pp. 215–233, 2007.

[40] L. Scardovi and R. Sepulchre, "Synchronization in networks of identical linear systems", *Automatica*, vol. 45, pp. 2557–2562, 2009.

[41] S. H. Strogatz, "From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators", *Physica D: Nonlinear Phenomena*, vol. 143(1–4), pp. 1–20, 2000.

[42] S. H. Strogatz, *Sync: The emerging science of spontaneous order*, Hyperion, 2003.

[43] J. Surowiecki, *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*, Little, Brown, 2004. Traduzione italiana: *La saggezza della folla*, Fusi Orari, 2007.

[44] J. Tsitsiklis, *Problems in decentralized decision making and computation*, Ph.D. thesis, Department of EECS, MIT, 1984.

[45] T. Vicsek, A. Czirk, E. Ben-Jacob, O. Shochet, "Novel type of phase transition in a system of self-driven particles", *Physical Review Letters*, vol. 75, pp. 1226–1229, 1995.

[46] J. Wolfers, E. Zitzewitz, "Prediction markets", *Journal of Economic Perspectives*, vol. 18(2), pp. 107–126, 2004.

[47] L. Xiao, S. Boyd, S. Lall, "A scheme for robust distributed sensor fusion based on average consensus", *International Conference on Information Processing in Sensor Networks*, pp. 63–70, 2005.