Communication Constraints in the Average Consensus Problem

Ruggero Carli [a], Fabio Fagnani [b], Alberto Speranzon [c], Sandro Zampieri [a]

[a] *Department of Information Engineering, Università di Padova, Via Gradenigo 6/a, 35131 Padova, Italy*

[b] *Dipartimento di Matematica, Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129 Torino, Italy*

[c] *School of Electrical Engineering, Royal Institute of Technology, Osquldasväg 10, SE-10044 Stockholm, Sweden*

In memory of Antonio Lepschy

Abstract

The interrelationship between control and communication theory is becoming of fundamental importance in many distributed control systems, such as the coordination of a team of autonomous agents. In such a problem, communication constraints impose limits on the achievable control performance. We consider as instance of coordination the consensus problem. The aim of the paper is to characterize the relationship between the amount of information exchanged by the agents and the rate of convergence to the consensus. We show that time-invariant communication networks with circulant symmetries yield slow convergence if the amount of information exchanged by the agents does not scale well with their number. On the other hand, we show that randomly time-varying communication networks allow very fast convergence rates. We also show that, by adding logarithmic quantized data links to time-invariant networks with symmetries, control performance significantly improves with little growth of the required communication effort.

Key words: **Consensus, Multi-agent Coordination, Convergence Rate, Logarithmic Quantization, Random Networks, Mixing Rate of Markov Chains**

1 Introduction

Multi-agent systems have many advantages compared to single-agent systems, including improved flexibility, sensing and reliability. When it comes to design control strategies for coordination, mobile agent systems need to be able to exchange information, such as the position, velocity, or other relevant quantities to solve a given task. For the coordination to be effective they need to rapidly reach a consensus on the shared data. The problem of designing strategies that guarantee the shared data to convergence (asymptotically) to common value is called *coordinated consensus* or *state agreement* problem. From the seminal work by Tsitsiklis [45], Olfati-Saber and Murray [36] and Jadbabaie et al. [24], in which the consensus problem was firstly defined, in system theoretical terms, the field has rapidly grown and attracted the attention of many researchers, see for example [43,17,26,18,39,41,29], and the recent survey paper [35]. The interest in these type of problems is not limited to the field of mobile agents coordination but also involves problems of synchronization [42,28,27] and distributed estimation [31,9].

Most of the literature is concerned with the design of control strategies that yield consensus. In the classical framework, each agent is modelled as an omnidirectional antenna with a short reliable communication range [24,43,10]. This results in a communication network whose topology changes with the agents' position. Design and analysis of decentralized control laws for these systems are in general hard tasks. One of the main difficulties is that the connectivity of the network is not guaranteed to be preserved under dynamical constraints. Simplified models have been proposed in [24,36,38] where the authors consider switching systems with switching rule that does not dependent on the agents' position and for which they derive only sufficient conditions for consensus. In [33], in the context of multi-agent flocking, virtual potential functions are

Email addresses: `carlirug@dei.unipd.it` (Ruggero Carli), `fabio.fagnani@polito.it` (Fabio Fagnani), `alberto.speranzon@unilever.com` (Alberto Speranzon), `zampi@unipd.it` (Sandro Zampieri).

[1] A. Speranzon was with Royal Institute of Technology, Sweden. He is currently with Unilever R&D Port Sunlight, UK.

also used in order to constraint the agents to form particular lattices, thus relaxing the connectivity condition of the networks. A similar approach is considered in [43] where authors use tools from non-smooth analysis to design and analyze consensus controllers. Robustness to communication link failure [10] and the effects of time delays [36] have been also considered.

The aim of this paper is to characterize the relationship between the amount of information exchanged by the agents and the achievable control performance. We model the communication network by a directed graph, in which an arc represents information transmission from one agent to another one. With this model the amount of information exchanged, or communication effort, is related to the number of neighbors of each agent. If we consider convergence rate to the average value of the initial conditions as control performance index, we expect that the more the graph is connected the better the performance. The main result of the paper is a mathematical characterization of this fact.

We assume that the graph topology is independent of the relative agents' positions and we analyze both deterministic time-invariant communication graphs (as in [17,41,18]) and stochastically varying communication graphs (as in [22]). Furthermore, since the focus of the paper is on how communication affects coordination, we assume that the agents are described by a first order model, as considered in [24,36,10]. This results in a tractable mathematical problem although some ideas can be partially extended to more general linear models.

We first study time-invariant communication networks. Under some assumptions, described in sections 2 and 3, it turns out that weighted directed graphs, for which the adjacency matrix is doubly stochastic, are communication graphs that guarantee the average consensus, with a degree of efficiency that is related to the spectral properties of such matrix. Such matrix can be interpreted as a Markov chain. The consensus convergence rate turns out to be related to the mixing rate of the chain, for which bounds are available in literature [4]. Here we have gathered them and presented from a different viewpoint. Spectral properties of doubly stochastic matrices can be characterized in a easier way if we impose symmetries on the matrices themselves, and thus on the associated communication graph. Markov chains and graphs satisfying symmetries, called Cayley graphs, are widely studied in the literature [3,30,46]. It is known that symmetries described by Abelian groups yield rather poor convergence rates [2]. By modelling the communication network as Cayley graphs defined on Abelian groups we determine a new bound on the consensus convergence rate. This extends available bounds on the mixing rate of Markov chains defined on such groups [12,4,40]. The main result, presented in section 4, shows that, imposing symmetries in the communication network, and thus in the control structure, yields convergence rates that degrades, as the number of agents increases, if the amount of information exchanged by the agents does not scale well with their total number.

The idea of imposing symmetries on the communication graph is not new [11,37,41]. In particular in [41] the authors show, for particular symmetries, that it is possible to obtain better performance by increasing the number of incoming arcs on each vertex. Further results have been obtained in [9]. In this contribution we extend these results to a broader class of graphs with symmetries and we propose a tight bound on the performance that is achievable in this case.

In section 5 we consider stochastically time-varying solutions. In these strategies the communication graph is chosen randomly at each time step over a family of graphs with the constraint that the number of incoming arcs in each vertex is constant. A mean square analysis shows that we can improve the convergence rate obtained with fixed communication graphs. This fact continue to hold true even if the random choice is restricted to families of Cayley graphs. In this case, compared to time-invariant solutions, imposing symmetries does not yield a performance degradation. A similar analysis has been proposed in [22,9] where a different model of randomly time-varying communication graph was proposed.

Another important contribution of the paper, described in section 6, consists in using other types of data transmission in coordinated control. More precisely, we introduce in the communication graph another type of arc that represents transmission of logarithmic quantized data. Exact data transmission is very expensive with respect to the required communication rate and it is well-known [14,15] that logarithmic quantization allows a more efficient use of the available communication bandwidth. A preliminary analysis of coordinated control strategies involving logarithmic quantized data transmission has been proposed in [25]. The analysis is very complicated in general whereas it is tractable for Cayley graphs. Through some examples it is showed that logarithmic quantized data transmission improves substantially the control performance with a limited increase of the total bandwidth.

2 Problem Formulation

Consider $N > 1$ identical systems whose dynamics are described by the following discrete time state equations

$$x_i^+ = x_i + u_i \qquad i = 1, \ldots, N,$$

where $x_i \in \mathbb{R}$ is the state of the i-th system, x_i^+ represents the updated state and $u_i \in \mathbb{R}$ is the control input. More compactly we can write

$$x^+ = x + u, \tag{1}$$

where $x, u \in \mathbb{R}^N$. The goal, in the consensus problem, is the design of a feedback control law $u = Kx$ with $K \in \mathbb{R}^{N\times N}$ such that, for any initial condition $x(0) \in \mathbb{R}^N$, the closed loop system $x^+ = (I + K)x$ yields

$$\lim_{t\to\infty} x(t) = \alpha \mathbb{1} \tag{2}$$

where $\mathbb{1} := (1, \ldots, 1)^T$ and where α is a scalar depending on $x(0)$ and K.

The fact that in the matrix K the element i, j is different from zero, means that the system i needs the state of the system j in order to compute its feedback action and thus communication needs to occur between the systems. A good description of the information flow required by a specific feedback K is given by the directed graph $\mathcal{G}_K$ with set of vertices $\{1, \ldots, N\}$ in which there is an arc from j to i whenever in the feedback matrix K the element $K_{ij} \neq 0$. The graph $\mathcal{G}_K$ is said to be the *communication graph* associated with K. Conversely, given any directed graph $\mathcal{G}$ with set of vertices $\{1, \ldots, N\}$, we say that a feedback K is *compatible* with $\mathcal{G}$ if $\mathcal{G}_K$ is a subgraph of $\mathcal{G}$ (we use the notation $\mathcal{G}_K \subseteq \mathcal{G}$). We say that the consensus problem is solvable on a graph $\mathcal{G}$ if there exists a feedback K compatible with $\mathcal{G}$ and solving the consensus problem. From now on we always assume that $\mathcal{G}$ contains all loops (i, i) since each system has access to its own state.

With such model of the network, we are interested in obtaining a matrix K compatible with a given graph, yielding the consensus and maximizing a suitable performance index. The simplest control performance index is the exponential rate of convergence to the consensus point. Clearly, any effective feedback matrix K must ensure that nonzero states having equal components correspond to equilibrium points of the closed loop system, because in this case no control action is necessary. This happens if and only if $K\mathbb{1} = 0$. From now on we impose this condition on K. In this context it is easy to see that the consensus problem is solved if and only if the following three conditions hold:

(i) 1 is the only eigenvalue of $I + K$ on the unit circle centered in 0;
(ii) the eigenvalue 1 has algebraic multiplicity one and $\mathbb{1}$ is its eigenvector;
(iii) all the other eigenvalues are strictly inside the unit disk centered in 0.

Under these conditions the convergence rate can be defined as follows. Let P be any matrix such that $P\mathbb{1} = \mathbb{1}$ and assume that its spectrum $\sigma(P)$ is contained in the closed unit disk centered in 0. We define the *essential spectral radius* of P as

$$\rho(P) = \max\{|\lambda| \text{ s.t. } \lambda \in \sigma(P) \setminus \{1\}\}\,. \tag{3}$$

As in [35], the goal this paper is to clarify the relation between the graph connectivity and $\rho(P)$. An interesting particular case considered in the literature is the average consensus [36]. This corresponds to a situation where the control law yields the consensus at the average of the initial states. These control laws are called average consensus controllers in [36]. It is easy to see that K is an average consensus controller if and only if $\mathbb{1}^T K = 0$. Notice that this condition is automatically true for symmetric matrices K satisfying $K\mathbb{1} = 0$. From this choice of performance we can formulate the following control problem: Given a graph $\mathcal{G}$, find a matrix K such that $K\mathbb{1} = 0$, $\mathbb{1}^T K = 0$, $\mathcal{G}_K \subseteq \mathcal{G}$ and minimizing $\rho(I + K)$.

When we are dealing with average consensus controllers it is meaningful to consider the displacement from the average, or disagreement vector as defined in [36],

$$\Delta(t) := x(t) - \left(N^{-1}\mathbb{1}^T x(0)\right)\mathbb{1}\,. \tag{4}$$

Since $K\mathbb{1} = 0$, $\Delta(t)$ satisfies the closed loop equation

$$\Delta^+ = (I + K)\Delta\,. \tag{5}$$

Moreover, since $\mathbb{1}^T K = 0$, we have that $\mathbb{1}^T x(t) = \mathbb{1}^T x(0)$ for every t. In this case we can represent

$$\Delta(t) = x(t) - \left(N^{-1}\mathbb{1}^T x(t)\right)\mathbb{1}\,.$$

Notice finally that the initial conditions $\Delta(0)$ are such that

$$\mathbb{1}^T \Delta(0) = 0\,. \tag{6}$$

Hence the asymptotic behavior of our consensus problem can equivalently be studied by looking at the evolution (5) on the hyperplane characterized by condition (6). The index $\rho(I + K)$ seems in this context appropriate for analyzing how performance is related to the communication effort associated to a graph.

3 Doubly Stochastic Matrices in Consensus

If we restrict to control laws K making $I + K$ a nonnegative matrix, namely a matrix with all elements nonnegative, condition $K\mathbb{1} = 0$ imposes that $I + K$ is a stochastic matrix. If, moreover, we also have $\mathbb{1}^T K = 0$, then $I + K$ is doubly stochastic. Since the spectral structure of stochastic and doubly stochastic matrices is quite well known, this observation allows to understand easily what conditions on the graph ensure the solvability of the consensus problem. To exploit this we need to recall some notation and results on directed graphs (the reader can further refer to textbooks on graph theory such as [21] or [13]).

Fix a directed graph $\mathcal{G}$ with set of vertices V and set of arcs $\mathcal{E} \subseteq V \times V$. The adjacency matrix A is a $\{0, 1\}$-valued square matrix indexed by the elements in V defined by letting $A_{ij} = 1$ if and only $(i, j) \in \mathcal{E}$. Define the in-degree of a vertex j as $\operatorname{indeg}(j) := \sum_i A_{ij}$ and the out-degree of a vertex i as $\operatorname{outdeg}(i) := \sum_j A_{ij}$. Vertices with out-degree equal to 0 are called sinks. A graph is called in-regular (out-regular) of degree k if each vertex has in-degree (out-degree) equal to k. A path in $\mathcal{G}$ consists of a sequence of vertices $i_1 i_2 \ldots\ldots i_r$ such that $(i_\ell, i_{\ell+1}) \in \mathcal{E}$ for every $\ell = 1, \ldots, r-1$; i_1 (resp. i_r) is said to be the initial (resp. terminal) vertex of the path. A cycle is a path in which the initial and the terminal vertices coincide. A vertex i is said to be connected to a vertex j if there exists a path with initial vertex i and terminal vertex j. A directed graph is said to be connected if, given any pair of vertices i and j, either i is connected to j or j is connected to i. A directed graph is said to be strongly connected if, given any pair of vertices i and j, i is connected to j.

Given any directed graph $\mathcal{G}$ we can consider its strongly connected components, namely maximal strongly connected subgraphs $\mathcal{G}_k$, $k = 1, \ldots, s$, with set of vertices $V_k \subseteq V$ and set of arcs $\mathcal{E}_k = \mathcal{E} \cap (V_k \times V_k)$ such that the sets V_k form a partition of V. The various components may have connections among each other. We define another directed graph $T_\mathcal{G}$ with set of vertices $\{1, \ldots, s\}$ such that there is an arc from h to k if there is an arc in $\mathcal{G}$ from a vertex in V_k to a vertex in V_h. It can be shown that $T_\mathcal{G}$ is a graph without cycles. The following proposition is the straightforward consequence of a standard results on stochastic matrices [20, pag. 88 and pag. 95].

Proposition 1 *Let $\mathcal{G}$ be a directed graph and assume that $\mathcal{G}$ contains all loops (i,i). Then, the consensus problem is solvable on $\mathcal{G}$ iff $T_\mathcal{G}$ is connected and has only one sink vertex. Moreover, if the above conditions are satisfied, any K such that $I + K$ is stochastic, $\mathcal{G}_K = \mathcal{G}$ and $K_{ii} \neq -1$ for every $i = 1, \ldots, n$ solves the consensus problem.*

Remark 2 *An analogous result has been proposed in [38] where the solvability of the consensus is related to the existence of a spanning tree in the graph $\mathcal{G}$.*

When the graph $\mathcal{G}$ satisfies the properties of Proposition 1, a particularly simple solution of the consensus problem can be obtained by defining P as follows

$$P_{ij} = \begin{cases} 1/\operatorname{indeg}(i) & \text{if } (j,i) \in \mathcal{E}, \\ 0 & \text{otherwise} \end{cases}$$

and by letting $K := P - I$. In this case the closed loop dynamics have the following form

$$x_i^+ = x_i + \frac{1}{\operatorname{indeg}(i)} \sum_{j \neq i,\, (j,i) \in \mathcal{E}} (x_j - x_i)\,. \qquad (7)$$

Again, if we restrict to K such that $I + K$ is nonnegative, we can relate the existence of average consensus controllers to the structure of the graph by mean of standard results on stochastic matrices.

Proposition 3 *Let $\mathcal{G}$ be a directed graph and assume that $\mathcal{G}$ contains all loops (i,i). Then, the average consensus problem is solvable on $\mathcal{G}$ iff $\mathcal{G}$ is strongly connected. Moreover, if the above conditions are satisfied, any K such that $I + K$ is doubly stochastic, $\mathcal{G}_K = \mathcal{G}$ and $K_{ii} \neq -1$ for every $i = 1, \ldots, n$ solves the average consensus problem.*

Notice that, in the special case when the graph $\mathcal{G}$ is undirected, namely $(i,j) \in \mathcal{E}$ if and only if $(j,i) \in \mathcal{E}$, we can find solutions K to the consensus problem that are symmetric and that therefore are automatically doubly stochastic. One example can be obtained as follows [35]. Let A be the adjacency matrix of the undirected graph $\mathcal{G}$, which is a symmetric matrix. Take the Laplacian matrix $L \in \mathbb{R}^{N \times N}$ of $\mathcal{G}$ being defined by letting $L_{ij} = -A_{ij}$ if $i \neq j$ and by letting $L_{ii} = \sum_{k \neq i} A_{ik} = \operatorname{indeg}(i) - 1 = \operatorname{outdeg}(i) - 1$. Then we have that $K = -\epsilon L$ yields the average consensus for all ϵ such that $0 < \epsilon < (\max_i \{\operatorname{indeg}(i) - 1\})^{-1}$. Instead, again taking a undirected graph $\mathcal{G}$, in general the law given by (7) does not yield a symmetric matrix K and neither an average consensus controller.

When P is a stochastic matrix, the problem of minimizing the essential spectral radius $\rho(P)$ or, equivalently, of maximizing $1 - \rho(P)$ (which is called the spectral gap of the associated Markov chain) over the matrices P's compatible with a given graph is a very classical problem in the theory of Markov chains and recently some very effective algorithms have been proposed for this maximization limited to the case in which P is a symmetric matrix [7].

4 Symmetric Controllers

The analysis of the consensus problem and the corresponding controller synthesis becomes more treatable if we limit our search to graphs $\mathcal{G}$ and matrices K exhibiting symmetries. We show, however, that these symmetries limit the achievable performance in terms of convergence rate.

In order to treat symmetries on a graph $\mathcal{G}$ in a general setting, we introduce the concept of Cayley graph defined on Abelian groups [3,2]. Let G be any finite Abelian group (internal operation will always be denoted $+$) of order $|G| = N$, and let S be a subset of G containing zero. The Cayley graph $\mathcal{G}(G,S)$ is the directed graph with vertex set G and arc set $\mathcal{E} = \{(g,h) : h - g \in S\}$. Notice that a Cayley graph is always in-regular, the in-degree of each vertex is $|S|$. Notice a Cayley graph $\mathcal{G}(G,S)$ is strongly connected if and only if the set S generates the group G. If S is such that $-S = S$ we say that S is inverse-closed. In this case the graph obtained is undirected.

Symmetries can be introduced also on matrices. Let G be any finite Abelian group of order N. A matrix $P \in \mathbb{R}^{G \times G}$ is said to be a Cayley matrix over the group G if

$$P_{i,j} = P_{i+h,j+h} \qquad \forall\, i,j,h \in G\,.$$

It is clear that for a Cayley matrix P there exists a $\pi : G \to \mathbb{R}$ such that $P_{i,j} = \pi(i-j)$. The function π is called the generator of the Cayley matrix P. Notice that, if π and π' are generators of the Cayley matrices P and P' respectively, then $\pi + \pi'$ is the generator of $P + P'$ and $\pi * \pi'$ is the generator of PP', where $(\pi * \pi')(i) := \sum_{j\in G} \pi(j)\pi'(i-j)$ for all $i \in G$. This shows that P and P' commute. Notice finally that, if P is a Cayley matrix generated by π, then $\mathcal{G}_P$ is a Cayley graph with $S = \{h \in G : \pi(h) \neq 0\}$. Moreover, it is easy to see that for any Cayley matrix P we have that $P\mathbb{1} = \mathbb{1}$ if and only if $\mathbb{1}^T P = \mathbb{1}^T$. This implies that a Cayley stochastic matrix is automatically doubly stochastic. In this case the function π associated with the matrix P is a probability distribution on the group G. Among the multiple possible choices of the probability distribution π, there is one which is particularly simple, namely $\pi(g) = 1/|S|$ for every $g \in S$. Given a Cayley graph $\mathcal{G}$ we can define

$$\rho_{\mathcal{G}}^{\text{Cayley}} = \min\{\rho(I+K) | I+K \text{Cayley stochastic}, \mathcal{G}_K \subseteq \mathcal{G}\}\,.$$

It turns out that $\rho_{\mathcal{G}}^{\text{Cayley}}$ can be evaluated or estimated in many cases. Moreover, it clearly holds that $\rho_{\mathcal{G}}^{\text{Cayley}} \geq \rho_{\mathcal{G}}^{\text{ds}}$. Before continuing we give some short background notions on group characters and on harmonic analysis on groups, which are the basis of our main results.

4.1 Cayley stochastic matrices on finite Abelian groups

We briefly review the theory of Fourier transform over finite Abelian groups (see [44] for a comprehensive treatment of the topic). Let G be a finite Abelian group of order N, and let $\mathbb{C}^*$ be the multiplicative group of the nonzero complex numbers. A character on G is a group homomorphism $\chi : G \to \mathbb{C}^*$, namely a function χ from G to $\mathbb{C}^*$ such that $\chi(g+h) = \chi(g)\chi(h)$ for all $g, h \in G$. Since we have that $\chi(g)^N = \chi(Ng) = \chi(0) = 1$ for any $g \in G$, it follows that χ takes values on the N^{th}-roots of unity. The character $\chi_0(g) = 1$ for every $g \in G$ is called the trivial character. The set of all characters of the group G forms an Abelian group with respect to the pointwise multiplication. It is called the character group and denoted by $\hat{G}$. The trivial character χ_0 is the zero of $\hat{G}$. Moreover, $\hat{G}$ is isomorphic to G, and its cardinality is N. If we consider the vector space $\mathbb{C}^G$ of all functions from G to $\mathbb{C}$ with the canonical Hermitian form

$$< f_1, f_2 >= \sum_{g\in G} f_1(g) f_2(g)^*\,,$$

it follows that the set $\{N^{-1/2}\chi \mid \chi \in \hat{G}\}$ is an orthonormal basis of $\mathbb{C}^G$. The Fourier transform of a function $f : G \to \mathbb{C}$ is defined as

$$\hat{f} : \hat{G} \to \mathbb{C}\,, \quad \hat{f}(\chi) = \sum_{g\in G} \chi(-g) f(g)\,.$$

The cyclic case is instrumental to study characters for any finite Abelian group. Indeed, a well known result in algebra [23] states that any finite Abelian group G is isomorphic to a finite direct sum of cyclic groups. Hence, in order to study characters of G, we can assume that $G = \mathbb{Z}_{N_1} \oplus \cdots \oplus \mathbb{Z}_{N_r}$. It can be shown [4] that the characters of G are precisely the maps $(g_1, g_2, \ldots, g_r) \mapsto \chi^{(1)}(g_1)\chi^{(2)}(g_2)\cdots\chi^{(r)}(g_r)$ with $\chi^{(i)} \in \hat{\mathbb{Z}}_{N_i}$ for $i = 1, \ldots, r$. In other terms, $\hat{G}$ is isomorphic to $\hat{\mathbb{Z}}_{N_1} \oplus \cdots \oplus \hat{\mathbb{Z}}_{N_r}$.

Fix now a Cayley matrix P on the Abelian group G generated by the function $\pi : G \to \mathbb{R}$. The spectral structure of P is very simple. To see this, first notice that P can be interpreted as a linear function from $\mathbb{C}^G$ to itself simply by considering, for $f \in \mathbb{C}^G$, $(Pf)(g) := \sum_h P_{gh} f(h)$. Notice that the trivial character χ_0 corresponds to the vector $\mathbb{1}$ having all components equal to 1. It is easy to see that each character χ is an eigenfunction of P with eigenvalue $\hat{\pi}(\chi)$. Since the characters form an orthonormal basis it follows that P is diagonalizable and its spectrum is given by $\sigma(P) = \{\hat{\pi}(\chi) \mid \chi \in \hat{G}\}$. We can interpret a character as a linear function from $\mathbb{C}$ to $\mathbb{C}^G$ as the mapping $\chi : \mathbb{C} \to \mathbb{C}^G : z \mapsto z\chi$. Its adjoint is the linear functional $\chi^* : \mathbb{C}^G \to \mathbb{C}\; f \mapsto < f, \chi >$. With this notation, $N^{-1}\chi\chi^*$ is a linear function from $\mathbb{C}^G$ to itself, projecting $\mathbb{C}^G$ on the eigenspace generated by χ. In this way, P can be represented as

$$P = \sum_{\chi\in\hat{G}} \hat{\pi}(\chi) N^{-1} \chi\chi^*\,.$$

Conversely, it can easily be shown that, given any $\hat{\theta} : \hat{G} \to \mathbb{C}$, the matrix

$$P = \sum_{\chi\in\hat{G}} \hat{\theta}(\chi) N^{-1} \chi\chi^*\,,$$

is a Cayley matrix generated by the Fourier transform $\hat{\theta}$. Suppose now that $P = I + K$ is the closed loop matrix of the system. The displacement from the average $\Delta(t)$, defined in (4), can be represented as

$$\Delta(t) = (I - N^{-1}\chi_0\chi_0^*)x\,.$$

Notice that since (5) and (6) hold, then

$$\Delta(t) = P^t\Delta(0) = \frac{1}{N}\sum_{\chi\neq\chi_0} \hat{\pi}(\chi)^t \chi < \Delta(0), \chi > \,.$$

Hence,

$$||\Delta(t)||^2 = \frac{1}{N}\sum_{\chi\neq\chi_0} |\hat{\pi}(\chi)|^{2t} | < \Delta(0), \chi > |^2\,.$$

This shows in a very simple way, in this case, the role of $\rho(P) = \max_{\chi\neq\chi_0} |\hat{\pi}(\chi)|$ in the rate of convergence.

4.2 The essential spectral radius of Cayley matrices.

The particular spectral structure of Cayley matrices allows to obtain asymptotic results on the behavior of the essential spectral radius $\rho(P)$ and therefore on the rate of convergence of the corresponding control scheme. Let us start from some examples.

Example 4 *Consider the group $\mathbb{Z}_N$ and the Cayley graph $\mathcal{G}(\mathbb{Z}_N, S)$, where $S = \{0, 1\}$. Consider the probability distribution π on S described by $\pi(0) = 1 - k$ and $\pi(1) = k$, where $k \in [0, 1]$. The characters are given by*

$$\chi_\ell(j) = e^{i\frac{2\pi}{N}\ell j}, \quad j \in \mathbb{Z}_N, \quad \ell = 0, \ldots, N-1.$$

The Fourier transform of π is

$$\hat{\pi}(\chi_\ell) = \sum_{g\in S} \chi(-g)\pi(g) = 1 - k + ke^{-i\frac{2\pi}{N}\ell},$$

with $\ell = 1, \ldots, N-1$. It can be shown that we have consensus stability if and only if $0 < k < 1$ and, in this case we have

$$\rho_{\mathcal{G}}^{\text{Cayley}} = \min_k \max_{1\le\ell\le N-1} \left|1 - k + ke^{-i\frac{2\pi}{N}\ell}\right|.$$

The optimality is obtained when $\ell = 1$ and $k = 1/2$ yielding

$$\rho_{\mathcal{G}}^{\text{Cayley}} = \left(\frac{1}{2} + \frac{1}{2}\cos\left(\frac{2\pi}{N}\right)\right)^{\frac{1}{2}} \simeq 1 - \frac{\pi^2}{2}\frac{1}{N^2}$$

where the last approximation is for $N \to \infty$.

Example 5 *Consider the group $\mathbb{Z}_N$ and the Cayley graph $\mathcal{G}(\mathbb{Z}_N, S)$, where $S = \{-1, 0, 1, \}$. For the sake of simplicity we assume that N is even; similar results can be obtained for odd N. Consider the probability distribution π on S described by $\pi(0) = k_0$, $\pi(1) = k_1$, and $\pi(-1) = k_{-1}$. The Fourier transform of π is in this case given by*

$$\hat{\pi}(\chi_\ell) = \sum_{g\in S} \chi(-g)\pi(g) = k_0 + k_1 e^{-i\frac{2\pi}{N}\ell} + k_{-1}e^{i\frac{2\pi}{N}\ell},$$

with $\ell = 1, \ldots, N-1$. We thus have

$$\rho_{\mathcal{G}}^{\text{Cayley}} = \min_{(k_0,k_1,k_{-1})} \max_{1\le\ell\le N-1} \left|k_0 + k_1 e^{-i\frac{2\pi}{N}\ell} + k_{-1}e^{i\frac{2\pi}{N}\ell}\right|.$$

Symmetry and convexity arguments [8,6] allow to conclude that a minimum is of the type $k_1 = k_{-1}$. With this assumption the minimum is achieved for

$$k_0 = \frac{1 - \cos\left(\frac{2\pi}{N}\right)}{3 - \cos\left(\frac{2\pi}{N}\right)}, \quad k_1 = k_{-1} = \frac{1}{3 - \cos\left(\frac{2\pi}{N}\right)}$$

and we have

$$\rho_{\mathcal{G}}^{\text{Cayley}} = \frac{1 + \cos\left(\frac{2\pi}{N}\right)}{3 - \cos\left(\frac{2\pi}{N}\right)} \simeq 1 - 2\pi^2\frac{1}{N^2} \tag{8}$$

where the last approximation is meant for $N \to \infty$.

Notice that in the first example the optimality is obtained when all the nonzero elements of π are equal. This is not a general feature since the same does not happen in the second example. Notice moreover that in this example, as N tends to infinity, the optimal solution tends to $k_0 = 0, k_1 = k_{-1} = 1/2$. This shows that the solution that optimizes $\rho_{\mathcal{G}}^{\text{Cayley}}$ can be very different from the law suggested in (7). The case of communication exchange with two neighbors (example 5) offers a better performance compare to the case with one neighbor (example 4). However, in both cases $\rho_{\mathcal{G}}^{\text{Cayley}} \to 1$ for $N \to +\infty$. This fact is more general: if we keep bounded the number of incoming arcs in a vertex, the essential spectral radius for Abelian stochastic Cayley matrices always converges to 1. This negative behavior has already been noticed in the literature [6,41,32,9]. In [32] it is shown that some random rewiring can correct this slow convergence rate. The next result provides a bound which proves that this bad performance is a general feature of this class of consensus algorithms.

Theorem 6 *Let G be any finite Abelian group of order N and $S \subseteq G$ be a subset containing zero. Let moreover $\mathcal{G}$ be the Cayley graph associated with G and S. If $|S| = \nu + 1$, then*

$$\rho_{\mathcal{G}}^{\text{Cayley}} \ge 1 - CN^{-2/\nu}, \tag{9}$$

where $C > 0$ is a constant independent of G and S.

In order to prove theorem 6 we need the following technical lemma.

Lemma 7 *Let $\mathbb{T} = \mathbb{R}/\mathbb{Z} \cong [-1/2, 1/2[$. Let $0 \le \delta < 1/2$ and consider the hypercube $V = [-\delta, \delta]^k \subseteq \mathbb{T}^k$. For every finite set $\Lambda \subseteq \mathbb{T}^k$ such that $|\Lambda| \ge \delta^{-k}$, there exist $\bar{x}_1$, $\bar{x}_2 \in \Lambda$ with $\bar{x}_1 \neq \bar{x}_2$ such that $\bar{x}_1 - \bar{x}_2 \in V$.*

PROOF. For any $x \in \mathbb{T}$ and $\delta > 0$, define the following set

$$L(x, \delta) = [x, x + \delta] + \mathbb{Z} \subseteq \mathbb{T}.$$

Observe that for all $y \in \mathbb{T}$, $L(x, \delta) + y = L(x + y, \delta)$. Now let $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_k) \in \mathbb{T}^k$ and define

$$L(\bar{x}, \delta) = \prod_{i=1}^k L(\bar{x}_i, \delta).$$

Also in this case we observe that $L(\bar{x}, \delta) + \bar{y} = L(\bar{x} + \bar{y}, \delta)$ for every $\bar{y} \in \mathbb{T}^k$. Consider now the family of subsets

$$\{L(\bar{x}, \delta), \ \bar{x} \in \Lambda\}.$$

We claim that there exist $\bar{x}_1$ and $\bar{x}_2$ in Λ such that $\bar{x}_1 \neq \bar{x}_2$ and such that $L(\bar{x}_1,\delta)\cap L(\bar{x}_2,\delta)\neq\emptyset$. Indeed, if not, we would have that

$$1 \geq m\left(\bigcup_{\bar{x}\in\Lambda} L(\bar{x},\delta)\right) = \sum_{\bar{x}\in\Lambda} m\left(L(\bar{x},\delta)\right) = |\Lambda|\delta^k \geq 1$$

where $m(\cdot)$ is the Lebesgue measure on $\mathbb{T}^k$ and where we used the hypothesis $|\Lambda| \geq \delta^{-k}$. However, since all $L(\bar{x}_1,\delta)$ are closed, it is not possible that $m\left(\bigcup_{\bar{x}\in\Lambda} L(\bar{x}_1,\delta)\right)=1$. Notice finally that

$$\begin{aligned} L(\bar{x}_1,\delta)\cap L(\bar{x}_2,\delta)\neq\emptyset \;&\Leftrightarrow\; L(0,\delta)\cap L(\bar{x}_2-\bar{x}_1,\delta)\neq\emptyset \\ &\Leftrightarrow\; \bar{x}_2-\bar{x}_1\in V\,. \end{aligned}$$

□

PROOF. [Theorem 6] With no loss of generality we can assume that $G=\mathbb{Z}_{N_1}\oplus\ldots\oplus\mathbb{Z}_{N_r}$. Assume we have fixed a probability distribution π supported on S. Let P be the corresponding stochastic Cayley matrix. It follows from previous considerations that the spectrum of P is given by

$$\sigma(P)=\Big\{\sum_{k_1=0}^{N_1-1}\cdots\sum_{k_r=0}^{N_r-1}\pi(k_1,\ldots,k_r)e^{i\frac{2\pi}{N_1}k_1\ell_1}\cdots e^{i\frac{2\pi}{N_r}k_r\ell_r} \\ :\ell_1\in\mathbb{Z}_{N_1},\ldots,\ell_r\in\mathbb{Z}_{N_r}\Big\}$$

Denote by $\bar{k}^j=(k_1^j,\ldots,k_r^j)$, for $j=1,\ldots,\nu$, the non-zero elements in S, and consider the subset

$$\Lambda=\Big\{\left(\sum_{i=1}^r\frac{k_i^1\ell_i}{N_i},\ldots,\sum_{i=1}^r\frac{k_i^\nu\ell_i}{N_i}\right)+\mathbb{Z}^\nu \\ :\ \ell_1\in\mathbb{Z}_{N_1},\ldots,\ell_r\in\mathbb{Z}_{N_r}\Big\}\subseteq\mathbb{T}^\nu\,.$$

Let $\delta=(\prod_i N_i)^{-1/\nu}$ and let V be the corresponding hypercube in $\mathbb{T}^\nu$ defined as in Lemma 7. We want to show that there exists $\bar{\ell}=(\ell_1,\ldots,\ell_r)\in\mathbb{Z}_{N_1}\times\cdots\times\mathbb{Z}_{N_r}$, $\bar{\ell}\neq 0$ such that

$$\left(\sum_{i=1}^r\frac{k_i^1\ell_i}{N_i},\ldots,\sum_{i=1}^r\frac{k_i^\nu\ell_i}{N_i}\right)+\mathbb{Z}^\nu\in V\,.$$

We consider two cases.

(1) If there exists $\bar{\ell}=(\ell_1,\ldots,\ell_r)\in\mathbb{Z}_{N_1}\times\cdots\times\mathbb{Z}_{N_r}$, $\bar{\ell}\neq 0$ such that

$$\left(\sum_{i=1}^r\frac{k_i^1\ell_i}{N_i},\ldots,\sum_{i=1}^r\frac{k_i^\nu\ell_i}{N_i}\right)+\mathbb{Z}^\nu=0\in V \qquad (10)$$

then clearly we can conclude.

(2) Assume now there are no $\bar{\ell}=(\ell_1,\ldots,\ell_r)\in\mathbb{Z}_{N_1}\times\cdots\times\mathbb{Z}_{N_r}$, $\bar{\ell}\neq 0$ satisfying condition (10). In this case it can be shown that two different $\bar{\ell}',\bar{\ell}''\in\mathbb{Z}_{N_1}\times\cdots\times\mathbb{Z}_{N_r}$ yield

$$\left(\sum_{i=1}^r\frac{k_i^1\ell_i'}{N_i},\ldots,\sum_{i=1}^r\frac{k_i^\nu\ell_i'}{N_i}\right)+\mathbb{Z}^\nu\neq \\ \left(\sum_{i=1}^r\frac{k_i^1\ell_i''}{N_i},\ldots,\sum_{i=1}^r\frac{k_i^\nu\ell_i''}{N_i}\right)+\mathbb{Z}^\nu\,,$$

namely different elements in $\mathbb{Z}_{N_1}\times\cdots\times\mathbb{Z}_{N_r}$ always lead do distinct elements in Λ. This implies that $|\Lambda|=\prod_i N_i=\delta^{-\nu}$ and so we are in a position to apply Lemma 7 and conclude that there exist two different $\bar{\ell}',\bar{\ell}''\in\mathbb{Z}_{N_1}\times\cdots\times\mathbb{Z}_{N_r}$ such that

$$\left[\left(\sum_{i=1}^r\frac{k_i^1\ell_i'}{N_i},\ldots,\sum_{i=1}^r\frac{k_i^\nu\ell_i'}{N_i}\right)+\mathbb{Z}^\nu\right]- \\ \left[\left(\sum_{i=1}^r\frac{k_i^1\ell_i''}{N_i},\ldots,\sum_{i=1}^r\frac{k_i^\nu\ell_i''}{N_i}\right)+\mathbb{Z}^\nu\right]\in V$$

and hence

$$\left(\sum_{i=1}^r\frac{k_i^1\ell_i}{N_i},\ldots,\sum_{i=1}^r\frac{k_i^\nu\ell_i}{N_i}\right)+\mathbb{Z}^\nu\in V\,,$$

where $\bar{\ell}=\bar{\ell}'-\bar{\ell}''\neq 0$.

Consider now the eigenvalue

$$\begin{aligned}\lambda&=\sum_{k_1=0}^{N_1-1}\cdots\sum_{k_r=0}^{N_r-1}\pi(k_1,\ldots,k_r)e^{i(\frac{2\pi}{N_1}k_1\ell_1+\cdots+\frac{2\pi}{N_r}k_r\ell_r)}\\ &=\pi(0)+\sum_{j=1}^{\nu}\pi(k_1^j,\ldots,k_r^j)e^{i(\frac{2\pi}{N_1}k_1^j\ell_1+\cdots+\frac{2\pi}{N_r}k_r^j\ell_r)}\,.\end{aligned}$$

Its norm can be estimated as follows

$$\begin{aligned}|\lambda|&\geq\pi(0)+\sum_{j=1}^{\nu}\pi(k_1^j,\ldots,k_r^j)\cos\Big(\frac{2\pi}{N_1}k_1^j\ell_1+\ldots\\ &\quad\cdots+\frac{2\pi}{N_r}k_r^j\ell_r\Big)\\ &\geq\pi(0)+\sum_{j=1}^{\nu}\pi(k_1^j,\ldots,k_r^j)\;-\sum_{j=1}^{\nu}\pi(k_1^j,\ldots,k_r^j)\frac{2\pi^2}{N^{2/\nu}}\\ &\geq 1-2\pi^2\frac{1}{N^{2/\nu}}\end{aligned}$$

and so we can conclude. □

Theorem 6 in particular implies that, if we consider a sequence of Abelian Cayley graphs $\mathcal{G}(G_N,S_N)$ such that

$|G_N| = N$ and $|S_N|$ grows less then logarithmically in N and we consider a sequence of Cayley stochastic matrices P_N compatible with $\mathcal{G}(G_N, S_N)$, then, necessarily, $\rho(P_N)$ converges to 1. This had already been shown, for adjacency matrices, in [2]. Notice that in Example 5 we have that $\nu = 2$ and we have an asymptotic behavior $\rho_{\mathcal{G}}^{\text{Cayley}} \simeq 1 - 2\pi^2 N^{-2}$, while the lower bound of Theorem 6 is, in this case, $1 - 2\pi^2 N^{-1}$. We can wonder whether it is possible to achieve the bound performance. In other words, we would like to understand whether the lower bound we have just found is tight or not. In the following example we show that this is the case.

Example 8 *Consider the group $\mathbb{Z}_N$ where we suppose that $N = M^\nu$. Consider the Cayley graph $\mathcal{G}(\mathbb{Z}_N, S)$, where $S = \{0, 1, M, M^2, \ldots, M^{\nu-1}\}$ and assume that the probability distribution π on S is described by $\pi(0) = \pi(1) = \pi(M) = \ldots = \pi(M^{\nu-1}) = \frac{1}{\nu+1}$. The Fourier transform of π is in this case given by*

$$\hat{\pi}(\chi_\ell) = \sum_{g \in S} \chi(-g)\pi(g) = \frac{1}{\nu+1}\left(1 + \sum_{h=0}^{\nu-1} e^{i\frac{2\pi}{N}M^h \ell}\right)$$

with $\ell = 1, \ldots, N-1$. We will show that, for all $\ell = 1, \ldots, N-1$ we have that

$$|\hat{\pi}(\chi_\ell)| \leq 1 - \frac{1}{\nu+1}\frac{1}{M^2} \tag{11}$$

This fact will be shown by induction on ν. The fact that the assertion holds for $\nu = 1$ follows from Example 4. Assume now that the assertion holds for $\nu - 1$. Let ℓ_0, ℓ_1 such that $0 \leq \ell_0 \leq M-1$, $0 \leq \ell_1 \leq M^{\nu-1} - 1$ and $\ell = \ell_0 + M\ell_1$. If $\ell_0 \neq 0$ then

$$\begin{aligned}
|\hat{\pi}(\chi_\ell)| &\leq \frac{1}{\nu+1}\left|1 + e^{i\frac{2\pi}{M^\nu}M^{\nu-1}\ell}\right| + \frac{1}{\nu+1}\left|\sum_{h=0}^{\nu-2} e^{i\frac{2\pi}{M^\nu}M^h \ell}\right| \\
&\leq \frac{1}{\nu+1}\left|1 + e^{j\frac{2\pi}{M}\ell_0}\right| + \frac{\nu-1}{\nu+1}
\end{aligned}$$

Since (11) holds for $\nu = 1$ we have that

$$\frac{1}{2}\left|1 + e^{j\frac{2\pi}{M}\ell_0}\right| \leq 1 - \frac{1}{2}\frac{1}{M^2}$$

and hence

$$\begin{aligned}
|\hat{\pi}(\chi_\ell)| &\leq \frac{2}{\nu+1}\left(1 - \frac{1}{2}\frac{1}{M^2}\right) + \frac{\nu-1}{\nu+1} \\
&\leq 1 - \frac{1}{\nu+1}\frac{1}{M^2}
\end{aligned}$$

If $\ell_0 = 0$, then $\ell = M\ell_1$ and so

$$\begin{aligned}
|\hat{\pi}(\chi_\ell)| &= \frac{1}{\nu+1}\left|1 + \sum_{h=0}^{\nu-1} e^{i\frac{2\pi}{M^{\nu-1}}M^h \ell_1}\right| \\
&= \frac{1}{\nu+1}\left|2 + \sum_{h=0}^{\nu-2} e^{i\frac{2\pi}{M^{\nu-1}}M^h \ell_1}\right| \\
&\leq \frac{\nu}{\nu+1}\left|\frac{1}{\nu}\left(1 + \sum_{h=0}^{\nu-2} e^{i\frac{2\pi}{M^{\nu-1}}M^h \ell_1}\right)\right| + \frac{1}{\nu+1}
\end{aligned}$$

From the inductive hypothesis it follows that

$$\left|\frac{1}{\nu}\left(1 + \sum_{h=0}^{\nu-2} e^{i\frac{2\pi}{M^{\nu-1}}M^h \ell_1}\right)\right| \leq 1 - \frac{1}{\nu}\frac{1}{M^2}.$$

Hence

$$\begin{aligned}
|\hat{\pi}(\chi_\ell)| &\leq \frac{\nu}{\nu+1}\left(1 - \frac{1}{\nu}\frac{1}{M^2}\right) + \frac{1}{\nu+1} \\
&\leq 1 - \frac{1}{\nu+1}\frac{1}{M^2}
\end{aligned}$$

This bound proves that there exists a circulant graph $\mathcal{G}$ with ν incoming edges in any vertex such that

$$\rho_{\mathcal{G}}^{\text{Cayley}} \leq 1 - \frac{1}{\nu+1}\frac{1}{N^{2/\nu}}.$$

proving in this way that the bound proposed by the previous theorem is tight.

The question at this point is the following: Is the Cayley structure on the matrix or the Cayley structure on the graph that prevents to obtain good performance? In other words, do there exist stochastic matrices supported by Abelian Cayley graphs that exhibit better performance than what imposed by the bound (9)? Notice that, in order to make fair comparisons, we need to limit to doubly stochastic matrices. We conjecture that for doubly stochastic matrices supported on Abelian Cayley graphs the bound (9) continues to hold. What about other graphs? An easy way to restrict to doubly stochastic matrices is by imposing that they are symmetric and so that the corresponding graphs are undirected. If A is the adjacency matrix of a ν-regular undirected graph, then, $P = \nu^{-1}A$ is doubly stochastic. For these graphs, we recall a basic asymptotic lower bound by Alon and Boppana [1] on the second eigenvalue

$$\liminf_{N \to +\infty} \rho(P) \geq \frac{2\sqrt{\nu-1}}{\nu},$$

where the lim inf is intended to be performed along the family of all ν-regular undirected graphs having N vertices. Ramanujan graphs (see [30] and references therein) are those ν-regular undirected graphs achieving the previous bound, namely such that $\rho(P) = 2\nu^{-1}\sqrt{\nu-1}$. Hence, through these graphs, it would be possible to keep the essential spectral radius bounded away from 1, while keeping the degree fixed (see also [34]). In fact, there are plenty of Ramanujan graphs (for instance any complete graph), but it is still an open problem if for any N and ν there exists a Ramanujan graph with N vertices and degree ν. There are only partial results in this direction. For example it is possible to prove that, if ν is such that $\nu-1$ is the power of a prime, then there exist a sequence of Ramanujan graphs with a growing number of vertices and of fixed degree ν. Moreover, when available, these constructions are quite complicated and the fact that they strictly depend on the choice of particular number of vertices makes them not so interesting from our point of view. However, it is interesting to notice that graphs behaving similarly to the Ramanujan ones are not so unlikely. Indeed Friedman [19] showed that for ν sufficiently large and fixed, in the average, $\rho(P)$ with $P = \nu^{-1}A$, remains bounded away from 1 as $N \to +\infty$.

5 Time-varying Strategies

In the previous sections we showed that controllers with symmetries behave quite poorly. One possibility to achieve better performance is to resort on Ramanujan graphs or to undirected regular graphs generated randomly. An alternative way to increase performance, while maintaining the symmetry of the controllers, is by a time-varying strategy in which at every time instant the communication graph is chosen randomly in a set of Cayley graphs. Such strategies yield a mean square convergence rate that is higher and, more importantly, independent of the number of systems.

5.1 Time-varying Cayley Graphs

Fix an Abelian group G and a number $\nu < |G|$. We consider a sequence of subsets $S_t \subseteq G$ that are randomly generated in the following way.
Let $\alpha_i(t)$, $i = 1, \dots, \nu$, be ν independent sequences of independent random variables taking value on G and uniformly distributed in such a set. We put

$$S_t = \{\alpha_0(t) = 0, \alpha_1(t), \dots, \alpha_\nu(t)\}\,.$$

Notice that in S_t there might be repetitions and so its cardinality may be less than $\nu + 1$.
Fix $k_0, k_1, \dots, k_\nu \geq 0$ such that $\sum_j k_j = 1$ consider the sequence of probability distributions π_t on G supported on the sequence of sets S_t defined as

$$\pi_t(g) = \begin{cases} k_j & \text{if } g = \alpha_j(t) \\ 0 & \text{otherwise.} \end{cases}$$

Let P_t be the stochastic Cayley matrix associated with π_t. If we consider the feedback matrix $K_t := I - P_t$, we obtain the closed loop system becomes $x(t+1) = P_t x(t)$, which is an instance of jump Markov linear system [16,5]. The state $x(t)$ becomes a random variable that evolves according to

$$x(t) = \prod_{s=1}^{t} P_s x(0)\,, \tag{12}$$

where $x(0)$ is a random variable independent of the processes $\alpha_i(t)$.
We want to study the asymptotic behavior of $x(t)$. Since we are interested in achieving average consensus, we consider the displacement from the average $\Delta(t) := x(t) - N^{-1}\mathbb{1}\mathbb{1}^T x(0) = (I - N^{-1}\chi_0\chi_0^*)x(t)$, which is governed by

$$\Delta(t) = \prod_{s=1}^{t} P_s \Delta(0)\,,$$

where $\Delta(0)$ is now a random variable taking values on $\mathbb{R}^G$ such that $< \Delta(0), \chi_0 >= 0$ and independent of the set of variables $\{\alpha_i(t)\}$. In this probabilistic context it is natural to study the asymptotic behavior of $\mathbb{E}||\Delta(t)||^2$. This is the result we obtain:

Proposition 9

$$\mathbb{E}||\Delta(t)||^2 = \left(\sum_{j=0}^{\nu} k_j^2\right)^t \mathbb{E}||\Delta(0)||^2\,.$$

PROOF. We know we can represent

$$P_t = \sum_{\chi \in \hat{G}} \hat{\pi}_t(\chi) N^{-1} \chi\chi^*\,.$$

Hence,

$$\prod_{s=1}^{t} P_s = \sum_{\chi \in \hat{G}} \left[\prod_{s=1}^{t} \hat{\pi}_s(\chi)\right] N^{-1}\chi\chi^*\,.$$

Let us study the average of the squared norm of the various eigenvalues.

$$\mathbb{E}\left[\left|\prod_{s=1}^{t} \hat{\pi}_s(\chi)\right|^2\right] = \prod_{s=1}^{t} \mathbb{E}\left[|\hat{\pi}_s(\chi)|^2\right]\,.$$

Since

$$\hat{\pi}_t(\chi) = k_0 + \sum_{j=1}^{\nu} k_j \chi(-\alpha_j(t))\,,$$

we obtain

$$\mathbb{E}\left[|\hat{\pi}_t(\chi)|^2\right] = k_0^2 + \sum_{j=1}^{\nu} k_0 k_j \left[\mathbb{E}\left[\chi(\alpha_j(t))\right] + \mathbb{E}\left[\chi(\alpha_j(t))^*\right]\right] + \sum_{j=1}^{\nu}\sum_{\ell=1}^{\nu} k_j k_\ell \mathbb{E}\left[\chi(\alpha_j(t))\chi(\alpha_\ell(t))^*\right] . \tag{13}$$

It is immediate to verify that $\mathbb{E}[\chi(\alpha_j(t))] = 0$ when $\chi \neq \chi_0$, $\mathbb{E}[\chi(\alpha_j(t))\chi(\alpha_\ell(t))^*] = 0$ when $j \neq \ell$ and $\mathbb{E}[|\chi(\alpha_j(t))|^2] = 1$. Substituting in (13) we then obtain

$$\mathbb{E}\left[|\hat{\pi}_t(\chi)|^2\right] = k_0^2 + \sum_{j=1}^{\nu} k_j^2 = \sum_{j=0}^{\nu} k_j^2 , \quad \forall \chi \neq 0 .$$

We finally have

$$\begin{aligned}\mathbb{E}||\Delta(t)||^2 &= \sum_{\chi\neq\chi_0} \mathbb{E}\left[\left|\prod_{s=1}^{t} \hat{\pi}(\chi)\right|^2\right] \frac{1}{N}\mathbb{E}| < \Delta(0), \chi > |^2 \\ &= \left(\sum_{j=0}^{\nu} k_j^2\right)^t \mathbb{E}||\Delta(0)||^2 .\end{aligned}$$

□

Notice that

$$\min\left\{\sum_{j=0}^{\nu} k_j^2 \;\middle|\; k_j \geq 0 , \sum_{j=1}^{\nu} k_j = 1\right\} = \frac{1}{\nu+1}$$

and it is obtained by choosing $k_j = 1/(\nu+1)$ for all j. With such a choice we have thus obtained the following mean convergence result

$$\mathbb{E}||\Delta(t)||^2 = \left(\frac{1}{1+\nu}\right)^t \mathbb{E}||\Delta(0)||^2 .$$

This performance is much better than what we had obtained so far, since in this case the rate of convergence remains constant with respect to N.

Remark 10 *As any average result, it is not immediately evident how the average computation above reflects on the behavior of the system when we consider a generic sequence S_t of subsets chosen at random. A simple standard probabilistic argument however allows us to show that such a convergence rate is indeed achieved by almost every sequence S_t. Fix any $c > 1$ and notice that for every $\chi \in \hat{G}$,*

$$\begin{aligned}P\left(\prod_{s=1}^{t} |\hat{\pi}_s(\chi)| \geq \left(\frac{c}{\nu+1}\right)^t\right) &= \\ = P\left(\sum_{s=1}^{t} \ln|\hat{\pi}_s(\chi)| \geq t\ln\frac{c}{\nu+1}\right) . \end{aligned} \tag{14}$$

Notice that $Y_s = \ln|\hat{\pi}_s(\chi)|$ is a sequence of independent identically distributed random variables taking values on $[-\infty, 0]$. The idea is to apply Chebyschev inequality. To overcome the problem of possible unboundedness of Y_s, we consider two different cases. Suppose that there exist a $\chi \in \hat{G}$ such that $\hat{\pi}_s(\chi)$ assumes the value 0 with probability $p > 0$. In this case we can simply estimate

$$P\left(\prod_{s=1}^{t} |\hat{\pi}_s(\chi)| \geq \left(\frac{c}{\nu+1}\right)^t\right) \leq (1-p)^t . \tag{15}$$

If instead the event $\{\hat{\pi}_s(\chi) = 0\}$ has probability zero, then the random variable Y_s is bounded and can be estimated as follows. First notice that, using Jensen inequality, we have

$$\mathbb{E}[Y_s] = \mathbb{E}[\ln|\hat{\pi}_s(\chi)|] \leq \ln\mathbb{E}[|\hat{\pi}_s(\chi)|] = \ln\frac{1}{\nu+1} .$$

Let

$$\delta := \ln\frac{c}{\nu+1} - \mathbb{E}[Y_s] = \ln c + \ln\frac{1}{\nu+1} - \mathbb{E}[Y_s] \geq \ln c > 0 .$$

We can now estimate

$$\begin{aligned}P\left(\sum_{s=1}^{t} \ln|\hat{\pi}_s(\chi)| \geq t\ln\frac{c}{\nu+1}\right) &= P\left(\sum_{s=1}^{t} Y_s \geq t\ln\frac{c}{\nu+1}\right) \\ = P\left(\sum_{s=1}^{t}(Y_s - \mathbb{E}[Y_s]) \geq t\delta\right) &\leq \frac{\mathrm{Var}[Y_s]}{\delta^2 t^2}\end{aligned} \tag{16}$$

A straightforward application of Borel-Cantelli lemma now allows to conclude from relations (15) and (16) that, for almost every sequence S_t of subsets,

$$\prod_{s=1}^{t} |\hat{\pi}_s(\chi)| < \left(\frac{c}{\nu+1}\right)^t$$

for t sufficiently large and for every $\chi \in \hat{G}$. From this we also obtain that, for almost every sequence S_t,

$$||\Delta(t)||^2 \leq \left(\frac{c}{\nu+1}\right)^t ||\Delta(0)||^2 \quad \text{for } t \text{ sufficiently large} .$$

Considering that c can be chosen arbitrarily close to 1, this proves our claim.

From an implementation point of view this strategy has an evident drawback: the same random choice done at every time instance has to be done by all systems. This seems to require a supervised communication of this information to every system. A possible way to overcome this limitation is by imposing that each agent uses the same pseudorandom number generator starting from the same seed.

5.2 Time-varying with Bounded In-degree

In this section we consider a time-varying strategy similar to the one presented in the previous section. The difference is that here we do not limit the time-varying matrices to be Cayley. We see that this generalization does not lead to better performance. In this case we assume that each system receives the state of ν systems chosen randomly and independently. Because of this it can happen that the resulting communication graph has multiple arcs connecting the same pair of nodes.
Fix $k_0, k_1, \ldots, k_\nu \geq 0$ such that $\sum_j k_j = 1$. The feedback matrix is in this case

$$K_t = (k_0 - 1)I + \sum_{i=1}^{\nu} k_i E_i(t)$$

where $E_i(t)$, $i = 1, \ldots, \nu$, are ν independent sequences of independent random variables taking values on the set of matrices

$$\mathcal{E} := \{E \in \{0,1\}^{N\times N} \ : \ E\mathbb{1} = \mathbb{1}\}$$

and uniformly distributed in such a set. The set $\mathcal{E}$ is constituted by all matrices with entries 0 or 1 which have exactly one 1 in each row. The closed loop system becomes $x(t+1) = P_t x(t)$ where

$$P_t = k_0 I + \sum_{i=1}^{\nu} k_i E_i(t)\,. \tag{17}$$

As before, the state $x(t)$ is a random variable which evolves according to (12). The initial condition $x(0)$ is a random variable independent of the processes $E_i(t)$.
Again, we want to study the asymptotic behavior of $x(t)$. Since the controllers we are using are not necessarily average controllers, we can not longer use the variable $\Delta(t) := x(t) - N^{-1}\mathbb{1}\mathbb{1}^T x(0)$ to study convergence to the consensus point. However we can prove the following result.

Theorem 11 *There exists a scalar random variable α^* such that*

$$\mathbb{E}||x(t) - \alpha^*\mathbb{1}||^2 \leq C\rho^t \mathbb{E}||(I - N^{-1}\mathbb{1}\mathbb{1}^T)x(0)||^2 \tag{18}$$

where

$$\rho = k_0^2 + \frac{N-1}{N}\sum_{i=1}^{\nu} k_i^2\,, \quad C = \frac{1 - 2k_0 + \sum_{i=1}^{\nu} k_i^2}{(1-\rho^{1/2})^2}$$

PROOF. Let $Q(t) := \mathbb{E}[x(t)x(t)^T]$. Notice that

$$\begin{aligned} Q^+ &= \mathbb{E}[P_t x x^T P_t^T] = \mathbb{E}[\mathbb{E}[P_t x x^T P_t^T | P_t]] = \mathbb{E}[P_t Q P_t^T] \\ &= k_0^2 Q + \sum_{i=1}^{\nu} k_0 k_i (Q\mathbb{E}[E_i^T] + \mathbb{E}[E_i]Q) \\ &+ \sum_{\substack{i,j=1 \\ i\neq j}}^{\nu} k_i k_j \mathbb{E}[E_i]Q\mathbb{E}[E_j^T] + \sum_{i=1}^{\nu} k_i^2 \mathbb{E}[E_i Q E_i^T] \end{aligned}$$

Notice that $\mathbb{E}[E_i] = N^{-1}\mathbb{1}\mathbb{1}^T$. Moreover, for any $M \in \mathbb{R}^{N\times N}$ it holds

$$\mathbb{E}[E_i M E_i^T] = \frac{1}{N}\mathrm{tr}\,\{M\}I + \frac{1}{N^2}\mathbb{1}^T M\mathbb{1}(\mathbb{1}\mathbb{1}^T - I)\,.$$

These relations imply that

$$\begin{aligned} Q^+ &= k_0^2 Q + \sum_{i=1}^{\nu} k_0 k_i (N^{-1}\mathbb{1}\mathbb{1}^T Q + Q N^{-1}\mathbb{1}\mathbb{1}^T) \\ &\quad + \sum_{i=1}^{\nu} k_i^2 \left(N^{-1}\mathrm{tr}\,(Q)I + N^{-2}\mathbb{1}^T Q\mathbb{1}(\mathbb{1}\mathbb{1}^T - I)\right) \\ &\quad + \sum_{\substack{i,j=1 \\ i\neq j}}^{\nu} k_i k_j \, N^{-1}\mathbb{1}\mathbb{1}^T Q N^{-1}\mathbb{1}\mathbb{1}^T\,. \end{aligned}$$

Let us define $w(t) = \mathrm{tr}\,(Q(t)) = \mathbb{E}||x(t)||^2$ and $s(t) = N^{-1}\mathbb{1}^T Q(t)\mathbb{1}$. Notice that

$$\begin{aligned} w^+ =& k_0^2 w + 2\left(\sum_{i=1}^{\nu} k_0 k_i\right) s + \left(\sum_{i=1}^{\nu} k_i^2\right) w \\ &+ \left(\sum_{\substack{i,j=1 \\ i\neq j}}^{\nu} k_i k_j\right) s = \left(\sum_{i=0}^{\nu} k_i^2\right) w + \left(1 - \sum_{i=0}^{\nu} k_i^2\right) s\,. \end{aligned}$$

Moreover we have that

$$\begin{aligned} s^+ &= k_0^2 s + 2\left(\sum_{i=1}^{\nu} k_0 k_i\right) s + \left(\sum_{i=1}^{\nu} k_i^2\right)\left(\frac{1}{N}w + \frac{N-1}{N}s\right) \\ &+ \left(\sum_{\substack{i,j=1 \\ i\neq j}}^{\nu} k_i k_j\right) s = \left(\frac{1}{N}\sum_{i=1}^{\nu} k_i^2\right) w + \left(1 - \frac{1}{N}\sum_{i=1}^{\nu} k_i^2\right) s\,. \end{aligned}$$

The previous two relations can be summarized as follows

$$\begin{bmatrix} w^+ \\ s^+ \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{\nu} k_i^2 & 1 - \sum_{i=0}^{\nu} k_i^2 \\ \frac{1}{N} \sum_{i=1}^{\nu} k_i^2 & 1 - \frac{1}{N} \sum_{i=1}^{\nu} k_i^2 \end{bmatrix} \begin{bmatrix} w \\ s \end{bmatrix}. \quad (19)$$

We want now to estimate $\mathbb{E}||x(t+1) - x(t)||^2$. Notice that

$$\begin{aligned}
&\mathbb{E}||x(t+1) - x(t)||^2 = \\
&= \operatorname{tr} \mathbb{E}(x(t+1) - x(t))(x(t+1) - x(t))^T = \\
&= \operatorname{tr} Q(t+1) + \operatorname{tr} Q(t) - 2\operatorname{tr}\,[(k_0 I + (1-k_0)\frac{1}{N}\mathbb{1}\mathbb{1}^T)Q] = \\
&= \left(\sum_{i=0}^{\nu} k_i^2\right) w + \left(1 - \sum_{i=0}^{\nu} k_i^2\right) s + w(t) - 2k_0 w \\
&\quad - 2(1-k_0)s(t) = \left(1 - 2k_0 + \sum_{i=0}^{\nu} k_i^2\right)(w(t) - s(t)).
\end{aligned}$$

Notice that, from equation (19) we can argue that

$$w(t) - s(t) = \left(k_0^2 + \frac{N-1}{N}\sum_{i=1}^{\nu} k_i^2\right)^t (w(0) - s(0))$$

and so

$$\mathbb{E}||x(t+1) - x(t)||^2 = \left(1 - 2k_0 + \sum_{i=0}^{\nu} k_i^2\right)\rho^t (w(0) - s(0)) \quad (20)$$

where

$$\rho := k_0^2 + \frac{N-1}{N}\sum_{i=1}^{\nu} k_i^2 .$$

Standard arguments on complete metrics show that the exponential convergence of the previous sequence implies that $x(t)$ must converge to some random vector x^* in the L^2-norm $(\mathbb{E}||y(t)||^2)^{1/2}$. Moreover,

$$\begin{aligned}
&(\mathbb{E}||x(t) - x^*||^2)^{1/2} \leq \sum_{s=t}^{+\infty} (\mathbb{E}||x(s+1) - x(s)||^2)^{1/2} \\
&= \left(1 - 2k_0 + \sum_{i=0}^{\nu} k_i^2\right)^{1/2} (w(0) - s(0))^{1/2} \sum_{s=t}^{+\infty} \rho^{s/2} = \\
&= \left(\frac{1 - 2k_0 + \sum_{i=0}^{\nu} k_i^2}{(1-\rho^{1/2})^2}(w(0) - s(0))\right)^{1/2} \rho^{t/2}.
\end{aligned}$$

Notice finally that, if $Y := I - N^{-1}\mathbb{1}\mathbb{1}^T$, then $\mathbb{E}||Yx(t)||^2 = w(t) - s(t)$ and, since $w(t) - s(t)$ tends to zero, we can argue $Yx^* = 0$ and this implies that there exists a scalar random variable α^* such that $x^* = \alpha^*\mathbb{1}$. $\square$

Notice that the ρ appearing in estimation (18) is the exact exponential rate of convergence of $\mathbb{E}||x(t) - \alpha^*\mathbb{1}||^2$ in the sense that

$$\lim_{t\to+\infty} \frac{\log \mathbb{E}||x(t) - \alpha^*\mathbb{1}||^2}{t} = \log\rho .$$

This is a straightforward consequence of relation (20). Notice moreover that the strongest exponential rate of convergence in (18) is given by

$$\begin{aligned}
\min\Big\{k_0^2 + \frac{N-1}{N}\sum_{i=1}^{\nu} k_i^2 \mid k_0, k_1, \ldots, k_\nu \geq 0, \\
k_0 + \sum_{i=1}^{\nu} k_i = 0\Big\} = \frac{N-1}{N(\nu+1)-1},
\end{aligned}$$

obtained by choosing

$$k_0 = \frac{N-1}{N(\nu+1)-1}, \quad \text{and} \quad k_i = \frac{N}{N(\nu+1)-1} \quad (21)$$

$i = 1, \ldots, \nu$. Notice that this convergence rate is smaller than $1/(\nu+1)$, which is the rate obtained through the time-varying strategy on Cayley graphs discussed before. However, for $N \to +\infty$, the two strategies yield the same rate. From the average behavior we can prove a performance result on generic random samples as we did in previous case. The probabilistic tools needed become however a bit more refined: matrices are not simultaneously diagonalizable and we have to use Oseledec ergodic theorem for products of random matrices. This will be presented elsewhere.

The most important difference between the two random strategies presented here is that the time-varying strategy on Cayley graphs yields convergence to the average of the initial configuration, whereas the one presented in this section does not reach the consensus at the initial average. Therefore, it is interesting to study how far from the initial average the systems reach consensus. We have the following exact result:

Proposition 12 *Let α^* be the random variable defined in Theorem 11. Then*

$$\mathbb{E}|\alpha^* - N^{-1}\mathbb{1}^T x(0)|^2 = \frac{\sum_{i=1}^{\nu} k_i^2 \mathbb{E}||\left(I - N^{-1}\mathbb{1}\mathbb{1}^T\right)x(0)||^2}{N^2[N(1-k_0^2) + (1-N)\sum_{i=1}^{\nu} k_i^2]},$$

PROOF. Consider $\Delta(t) := x(t) - N^{-1}\mathbb{1}\mathbb{1}^T x(0)$. We know from (5) that the dynamics of $\Delta(t)$ is described by the equation $\Delta^+ = P_t\Delta$ where P_t is given in (17). For this reason, by denoting $Q(t) := \mathbb{E}[\Delta(t)\Delta(t)^T]$, $w(t) = \operatorname{tr}(Q(t)) = \mathbb{E}||\Delta(t)||^2$ and $s(t) = \mathbb{1}^T Q(t)\mathbb{1}/N$, exactly the same computation done in the proof of the previous result show that equation (19) still holds true. The transition matrix has eigenvalues $\lambda_1 = 1$, and $\lambda_2 =$

$k_0^2 + \frac{N-1}{N}\sum_{i=1}^{\nu} k_i^2$. The second eigenvalue coincides with the convergence rate ρ computed before. The time evolution of $w(t)$ and $s(t)$ is thus given by

$$[w(t), s(t)]^T = c_1\lambda_1^t a_1 + c_2\lambda_2^t a_2$$

where c_1, c_2 are constants and a_1, a_2 are the eigenvectors associated to λ_1 and λ_2. Notice that $a_1 = (1\ 1)^T$. At steady state the vector $(w(\infty), s(\infty))^T$ is aligned to the dominant eigenvector a_1 and thus $w(\infty) = c_1$. Simple calculations yield

$$w(\infty) = \frac{\sum_{i=1}^{\nu} k_i^2\, \mathbb{E}||\left(I - N^{-1}\mathbb{1}\mathbb{1}^T\right)x(0)||^2}{N[N(1-k_0^2) + (1-N)\sum_{i=1}^{\nu} k_i^2]}\,,$$

This yields the result. $\square$

If we use the control gains $k_0, k_1, \ldots, k_\nu$ as in (21), which yield the fastest convergence rate, then we have

$$\mathbb{E}|\alpha^* - N^{-1}\mathbb{1}^T x(0)|^2 = \frac{\mathbb{E}||\left(I - N^{-1}\mathbb{1}\mathbb{1}^T\right)x(0)||^2}{N^2(N(1+\nu)-1)}\,.$$

Notice that, if the initial states $x_i(0)$ of the systems are independent and $\mathbb{E}\left(x_i(0)^2\right) = \sigma^2$ is the same for all i, then,

$$\mathbb{E}||\left(I - N^{-1}\mathbb{1}\mathbb{1}^T\right)x(0)||^2 = (N-1)\sigma^2\,.$$

In this case the final formula becomes

$$\mathbb{E}|\alpha^* - N^{-1}\mathbb{1}^T x(0)|^2 = \frac{N-1}{N^2(N(1+\nu)-1)}\sigma^2\,,$$

which in particular shows that, as $N \to \infty$, the mean square distance of the consensus to the initial average tends to zero as N^{-2}.

6 Logarithmic Quantizers

In this section we present another strategy that allows us to overcome the poor performance that are achievable by time-invariant communication networks with symmetries. This can be done by allowing data exchange over communication links that transmit logarithmic quantized data. As well-known in the literature, logarithmic quantizers provide a very efficient way of transmitting control signals. More precisely, assume we want to drive the state from a state region I to a target region J and let C, called the *contraction rate*, be the ratio between the measure of I and the measure of J. This parameter describes the required relative precision of the consensus. It is known that [14,15], while exact communication links, modelled by uniform quantizers, require channels able to transmit over an alphabet with cardinality proportional to C, logarithmic quantizers need instead an alphabet with cardinality growing only logarithmically in C. The simplest way to model for the effect of a logarithmic quantizer is by introducing a multiplicative noise. In this section we provide the instruments for analyzing what happens if we introduce this kind of links in the consensus problem.

Assume we have fixed an Abelian group G having N elements and a subset $S \subseteq G$ such that $0 \in S$. Consider the Cayley graph $\mathcal{G}$ associated with G and S. This has to be interpreted as the un-noisy communication graph with which we associate a Cayley stochastic matrix P_0 compatible with $\mathcal{G}$. Such a matrix corresponds to the closed loop matrix obtained using these perfect communication links. We now consider the possibility that each system i can transmit functions of the exact information available at system i to some other systems. Such transmissions are logarithmically quantized and this effect is approximated by introducing a multiplicative noise. We impose that the Cayley symmetry of the overall structure is maintained. In order to achieve this, we define q outputs

$$z_s := H_s x, \quad s = 1, \ldots, q \tag{22}$$

where H_s are Cayley matrices still compatible with $\mathcal{G}$. The i-th components of the outputs $z_1, \ldots, z_q$ represent the information the i-th system transmits to the other systems. In this way every system transmits q scalar messages. We assume that each component of the output z_{si} gets distorted by the multiplicative noise $1 + e_{s,i}$. To complete the model we have to specify which systems receive this information and how this information is used for the control. We assume again Cayley structure at the level of controllers, namely we assume there exist Cayley matrices P_s such that the closed loop dynamics can be described as

$$x^+ = P_0 x + \sum_{s=1}^{q} P_s(I + E_s)H_s x\,,$$

where $E_s = \mathrm{diag}\{e_{s,1}, \ldots, e_{s,N}\}$ is a diagonal matrix of noise random variables. All noises $e_{s,i}$ are assumed to be independent, having zero mean and finite variance δ_s^2. Notice that the nonzero elements of the matrix P_s specify what logarithmic link is active. More specifically, $(P_s)_{ij} \neq 0$ means that the signal $(H_s x)_j$ is transmitted to the system i after being logarithmically quantized.

It is reasonable to assume that consensus configurations $x = c\chi_0$ are equilibrium points, namely $x^+ = x$ under any possible multiplicative noise. This happens if and only if $P_0\mathbb{1} = \mathbb{1}$, and $H_s\mathbb{1} = 0$ for $s = 1, \ldots, q$. This is quite natural: data affected by multiplicative noise maintain the consensus convergence only if they converge to 0. Hence they must consist in differences.

The asymptotic behavior of this dynamical system can be studied in a similar way to the random case treated in Subsection 5.2 by considering $Q = \mathbb{E}[xx^T]$. With the position $P = P_0 + \sum_s P_s H_s$, the evolution law for Q can be described as follows

$$Q^+ = PQP^T + \sum_{s=1}^{q} P_s \mathbb{E}\left(E_s H_s Q H_s^T E_s\right) P_s^T ,$$

Observe that, if M is any square matrix, then

$$\mathbb{E}(E_s M E_s^T)_{ij} = \mathbb{E}(e_{si} M_{ij} e_{sj}) = \begin{cases} M_{ii}\mathbb{E}(e_{si}^2) & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} .$$

This implies that

$$Q^+ = PQP^T + \sum_{s=1}^{q} \delta_s^2 P_s \mathrm{diag}(H_s Q H_s^T) P_s^T , \qquad (23)$$

where we use the notation

$$\mathrm{diag}\{M\} := \mathrm{diag}\{M_{1,1}, \ldots, M_{N,N}\}.$$

Let $Y := I - N^{-1}\mathbb{1}\mathbb{1}^T$ and define the signals $y(t) := Yx(t)$ and $x_B(t) = N^{-1}\mathbb{1}\mathbb{1}^T x(t)$. Let moreover

$$\begin{aligned} w(t) &:= \mathbb{E}[||y(t)||^2] = \mathrm{tr}\, \mathbb{E}[y(t)y(t)^T] = \mathrm{tr}\,(YQ(t)Y^T) \\ w_s(t) &:= \mathbb{E}[||z_s(t)||^2] = \mathrm{tr}\,(H_s Q(t) H_s^T) \qquad (24) \\ s(t) &:= \mathbb{E}[||x_B(t)||^2] = \mathrm{tr}\,(N^{-1}\chi_0\chi_0^* Q(t) N^{-1}\chi_0\chi_0^*) \\ &= N^{-1}\chi_0^* Q(t)\chi_0 . \end{aligned}$$

where the signals $z_s(t)$ are defined in (22).
In order to study the evolution of the above quantities, we need a technical result on the trace operator for Cayley matrices. Assume P is a Cayley matrix. We know that P can be written as

$$P = \sum_{\chi\in\hat{G}} \theta(\chi) N^{-1} \chi\chi^* . \qquad (25)$$

Consider now the norm $||P||^2 := \sum_{\chi\in\hat{G}} |\theta(\chi)|^2$. Notice that, if π id the generator of P, then $\theta(\chi) = \hat{\pi}(\chi)$ for all $\chi \in \hat{G}$, where $\hat{\pi}$ is the Fourier transform of π. Moreover by Parseval theorem we have that $||P||^2 = \sum_{\chi\in\hat{G}} |\hat{\pi}(\chi)|^2 = N\sum_{g\in G} |\pi(g)|^2$. We have the following result.

Lemma 13 *Assume that P is a Cayley matrix and D is a diagonal matrix. Then,*

$$\mathrm{tr}\,(PDP^*) = N^{-1}||P||^2 \mathrm{tr}\,(D) .$$

PROOF. Assume that P is represented as in (25)). We can write

$$\begin{aligned} PDP^* &= \sum_{\chi,\bar{\chi}} \theta(\chi)\theta(\bar{\chi}) N^{-1}\chi\chi^* D N^{-1}\bar{\chi}\bar{\chi}^* \\ &= N^{-1}\sum_{\chi,\bar{\chi}} \theta(\chi)\theta(\bar{\chi})(\chi^* D\bar{\chi}) N^{-1}\chi\bar{\chi}^* . \end{aligned}$$

Hence,

$$\mathrm{tr}\,(PDP^*) = N^{-1}\sum_{\chi,\bar{\chi}} \theta(\chi)\theta(\bar{\chi})(\chi^* D\bar{\chi})\mathrm{tr}\,(N^{-1}\chi\bar{\chi}^*) .$$

It is immediate to verify that

$$\mathrm{tr}\,(N^{-1}\chi\bar{\chi}^*) = \begin{cases} 0 & \text{f } \chi \neq \bar{\chi} \\ 1 & \text{if } \chi = \bar{\chi} \end{cases}$$

Moreover, we have that

$$\chi^* D\chi = \sum_{g\in G} \chi(g)^* D_{gg}\chi(g) = \sum_{g\in G} D_{gg} = \mathrm{tr}\,(D)$$

Substituting in the expression above we finally obtain

$$\mathrm{tr}\,(PDP^*) = N^{-1}\sum_{\chi} |\theta(\chi)|^2 \mathrm{tr}\,(D) = N^{-1}||P||^2 \mathrm{tr}\,(D) .$$

□

Using the above lemma, we obtain from (23) that

$$\begin{aligned} w^+ &= \mathrm{tr}\,(YPQP^TY^T) + N^{-1}\sum_{s=1}^{q} \delta_s^2 ||YP_s||^2 w_s \\ w_r^+ &= \mathrm{tr}\,(H_r PQP^T H_r^T) + N^{-1}\sum_{s=1}^{q} \delta_s^2 ||H_r P_s||^2 w_s \\ s^+ &= s + N^{-1}\sum_{s=1}^{q} \delta_s^2 |\lambda_s|^2 w_s \qquad (26) \end{aligned}$$

where λ_s is defined by $P_s\chi_0 = \lambda_s\chi_0$ (equivalently, $\lambda_s = \hat{\pi}_{P_s}(\chi_0)$). Define $\mathbf{w}(t)$ to be the q-dimensional vector with $w_s(t)$ at position s and moreover the $q \times q$-matrix L with

$$L_{rs} = N^{-1}\delta_s^2 ||H_r P_s||^2$$

We have the following result.

Lemma 14 *Assume that $L \neq 0$. Then we have*

$$\mathbf{w}(t) \leq (\rho^2 I + L)^t \mathbf{w}(0)$$

where the inequality is meant componentwise and where $\rho := \rho(P)$ (namely the essential spectral radius of P as defined in (3)).

PROOF. Since P is a Cayley matrix, it can be written as in (25). Then

$$\mathrm{tr}\,(H_r P Q P^T H_r^T) = \frac{1}{N}\sum_{\chi\neq\chi_0}|\theta(\chi)|^2 \mathrm{tr}\,(H_r Q H_r^T \chi\chi^*)$$
$$\leq \frac{1}{N}\max\{|\theta(\chi)|^2 : \chi\neq\chi_0\}\sum_{\chi\neq\chi_0}\mathrm{tr}\,(H_r Q H_r^T \chi\chi^*)$$
$$= \rho^2 \mathrm{tr}\left(H_r Q H_r^T \frac{1}{N}\sum_{\chi}\chi\chi^*\right) = \rho^2 \mathrm{tr}\left(H_r Q H_r^T\right) = \rho^2 w_r$$

Define now a sequence of q dimensional vectors $\bar{\mathbf{w}}(t)$ as follows. Let $\bar{\mathbf{w}}(0) = \mathbf{w}(0)$ and let

$$\bar{\mathbf{w}}^+ = (\rho^2 I + L)\bar{\mathbf{w}}$$

By induction it can be proved that $\mathbf{w}(t) \leq \bar{\mathbf{w}}(t)$ for all t and this proves the inequality. $\square$

Define now the q-dimensional column vectors a, b defined by letting

$$a_s = N^{-1}\delta_s^2 ||YP_s||^2 \qquad b_s = N^{-1}\delta_s^2|\lambda_s|^2$$

We can now state and prove a general convergence result.

Theorem 15 *Let $\rho := \rho(P)$ and let $\bar{\rho}^2$ be the induced 2-norm of the matrix $\rho^2 I + L$. Assume that $L \neq 0$ and that $\bar{\rho}^2 < 1$. Then, there exists a scalar random variable α^* such that*

$$\mathbb{E}||x(t) - \alpha^* \mathbb{1}||^2 \leq A\rho^{2t} + B\bar{\rho}^{2t} \tag{27}$$

where

$$A = w(0) - \frac{||a||}{\bar{\rho}^2 - \rho^2}||\mathbf{w}(0)||$$
$$B = \left(\frac{||a||}{\bar{\rho}^2 - \rho^2} + \frac{||b||}{(1-\bar{\rho})^2}\right)||\mathbf{w}(0)||$$

and where $w(0)$ and $\mathbf{w}(0)$ are defined in (24).

PROOF. Notice that, as showed in the proof of Lemma 14, we have $\mathrm{tr}\,(YPQP^TY^T) \leq \rho^2 w$. Define the sequence $\bar{w}(t)$ as follows:

$$\bar{w}(0) = w(0), \quad \bar{w}^+ = \rho^2\bar{w} + ||a||\,||\bar{\mathbf{w}}||$$

where $\bar{\mathbf{w}}(t) := (\rho^2 I + L)^t\mathbf{w}(0)$. By induction it can be proved that $w(t) \leq \bar{w}(t)$ for all t. Using moreover the fact that $\bar{\rho}^2 > \rho^2$ (since $L \neq 0$), we can estimate

$$\bar{w}(t) = \rho^{2t}w(0) + ||a||\sum_{i=0}^{t-1}\rho^{2(t-1-i)}||(\rho^2 I + L)^i\mathbf{w}(0)||$$
$$\leq \rho^{2t}w(0) + ||a||\sum_{i=0}^{t-1}\rho^{2(t-1-i)}\bar{\rho}^{2i}||\mathbf{w}(0)|| =$$
$$= \left(w(0) - \frac{||a||||\mathbf{w}(0)||}{\bar{\rho}^2 - \rho^2}\right)\rho^{2t} + \left(\frac{||a||\,||\mathbf{w}(0)||}{\bar{\rho}^2 - \rho^2}\right)\bar{\rho}^{2t}$$

Notice now that

$$\mathbb{E}[||x_B(t+1) - x_B(t)||^2] = \mathbb{E}[||x_B(t+1)||^2] + \mathbb{E}[||x_B(t)||^2] - 2\mathrm{tr}\,\mathbb{E}[x_B(t+1)x_B(t)^T]\,.$$

On the other hand, since

$$x_B^+ = x_B + N^{-1}\sum_{s=1}^{q}\chi_0\chi_0^* P_s E_s H_s x$$

we have that

$$\mathrm{tr}\,\mathbb{E}[x_B(t+1)x_B(t)^T] = \mathrm{tr}\,\mathbb{E}[x_B(t)x_B(t)^T] + N^{-1}\sum_{s=1}^{q}\mathrm{tr}\,[\chi_0\chi_0^* P_s\mathbb{E}(E_s)H_s\mathbb{E}(x(t)x_B(t)^T)] = s(t)\,.$$

Using Lemma 14 we can then estimate as follows

$$\mathbb{E}[||x_B(t+1) - x_B(t)||^2] = s(t+1) - s(t) = b^T\mathbf{w}(t) \leq b^T(\rho^2 I + L)^t\mathbf{w}(0) \leq \bar{\rho}^{2t}||b||\,||\mathbf{w}(0)||\,. \tag{28}$$

This shows that $x_B(t)$ converges in mean square sense to a random variable $\alpha^*\mathbb{1}$ and that

$$(\mathbb{E}||x_B(t) - \alpha^*\mathbb{1}||^2)^{1/2} \leq \sum_{s=t}^{\infty}(\mathbb{E}||x_B(s+1) - x_B(s)||^2)^{1/2} \leq \frac{||b||^{1/2}\,||\mathbf{w}(0)||^{1/2}}{1-\bar{\rho}}\bar{\rho}^t\,. \tag{29}$$

Final estimation (27) now immediately follows from the splitting

$$\mathbb{E}||x(t) - \alpha^*\mathbb{1}||^2 = \mathbb{E}||x_B(t) - \alpha^*\mathbb{1}||^2 + \mathbb{E}||y(t)||^2\,. \tag{30}$$

$\square$

Notice that, since $\bar{\rho} > \rho$, then the rate of convergence is determined by the parameter $\bar{\rho}$, namely by the induced 2-norm of the matrix $\rho^2 I + L$.

As for the strategies illustrated in Chapter 5, it is also here interesting to evaluate the mean square distance of the consensus α^* from the initial average $N^{-1}\mathbb{1}^T x(0)$. We have the following result.

Proposition 16 *Let α^* be the random variable defined in Theorem 15. Under the same hypotheses of Theorem 15, we have that*

$$\mathbb{E}|\alpha^* - N^{-1}\mathbb{1}^T x(0)|^2 \leq \frac{1}{N}\frac{||b||}{1-\bar{\rho}^2}||\mathbf{w}(0)||\,.$$

PROOF. Consider $\Delta(t) := x(t) - N^{-1}\mathbb{1}\mathbb{1}^T x(0)$ and $Q(t) := \mathbb{E}[\Delta(t)\Delta(t)^T]$. It is immediate to check that $Q(t)$ satisfies the same evolution law (23). Moreover, we have that $y(t) = Yx(t) = Y\Delta(t)$ and $z_s(t) = H_s x(t) = H_s\Delta(t)$. If we define in this context $x_B(t) = N^{-1}\mathbb{1}\mathbb{1}^T x(t) - N^{-1}\mathbb{1}\mathbb{1}^T x(0)$ we have that the corresponding mean square values $w(t)$, $w_s(t)$, $s(t)$ have exactly the same expression in terms of the matrix Q and, as a consequence, they satisfy the same evolution equations (26). In particular, we obtain that

$$s(t) = b^T\sum_{j=0}^{t-1}\mathbf{w}(j)\,.$$

Using Lemma 14 we can now estimate

$$|s(\infty)| \leq \frac{||b||}{1-\bar{\rho}^2}||\mathbf{w}(0)||\,,$$

from which the thesis immediately follows. □

In the sequel we apply previous results to analyze a particular but significant example.

Example 17 *We assume we have the same exact communication graph of Example 4, namely, the Cayley graph $\mathcal{G}(\mathbb{Z}_N, S)$, where $S = \{0,1\}$. We assume that P_0 is the stochastic Cayley matrix generated by the distribution $\pi_{P_0}(0) = 1-k$, and $\pi_{P_0}(1) = k$, where $k \in [0,1]$. Assume moreover $q = 1$, namely that each system transmits just one scalar signal. Precisely, define H_1 to be the Cayley matrix generated by the distribution $\pi_{H_1}(0) = 1$, and $\pi_{H_1}(1) = -1$. This means that each system i transmits the difference between its own state x_i and the the state x_{i-1} which is known exactly by the system i. It remains to choose the matrix P_1. Our objective is to choose P_1 in such a way that $P = P_0 + P_1H_1 = N^{-1}\chi_0\chi_0^*$. This can be done by letting*

$$\pi_{P_1}(g) = \frac{g+1-N}{N} \qquad g = 1,\ldots,N-2$$

and $k = \frac{N-1}{N}$. Indeed, this definitions yield P_1H_1 with the following generator

$$\pi_{P_1H_1}(0) = 0\,,\ \pi_{P_1H_1}(1) = \frac{2-N}{N}\,,\ \pi_{P_1H_1}(g) = \frac{1}{N}$$

for all $g = 2,\ldots,N-1$. With such a choice we have that $PH_1 = PY = 0$. Notice moreover that $P_1\chi_0 = \lambda_1\chi_0$ implies that

$$\lambda_1 = \sum_{g=0}^{N-1}\pi_{P_1}(g) = \sum_{g=1}^{N-2}\frac{g+1-N}{N} = \sum_{g=1}^{N-1}-\frac{g}{N} = -\frac{(N-1)(N-2)}{2N}$$

and so

$$b = \frac{1}{N}\delta_1^2|\lambda_1|^2 = \delta_1^2\frac{(N-1)(N-2)}{2N^2}.$$

Moreover we have that

$$||H_1P_1||^2 = N\sum_{g=0}^{N-1}|\pi_{P_1H_1}(g)|^2 = \frac{(2-N)^2}{N} + (N-2)\frac{1}{N} = \frac{(N-1)(N-2)}{N}$$

which implies that

$$L = \frac{1}{N}\delta_1^2||H_1P_1||^2 = \delta_1^2\frac{(N-1)(N-2)}{N^2}$$

Finally notice that $YP_1 = (I-P)P_1 = P_1 - \lambda_1 P$ and so

$$\begin{aligned}||YP_1||^2 &= N\sum_{g=0}^{N-1}|\pi_{P_1}(g) - \lambda_1\pi_P|^2\\ &= 2N\frac{(N-1)^2(N-2)^2}{4N^2} + \sum_{g=1}^{N-2}\left|-\frac{g}{N} + \frac{(N-1)(N-2)}{2N}\right|^2 =\\ &= \frac{1}{N}\sum_{g=1}^{N-2}g^2 - \frac{(N-1)^2(N-2)^2}{4N^2}\\ &= \frac{(N-1)(N-2)(N^2+3N-6)}{12N^2}\end{aligned}$$

which implies that

$$a = \delta_1^2\frac{(N-1)(N-2)(N^2+3N-6)}{12N^3}$$

For big N we have that $L \simeq \delta_1^2$, $a \simeq \delta_1^2\frac{N}{12}$ and $b \simeq \frac{\delta_1^2}{2}$. In this case, since we have that $\rho(P) = 0$, applying Theorem 15, we obtain that

$$\mathbb{E}||x(t) - \alpha^*\mathbb{1}||^2 \leq B\delta_1^{2t} \tag{31}$$

where

$$B = \left(\frac{N}{12} + \frac{\delta_1^2}{2(1-\delta_1)^2}\right)\mathbb{E}||Hx(0)||^2$$

Instead, from Proposition 16 we obtain that

$$\mathbb{E}|\alpha^* - N^{-1}\mathbb{1}^T x(0)|^2 \leq \frac{1}{N}\frac{\delta_1^2}{2(1-\delta_1^2)}\mathbb{E}||Hx(0)||^2 \,. \quad (32)$$

Notice that, for small δ_1, the convergence rate towards the consensus established in (31) *is much better than what obtained without noisy data transmission. More precisely, suppose the our goal is to have convergence of the initial states $x_i(0) \in [-M, M]$ to a target configuration $x_i(\infty) \in [\alpha - \epsilon, \alpha + \epsilon]$ where α is a constant depending only on the initial condition $x(0)$ and ϵ describes the desired consensus precision. This is a "practical stability" requirement and it is the only goal achievable through finite data rate transmission. In this case the contraction rate is $C := M/\epsilon$. We assume that the exact data transmissions are substituted by transmissions of precision ϵ uniformly quantized data. In this framework it is known [15] that each uniform quantizer needs C different levels and so the transmission of its data needs an alphabet of C different symbols. On the other hand (see [15]) each logarithmic quantizer needs*

$$\frac{2\log C}{\log\dfrac{1+\delta_1}{1-\delta_1}}$$

different symbols. Let $\delta_1 = 1/2$. We know that the strategy proposed in this example allows a convergence rate $\rho \simeq 1/2$. In this case we need N uniform quantizers and $N(N-2)$ logarithmic quantizers. Thus, the total number of symbols L_{tot} that needs to be transmitted during each sampling period in order to obtain the consensus is

$$L_{tot} = NC + \frac{2}{\log 3}N(N-2)\log C \,.$$

Without the logarithmic quantizers we need only $L_{tot} = NC$ symbols but we obtain a convergence rate $\rho \simeq 1 - 2\pi^2 N^{-2}$. Observe that for large C the total number of symbols L_{tot} in the two cases are slightly different, but we obtain a manifest improvement in terms of rate of convergence.
Finally, notice that the mean square distance of the consensus from the initial average (32) is infinitesimal for $\delta_1 \to 0$.

7 Conclusions

We have derived bounds on the convergence rate to the average consensus for a team of mobile agents that exchange information over time-invariant and randomly time-varying communication networks with symmetries. We have showed that, in time-invariant networks, symmetries yield rather slow convergence to the consensus. In particular for such networks we have computed a tight bound for the convergence rate. We have also showed that, if the communication network is randomly time-varying over a class of networks with symmetries, the achievable performance is much higher. The last part of the paper have been devoted to study the control performance when agents also exchange logarithmically quantized data. It has been shown that adding such links in time-invariant networks with symmetries improves the convergence rate with little growth of the required bandwidth.

Acknowledgements

A. Speranzon's work has been supported by the European project RECSYS FP5-IST-32515 and partially by HYCON FP6-IST-511368.

References

[1] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–86, 1986.

[2] N. Alon and Y. Roichman. Random Cayley graphs and expanders. *Random Structures and Algorithms*, 5:271–284, 1994.

[3] L. Babai. Spectra of Cayley graphs. *Journal of Combinatorial Theory, Series B*, 27:180–189, 1979.

[4] E. Behrends. *Introduction to Markov Chains (with Special Emphasis on Rapid Mixing)*. Vieweg Verlag, 1999.

[5] P. Bolzen, P. Colaneri, and G. De Nicolao. On almost sure stability of discrete-time Markov jump linear systems. In *IEEE Conference on Decision and Control*, 2004.

[6] S. Boyd, P. Diaconis, P. Parrilo, and L. Xiao. Symmetry analysis of reversible Markov chains. Technical report, Department of Statistics, Stanford University, 2003.

[7] S. Boyd, P. Diaconis, and L. Xiao. Fastest mixing Markov chain on a graph. *SIAM Review*, 46:667–689, 2004.

[8] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Mixing times for random walks on geometric random graphs. In *SIAM ANALCO*, 2005.

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

[10] J. Cortés, S. Martínez, and F. Bullo. Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions. *IEEE Transaction on Automatic Control*, 2004. To appear.

[11] R. D'Andrea and G. E. Dullerud. Distributed control design for spatially interconnected systems. *IEEE Transactions on Automatic Control*, 48(9):1478–1495, 2003.

[12] P. Diaconis. *Group Representations in Probability and Statistics*, volume 11 of *IMS Lecture Notes - Monograph Series, S. S. Gupta (ed.)*. Institute of Mathematical Statistics, Hawyard Ca, 1988.

[13] R. Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, Heidelberg, third edition edition, 2005.

[14] N. Elia and S. J. Mitter. Stabilization of linear systems with limited information. *IEEE Transactions on Automatic Control*, 46(9), 2001.

[15] F. Fagnani and S. Zampieri. Quantized stabilization of linear systems: complexity versus performances. *IEEE Transactions on Automatic Control*, 49:1534–1548, 2004.

[16] Y. Fang and K. A. Loparo. Stochastic stability of jump linear systems. *IEEE Transaction on Automatic Control*, 47(7):1204–1208, 2002.

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

[18] G. Ferrari-Trecate, A. Buffa, and M. Gati. Analysis of coordination in multiple agents formations through partial difference equations. Technical Report 5-PV, Istituto di Matematica Applicata e Tecnologie Informatiche, C.N.R., Pavia, Italy, 2005. Submitted for publication.

[19] J. Friedman. A proof of Alon's second eigenvalue conjecture. *Accepted to the Memoirs of the A.M.S.*, 2004.

[20] F. R. Gantmacher. *The theory of matrices.* New York : Chelsea publ., 1959.

[21] C. Godsil and G. Royle. *Algebraic Graph Theory.* Springer, 2001.

[22] Y. Hatano and M. Mesbahi. Agreement of random networks. In *IEEE Conference on Decision and Control*, 2004.

[23] T. W. Hungerford. *Algebra*, volume 73 of *Graduate Texts in Mathematics.* Springer-Verlag, 1974.

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

[25] K. H. Johansson, A. Speranzon, and S. Zampieri. On quantization and communication topologies in multi-vehicle rendezvous. In *Proceedings of IFAC World Congress*, 2005.

[26] J. Lin, A. S. Morse, and B. D. O. Anderson. The multi-agent rendezvous problem: an extended summary. In A. S. Morse V. Kumar, N. E. Leonard, editor, *Proceedings of the 2003 Block Island Workshop on Cooperative Control*, volume 309 of *Lecture Notes in Control and Information Sciences*, pages 257–282. New York: Springer Verlag, 2004.

[27] Z. Lin, B. Francis, and M. Maggiore. On the state agreement problem for multiple nonlinear dynamical systems. In *Proceedings of 16th IFAC World Congress*, 2005.

[28] M. Marodi, F. d'Ovidio, and T. Vicsek. Synchronization of oscillators with long range interaction: Phase transition and anomalous finite size effects. *Phisical Review E*, 66, 2002.

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

[30] M. R. Murty. Ramanujan graphs. *Journal of Ramanujan Mathematical Society*, 18(1):1–20, 2003.

[31] R. Olfati-Saber. Distributed Kalman filter with embedded consensus filters. In *IEEE Conference on Decision and Control*, 2005.

[32] R. Olfati-Saber. Ultrafast consensus in small-world networks. In *American Control Conference*, 2005.

[33] R. Olfati-Saber. Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on Automatic Control*, 51(3):401–421, 2006.

[34] R. Olfati-Saber. Algebraic connectivity ratio of Ramunujan graphs. In *Proceedings of American Control Conference (ACC '07)*, 2007.

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

[36] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

[37] B. Recht and R. D'Andrea. Distributed control of systems over discrete groups. *IEEE Transactions on Automatic Control*, 49(9):1446–1452, 2004.

[38] W. Ren and R. W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transaction on Automatic Control*, 50(5), 2005.

[39] W. Ren, R. W. Beard, and E. M. Atkins. A survey of consensus problems in multi-agent coordination. In *American Control Conference*, 2005.

[40] L. Saloff-Coste. Random walks on finite groups. In Harry Kesten, editor, *Encyclopaedia of Mathematical Sciences*, pages 263–346. Springer, 2004.

[41] S. L. Smith, M. E. Broucke, and B. A. Francis. A hierarchical cyclic pursuit scheme for vehicle networks. *Automatica*, 41(6):1045–1053, 2005.

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

[43] H. G. Tanner, A. Jadbabaie, and G .J. Pappas. Stable flocking of mobile agents, part I and part III. In *IEEE Conference on Decision and Control*, 2003.

[44] A. Terras. *Fourier analysis on finite groups and applications*, volume 43 of *London Mathematical Society Student Texts.* Cambridge University Press, Cambridge Ma, 1999.

[45] Tsitsiklis. *Problems in decentralized decision making and computation.* PhD thesis, Department of EECs, MIT, 1984.

[46] A. Valette. Graphes de Ramanujan. *Astérisque*, 245:247–276, 1997.