Quadratic indices for the analysis of consensus algorithms

Ruggero Carli
Center for Control, Dynamical Systems and Computation,
University of California at Santa Barbara,
Santa Barbara, CA 93106, USA
Email: carlirug@engineering.ucsb.edu

Federica Garin and Sandro Zampieri
DEI (Department of Information Engineering)
Università di Padova
Via Gradenigo 6b, 35131 Padova, Italy
Email: {garin, zampi}@dei.unipd.it

***Abstract*—Average-consensus algorithms allow to compute the average of some agents' data in a distributed way, and they are used as a basic building block in algorithms for distributed estimation, load balancing, formation and distributed control.**

Traditional analysis of linear average-consensus algorithms studies, for a given communication graph, the convergence rate, given by the essential spectral radius of the transition matrix (i.e. the second largest eigenvalues' modulus). For many graph families, such analysis predicts a performance which degrades when the number of agents grows, basically because spreading information across a larger graph requires a longer time. However, if you consider other well-known quadratic performance indices (involving all the eigenvalues of the transition matrix), the scaling law with respect to the number of agents can be different. This is consistent with the fact that, in many applications, for example in estimation problems, it is natural to expect that a larger number of cooperating agents has a positive, not a negative effect on performance.

In this paper, we consider two different quadratic indices, which arise naturally from some estimation and control problems in which the consensus algorithm is used. We provide analytical results on their asymptotic behavior when the communication network belongs to some families of Cayley graphs. This framework includes grids on toruses in arbitrary dimensions, which are conjectured to be a good approximation of random geometric graphs; indeed, we show simulation results supporting this conjecture.

I. Introduction

Average-consensus algorithms allow to compute an average in a distributed way. These algorithms have been extensively applied to the solution of problems of distributed estimation [19] and of sensor calibration for sensor networks [12], to load balancing for distributed computing systems [7], and to mobile multi-vehicles coordination [6]. We refer the reader to the survey [20], the recent books [25], [4] and references therein.

In this paper, we focus on linear average-consensus. Starting from the natural and widely studied algorithm proposed by Tsitsiklis [26] since the early $80's$, the average consensus problem has been addressed in many different directions. Convergence conditions both for fixed and dynamically changing interaction topologies have been provided [21], [17], [5]. Randomized algorithms has been considered in [9], [3]. In [8], [22] the authors investigate the case of data loss in the transmission between the nodes. The effects of noisy communications has been analyzed in [15]. Recently a large attention has been captured by the realistic situation of quantized data transmission [16], [18], [1], [11].

In this paper we are interested in evaluating the performance of linear time-invariant average-consensus algorithms. Typically this kind of analysis exploits results from Markov chains literature, and is focused on predicting the speed of convergence to the average, when computation time grows. There has been an extensive literature on this topic, with both analysis and optimization of asymptotic convergence speed which is given by the dominant mode of the transition matrix [5], [27]. However, we believe that when convergence to the average is not an objective per se, but is used to solve an estimation or control problem, it is important to consider different performance measures, more tightly related to the actual objective pursued.

In this paper we introduce two different functional costs which arise quite naturally in control and estimation problems. The first one is a classical LQ functional cost which evaluates the performance of the average-consensus algorithm by calculating the L_2 norm of a suitable variable; this cost represents a different way to evaluate the transient phase of the algorithm. In fact, also in classical control theory, there are various ways to evaluate the transient performance of a control strategy: one is based on the dominant eigenvalues, and the corresponding control methodology relies on the possibility to suitably allocate such eigenvalues; a second one is based on the L_2-norm of the transient and this yields to the so-called linear quadratic optimal control methodology. Our functional is this second kind of cost, for the consensus algorithm. We will show that this functional cost depends on all the eigenvalues of the transition matrix, and our main contribution will be to characterize it for some relevant families of graphs. The second cost we propose is related to the estimation error made by a network of sensors when averaging their measurements. For many families of graphs (including geometric graphs), the speed of convergence, as evaluated by the essential spectral radius, deteriorates with larger networks, but on the contrary in estimation one would expect that a broader number of measurements should improve the quality of the final estimate. We propose a natural performance measure (average error variance) and, for some families of graphs, we find its scaling

laws with respect to both number of vertices and computation time, so that we can suggest useful criteria in the design of the size of large-scale sensor networks.

The paper is organized as follows. In Section II we formulate the problem we deal with. In Section III we introduce three families of graphs, the Cayley graphs, the grids and the random geometric graphs: these families are the subject of our investigations. In Section IV we illustrate our main theoretical and numerical results. Finally in the Appendix we collect the proofs of the theoretical results stated in the previous section.

Mathematical Preliminaries

Before proceeding, we collect some useful definitions and we fix some notation. More detailed definitions about Cayley graphs will be given in Sect. III-A. In this paper, $\mathcal{G} = (V, E)$ denotes an *directed graph* where V is the set of vertices, $N = |V|$ the number of vertices, and E is the set of directed edges, i.e., a subset of $V \times V$.

A matrix M is *nonnegative* if $M_{ij} \geq 0$ for all i and j. A square matrix M is *stochastic* if it is nonnegative and the sum along each row of M is equal to 1. Moreover, a square matrix M is *doubly-stochastic* if it is stochastic and the sum along each column of M is equal to 1. Given a nonnegative matrix $M \in \mathbb{R}^{N\times N}$, we define the induced graph $\mathcal{G}_M$ by taking N nodes and putting an edge (j, i) in E if $M_{ij} > 0$. Given a graph $\mathcal{G}$ on V, the matrix M is *adapted* to, or *compatible* with, $\mathcal{G}$ if $\mathcal{G}_M \subseteq \mathcal{G}$.

Now we give some notational conventions. Vectors will be denoted with bold letters. Given a vector $\boldsymbol{v} \in \mathbb{R}^N$ and a matrix $M \in \mathbb{R}^{N\times N}$, we let $\boldsymbol{v}^T$ and M^T respectively denote the transpose of $\boldsymbol{v}$ and of M. We let $\sigma(M)$ denote the set of eigenvalues of M. With the symbol $\mathbf{1}$ we denote the N-dimensional vector having all the components equal to 1.

Given any set A with finite cardinality $|A|$, $\mathbb{R}^A$ will denote the vector space isomorphic to $\mathbb{R}^{|A|}$, made of vectors where indices are elements of A instead of $\{1, 2, \ldots, |A|\}$. Analogously, $\mathbb{R}^{A\times A}$ will denote the vector space of all linear maps from $\mathbb{R}^A$ to $\mathbb{R}^A$.

II. Problem formulation

We start this section by briefly describing the standard discrete-time consensus algorithm. Assume that we have a set of agents V and a graph $\mathcal{G}$ on V describing the feasible communications among the agents. For each agent $i \in V$ we denote by $x_i(t)$ the estimate of the average of agent i at t-th iteration. Standard consensus algorithms are built by choosing a doubly-stochastic matrix $P \in \mathbb{R}^{N\times N}$ compatible with $\mathcal{G}$ and assuming that at every step t agent i updates its estimates according to

$$x_i(t+1) = \sum_{j=1}^{N} P_{ij} x_j(t). \tag{1}$$

More compactly we can write

$$\boldsymbol{x}(t+1) = P\boldsymbol{x}(t), \tag{2}$$

where $\boldsymbol{x}(t)$ is the column vector whose i-th entry is $x_i(t)$.

It is well-known in the literature [4] that, if P is primitive then the algorithm (2) solves the *average consensus problem*, namely

$$\lim_{t\to\infty} \boldsymbol{x}(t) = x_{\text{ave}}\mathbf{1}, \tag{3}$$

where $x_{\text{ave}} = \frac{1}{N}\sum_{i=1}^{N} x_i(0)$. From now on we assume the following property.

Assumption 1: P is a primitive doubly-stochastic matrix. □

Traditionally the performance of the average consensus algorithm is evaluated by considering the *asymptotic convergence factor*, defined as

$$r_{\text{asym}} = \sup_{x(0)} \limsup_{t\to\infty} \left(\|x(t) - x_{\text{ave}}\|_2\right)^{\frac{1}{t}}. \tag{4}$$

It is well known [4] that, if P satisfies Assumption 1 and the initial condition can be any arbitrary vector of $\mathbb{R}^N$, then r_{asym} coincides with the *essential spectral radius* of P, that we denote by $\rho_{\text{ess}}(P)$[1]. Typically, for many graph families, when considering sequences of matrices $\{P_N\}$ of increasing size, it turns out that

$$\lim_{N\to\infty} \rho_{\text{ess}}(P_N) = 1. \tag{6}$$

Notice that equation (6) predicts a performance which degrades as the number of agents increases. This is not surprising, since intuitively one should expect that the larger is the graph, the longer is the time required to spread the information across the nodes. A mathematical characterization of (5) has been carried out for graphs exhibiting Cayley symmetries in [5] and for the random geometric graphs in [3].

In this paper we evaluate the performance of the average consensus algorithm according to two different functional costs. The first one is a classical LQ functional cost which accounts for the speed of the average consensus algorithm by calculating the L_2-norm of a suitable trajectory; the second one is related to the estimation error made by a network of sensors when averaging their measurements. We proceed now by presenting them separately.

A. Transient performance evaluation by L_2-norm: a LQ cost

In this subsection we assume that the initial condition $\boldsymbol{x}(0)$ satisfies the following condition.

Assumption 2: The initial condition $\boldsymbol{x}(0)$ is a random variable such that $\mathbb{E}[\boldsymbol{x}(0)] = 0$ and $\mathbb{E}\left[\boldsymbol{x}(0)\boldsymbol{x}^T(0)\right] = \sigma_0^2 I$ for some $\sigma_0^2 > 0$. □

Without loss of generality, we will consider $\sigma_0^2 = 1$ throughout this paper.

When dealing with the average consensus problem it is convenient to introduce the following random variable

$$\Delta(t) = \boldsymbol{x}(t) - \boldsymbol{x}_{\text{ave}}\mathbf{1} = \left(I - \tfrac{1}{N}\mathbf{1}\mathbf{1}^T\right)\boldsymbol{x}(t),$$

[1]For the sake of the clarity, we recall that for a primitive doubly-stochastic matrix $\rho_{\text{ess}}(P)$ is given by the second largest eigenvalues' modulus, i.e.,

$$\rho_{\text{ess}}(P) := \max_{\lambda\in\sigma(P)\setminus\{1\}} |\lambda|. \tag{5}$$

where the last equality follows from the fact that, since P is doubly-stochastic then $\mathbf{1}^T\boldsymbol{x}(t) = \mathbf{1}^T\boldsymbol{x}(0)$ for all $t \geq 0$. Observe that $\Delta(t)$ represents the distance of $\boldsymbol{x}(t)$ from the average of the initial conditions. It is easy to see that Δ satisfies the same recursive equation as $\boldsymbol{x}$, i.e.,

$$\Delta(t+1) = P\Delta(t) \tag{7}$$

and that $\boldsymbol{x}(t) \to \boldsymbol{x}_{\text{ave}}\mathbf{1}$ if and only if $\Delta(t) \to \mathbf{0}$.

In control theory a classical way of evaluating the performance of system (7) would be considering a linear quadratic cost of the form

$$J_\Delta(P) = \frac{1}{N}\sum_{t=0}^{\infty}\mathbb{E}\left[\Delta^T(t)Q\Delta(t)\right]$$

where Q is an pre-assigned semidefinite positive matrix. In our setup we assume that $Q = I$ and hence the above expression reduces to the following functional cost

$$J_\Delta(P) := \frac{1}{N}\sum_{t=0}^{\infty}\mathbb{E}\|\Delta(t)\|^2. \tag{8}$$

It is worth noting that $J_\Delta(P)$ represents also as the L_2-norm of the random process $\{\Delta(t)\}_{t=0}^{\infty}$.

In Section IV we will characterize the behavior of $J_\Delta(P)$ for some relevant graph families as the number of agents N varies.

We provide now a characterization of J_Δ that will be useful later on. Notice that Assumption 2 implies that $\Delta(0)$ is a random variable such that $\mathbb{E}[\Delta(0)] = \mathbf{0}$ and $\mathbb{E}\left[\Delta(0)\Delta^T(0)\right] = I - \frac{1}{N}\mathbf{1}\mathbf{1}^T$. Hence,

$$\begin{aligned} J_\Delta(P) &= \frac{1}{N}\sum_{t=0}^{\infty}\text{trace}\left\{\mathbb{E}\left[\Delta(t)\Delta^T(t)\right]\right\} \\ &= \frac{1}{N}\sum_{t=0}^{\infty}\text{trace}\left[P^t\left(I - \tfrac{1}{N}\mathbf{1}\mathbf{1}^T\right)\left(P^t\right)^T\right] \end{aligned}$$

If P is normal, i.e., $PP^T = P^TP$, the above expression can be written as a function of only the eigenvalues of P. Precisely, by observing that $P^t\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right) = \left(P\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)\right)^t$, and that $\sigma\left(P\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)\right) = \{0\} \cup \sigma(P) \setminus \{1\}$, we can write

$$\begin{aligned} J_\Delta(P) &= \frac{1}{N}\sum_{t=0}^{\infty}\sum_{\lambda\in\sigma(P)\setminus\{1\}}|\lambda|^{2t} \\ &= \frac{1}{N}\sum_{\lambda\in\sigma(P)\setminus\{1\}}\frac{1}{1-|\lambda|^2} \end{aligned}$$

Remark 1: We end this subsection by remarking that the functional cost (8) has been also analyzed in [28] in a different context. The authors in [28] consider a stochastic model for distributed average consensus where each node, updates its local variable with a weighted average of its neighbors' values as in (1), but each new value is corrupted by an additive noise with zero mean, i.e.,

$$x_i(t+1) = \sum_{j=1}^{N}P_{ij}x_j(t) + v_i(t), \qquad i = 1,\dots,N, \tag{9}$$

where $v_i(t)$, $i = 1,\dots,N, t = 0,1,\dots$ are independent random variables, identically distributed, with zero mean and unit variance. To analyze the performance of the algorithm, the authors introduce the variable

$$z(t) = \left(I - \tfrac{1}{N}\mathbf{1}\mathbf{1}^T\right)\boldsymbol{x}(t),$$

and the corresponding functional cost

$$\delta_{ss}(P) := \frac{1}{N}\sum_{t=0}^{\infty}\mathbb{E}\|z(t)\|^2$$

Observe that, due to the presence of the noise, differently from $\Delta(t)$, $e(t) \neq \boldsymbol{x}(t) - \boldsymbol{x}_{\text{ave}}\mathbf{1}$; in other words $e(t)$ quantifies the distance of the states from their current average which, in general, differs from the average of their initial conditions. Thus the mean-square deviation $\delta_{ss}(P)$ can be viewed as a measure of how well the weight matrix P is able to enforce consensus (but not in general the average consensus), despite the additive errors introduced at each node at each step. Some straightforward manipulations show that, if P is normal, then $\delta_{ss}(P) = J_\Delta(P)$.

B. Quadratic error in distributed estimation

In this subsection we consider the following problem of distributed estimation: N sensors measure the same real quantity y plus independent. noises. To be more precise, if v_i denotes the measurement made by the i-th sensor, we have that $v_i = y + n_i$ where n_i, $i \in \{1,\dots,N\}$, are independent zero-mean noises with the same variance σ_0^2 (without loss of generality, we will assume $\sigma_0^2 = 1$). If all the measurements were available at the same location, it is well known that the optimal estimate would be given by the mean of all measurements, i.e., $1/N\sum_{i=1}^{N}v_i$.

In our setup, where the sensors could be constrained by the graph $\mathcal{G}$ to communicate only with a limited number of neighbors, the average of the measurements $v_i, i \in \{1,\dots,N\}$ can be computed efficiently in a distributed way by means of an average consensus algorithm. Let $\boldsymbol{x}(0)$ be such that its i-th component $x_i(0)$ is equal to v_i. Then the estimate $x(t)$ is updated by the sensors according to (1), where P is a doubly-stochastic matrix compatible with the graph $\mathcal{G}$. Clearly, under Assumption 1, $\boldsymbol{x}(t) \to \left(1/N\sum_{i=1}^{N}v_i\right)\mathbf{1}$. In this context, since the goal is estimating y, it is quite natural to introduce the error variable

$$\boldsymbol{e}(t) = \boldsymbol{x}(t) - y\mathbf{1}$$

and the corresponding quadratic functional cost

$$J_e(P,t) = \frac{1}{N}\mathbb{E}\left[\boldsymbol{e}^T(t)\boldsymbol{e}(t)\right].$$

For our problem, it is easy to show that the cost $J_e(P,t)$ can be re-written as

$$J_e(P,t) = \tfrac{1}{N}\text{ trace}\left((P^t)^TP^t\right)$$

If P is normal, then this is equivalent to

$$J_e(P,t) = \tfrac{1}{N}\sum_{\lambda\in\sigma(P)}|\lambda|^{2t}.$$

In the next sections, we will study the asymptotic behavior of $J_e(P,t)$ when both N and t grow to infinity, for some families of graphs. This result is particularly relevant because it suggests the right trade-off between number of nodes and computation time in the design of large-scale sensor networks.

III. Families of graphs

We introduce here the main family of graphs and transition matrices we are going to consider in our analysis.

A. Abelian Cayley graphs and matrices

First of all let's recall the definition of Cayley graphs: given a group $(G,+)$ and a set $S \subseteq G$, the Cayley graph $\mathcal{G}(G,S)$ is a directed graph with vertex set G and edge set $E = \{(g,h) : h-g \in S\}$. We will consider finite graphs, with $|G| = N$, and matrices associated with such graphs, which respect the strong symmetries of the graph: we say that a matrix $P \in \mathbb{R}^{G\times G}$ (i.e. with entries labeled by indexes belonging to G) is Cayley if $P_{g,h} = P_{g+k,h+k} \forall g,h,k \in G$. This is equivalent to say that there exists a map $\pi : G \to \mathbb{R}$ such that $P_{h,k} = \pi(h-k)$; such function is called the generator of the Cayley matrix P.

Throughout this paper, we will also assume that the graph associated with P is strongly connected and aperiodic, and that P is stochastic, i.e., $\pi(g) \geq 0 \, \forall g \in G$ and $\sum_{g\in G} \pi(g) = 1$. Notice that a stochastic Cayley matrix is also doubly-stochastic.

In this paper, we restrict our attention to the case when $G = \mathbb{Z}_n^d$ even though most results can be generalized to any finite Abelian group. Under this assumption, the eigenvalues and eigenvectors of P have the following simple expression: for any $\boldsymbol{h} = (h_1,\dots,h_d) \in \mathbb{Z}_n^d$,

$$\lambda_{\boldsymbol{h}} = \sum_{\boldsymbol{k}\in\mathbb{Z}_n^d} \pi(\boldsymbol{h}) e^{-i(\frac{2\pi}{n}h_1k_1+\dots+\frac{2\pi}{n}h_dk_d)}$$

is an eigenvalue with corresponding eigenvector $\boldsymbol{v}_{\boldsymbol{h}} \in \mathbb{R}^{\mathbb{Z}_n^d}$ defined by

$$\boldsymbol{v}_{\boldsymbol{h}}(\boldsymbol{k}) = \tfrac{1}{\sqrt{N}} \, e^{i(\frac{2\pi}{n}h_1k_1+\dots+\frac{2\pi}{n}h_dk_d)} \, .$$

Notice that such eigenvectors are orthonormal, so that P is a normal matrix.

As a simple example, when $d = 1$, i.e. $G = \mathbb{Z}_N$, you obtain that P is a circulant $N \times N$ matrix, with first row equal to the vector $[\pi(0),\dots,\pi(N-1)]^T$, and with eigenvalues/eigenvectors $\lambda_h = \sum_{k=0}^{N-1} \pi(k) e^{-i\frac{2\pi}{N}hk}$ and $\boldsymbol{v}_h = [1, e^{i\frac{2\pi}{N}h}, e^{2i\frac{2\pi}{N}h}, \dots, e^{(N-1)i\frac{2\pi}{N}h}]^T$, for $h = 0,\dots,N-1$.

In our analysis we want to consider families of Cayley graphs, with a growing number of vertices, but with constant degree, and with the same algebraic structure and same values for the non-zero entries of P. For example, we can look at a a circular graph where each node talk to itself, to its first two neighbours on the right and its first neighbour to the left, each with weight $1/3$, regardless the number of agents.

First of all, we consider a family of groups $G_n = \mathbb{Z}_n^d$, for some fixed d and growing n; let $N = |G_n| = n^d$. Then, we have to define the neighbours and the weights. We fix a positive integer δ, we define the set $D_\delta = \{-\delta, -\delta+1, \dots, +\delta\}^d$ and we fix $|D_\delta|$ real numbers $p_{\boldsymbol{h}}$, $\boldsymbol{h} = (h_1,\dots,h_d) \in D_\delta$ such that $p_{\boldsymbol{h}} \geq 0 \, \forall \boldsymbol{h}$ and $\sum_{\boldsymbol{h}\in D_\delta} p_{\boldsymbol{h}} = 1$. Then, for any $n > \delta$, we construct the Cayley matrix $P_n \in \mathbb{R}^{\mathbb{Z}_n^d \times \mathbb{Z}_n^d}$ with generator $\pi_n : \mathbb{Z}_n^d \to \mathbb{R}$ defined by $\pi_n(\boldsymbol{g}) = p_{\boldsymbol{h}}$ if there is an $\boldsymbol{h} \in D_\delta$ such that, for all $l = 1,\dots,d$ $g_l = h_l \bmod n$, and $\pi_n(\boldsymbol{g}) = 0$ otherwise. Note that for any $n \geq \delta$ π_n is well-defined.

We can also do a similar construction taking $G = \mathbb{Z}^d$ and defining $\pi(\boldsymbol{g}) = p_{\boldsymbol{g}}$ if $\boldsymbol{g} \in D_\delta$ and $\pi(\boldsymbol{g}) = 0$ otherwise. We assume that there are enough non-zero weights $p_{\boldsymbol{h}}$ so as to ensure that the corresponding infinite graph is connected. Moreover, we assume there are self-loops, i.e., $p_{\boldsymbol{0}} \neq 0$. These two assumptions guarantee that all matrices P_n of the sequence we have constructed above are primitive; also recall that P_n are doubly-stochastic and normal.

We introduce here also a useful notation, defining the Laurent polynomial $p(z) \in \mathbb{R}[z_1, z_1^{-1}, \dots, z_d, z_d^{-1}]$ given by

$$p(z) = \sum_{\boldsymbol{k}\in D_\delta} p_{\boldsymbol{k}} z_1^{k_1} \dots z_d^{k_d}$$

We will refer to the above construction of a family of Cayley matrices $\{P_n\}_{n\geq\delta}$ with the short name 'Cayley matrix family associated with $p(z_1,\dots,z_d)$'. With this notation, P_n has eigenvalues $\lambda_{\boldsymbol{h}} = p(e^{-i\frac{2\pi}{n}h_1}, \dots, e^{-i\frac{2\pi}{n}h_d})$, $\boldsymbol{h} = (h_1,\dots,h_d) \in \mathbb{Z}_n^d$.

Note that when we write $e^{i\frac{2\pi}{n}h_r}$ with $h_r \in \mathbb{Z}_n$, we mean that we can substitute h_r with any integer which is equal to $h_r \bmod n$. Later, we will need the specific choice of $h_r \in \{0,1,\dots,n-1\}$, which we will denote by $\boldsymbol{h} \in V_n$, $V_n = \{0,\dots,n-1\}^d$. When needed, we will actually identify vertices of the graph with V_n rather than G_n.

B. Grids in $\mathbb{R}^d$

The families of Cayley graphs on the group Z_n^d presented above can be seen as grids on a (multi-dimensional) torus. An interesting result by Boyd et al. [2] allows to compute the eigenvalues and eigenvectors also of grids on a cube in $\mathbb{R}^d$, which are the same as the one on a torus except that they are suitably modified at the borders.

More precisely, define the following family of matrices. Consider P_{2n} a Cayley matrix on $\mathbb{Z}_{2n}^d$ associated with $p(z_1,\dots,z_d)$, and assume that the coefficients $p_{\boldsymbol{h}}$ satisfy the following quadrantal symmetry: $p_{\boldsymbol{h}} = p_{\boldsymbol{k}}$ if $\forall i, h_i = \pm k_i$. This assumption implies that reflections σ_r on G_n defined by $\sigma_r(\boldsymbol{h}) = \boldsymbol{k}$ with $k_l = h_l$ if $l \neq r$ and $k_r = 2n-1-r$, are symmetries of the labeled grid on the torus. For example, Fig. 1 shows the axis of reflection of σ_1 for the case $d = 1$. It is convenient here to identify G_n with the set $V_n = \{0,\dots,n-1\}^d$, and consider $\sigma_r : V_n \to V_n$.

Now denote by H the group generated by $\sigma_1,\dots,\sigma_d$ and consider, for all $\boldsymbol{g} \in V_n \subseteq V_{2n}$, the orbit $O_{\boldsymbol{g}} = \{\eta(\boldsymbol{g}) : \eta \in H\} \subseteq V_{2n}$. Finally, define $\overline{P}_n : \mathbb{R}^{V_n} \to \mathbb{R}^{V_n}$ by $(P_n)_{\boldsymbol{h},\boldsymbol{k}} = \sum_{\boldsymbol{l}\in O_{\boldsymbol{k}}} P_{\boldsymbol{h},\boldsymbol{l}}$, for all $\boldsymbol{h},\boldsymbol{k} \in V_n$. Notice that $\overline{P}_n$ is symmetric and that, apart from the borders, $\overline{P}_n$ associates to edges of the

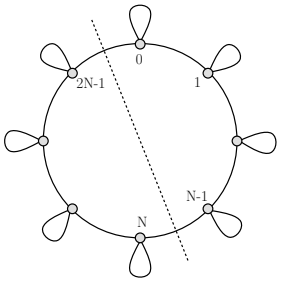

Fig. 1. Circle with $2N$ vertices and reflection axis corresponding to the map $l \mapsto 2N-1-l$, used in the construction of a line with N vertices.

grid the same coefficients that P_n associates to edges of the grid on the torus.

We will refer to the above construction of a family of matrices $\{\overline{P}_n\}_{n\geq\delta}$ with the short name 'grid matrix family associated with $p(z_1,\dots,z_d)$'.

Using [2, Prop. 3.2], we can find the eigenvalues of $\overline{P}_n$:

$$\bar{\lambda}_{\boldsymbol{h}} = p(e^{i\frac{\pi}{n}h_1},\dots,e^{i\frac{\pi}{n}h_d}), \qquad \boldsymbol{h}\in V_n\,.$$

C. Random geometric graphs

The random geometric graph is a random undirected graph drawn on a bounded region, e.g., the d-dimensional unitary cube $[0,1]^d$. It is generated by

- placing vertices at random, uniformly and independently inside the region, and
- connecting two vertices if and only if the euclidean distance between them is at most a pre-assigned threshold r.

The random geometric graph was first introduced in [13] and has been deeply studied under a communications and information-theoretic point of view in [14]. It has recently witnessed a large interest in many applications; particularly it has been successfully used to model wireless communication [10]. In Section IV-B, given a random geometric graph $\mathcal{G}(V,E)$, we will associate to it a doubly stochastic matrix P built according to the *Metropolis weights* rule [29]; precisely, if P_{ij} denotes the element of P in the i-th row and in the j-th column we will have

$$P_{ij} = \begin{cases} \frac{1}{1+\max\{d_i,d_j\}} & \text{if } (i,j)\in E \\ 1-\sum_{(i,k)\in E\setminus\{(i,i)\}} P_{ik} & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$$

where $d_i = |\mathcal{N}_i\setminus\{(i,i)\}|$ with $\mathcal{N}_i=\{j\in V\,|\,(i,j)\in E\}$. In other words the weight on each edge is one over one plus the largest degree at its two incident vertices, and the self-weights are chosen so the sum of weights at each node is 1.

IV. Main results

A. Asymptotic costs for Cayley graphs and grids

We state here our main theoretical results: an asymptotic analysis of the proposed quadratic indices for the families of Cayley graphs and of grids described in previous section. The proofs can be found in the Appendix.

Proposition 1 (LQ cost asymptotics): Given $\{P_n\}_{n\geq\delta}$ a Cayley or a grid matrix family associated with $p(z_1,\dots,z_d)$, there exist $C_d, C'_d>0$ (depending only on d) such that:

- if $d=1$,
$$C_1 N \leq J_\Delta(P_n) \leq C'_1 N\,;$$
- if $d=2$,
$$C_2 \log N \leq J_\Delta(P_n) \leq C'_2 \log N\,;$$
- if $d\geq 3$,
$$C_d \leq J_\Delta(P_n) \leq C'_d\,.$$

□

To give a better understanding of the above Theorem, we propose an example illustrating an interestig comparison between the behavior of the functional cost J_Δ and of the essential spectral ρ_{ess} as $n\to\infty$ of a particular sequence of Cayley graphs. We will see how the evaluation of the performance of the average consensus algorithms, in the asymptotic regime $n\to\infty$, is strictly related to the choice of the functional cost.

Example 1: Consider the sequence of Cayley matrices $\{P_n\}$ built as follows. For each n, let $\mathcal{G}=\mathbb{Z}_n^3$ and let $S=\{(0,0,0),\ (1,0,0),\ (0,1,0),\ (0,0,1),\ (-1,0,0),\ (0,-1,0),\ (0,0,-1)\}$. Moreover let $\pi(g)=\frac{1}{7}$ for all $g\in S$. It is well known (see [5]) that, in this case,

$$\rho_{\text{ess}}(P_n) \geq 1-\frac{C}{N^3}$$

where C is a constant independent from the topology of the graphs. From the above inequality it turns out that, if we consider as functional cost the *asymptotic covergence factor* defined in (4), then the performance of the average consensus algorithms associated to the sequence $\{P_n\}$ degrades drastically as $n\to\infty$. Instead, Proposition 1 guarantees the existence of constants C_3 and C'_3 such that $C_3\leq J_\Delta(P_n)\leq C'_3$ for all n. □

Proposition 2 (Quadratic estimation error asymptotics): Given $\{P_n\}_{n\geq\delta}$ a Cayley or a grid matrix family associated with $p(z_1,\dots,z_d)$, there exist constants $c_0=c'_0=1$, $c_1,\dots,c_d,c'_1,\dots,c'_d>0$ and $k\in(0,1)$ such that

$$\max_{l=0,\dots,d}\frac{1}{n^{d-l}}\frac{c_l}{t^{l/2}} \leq J_e\,(P_n,t) \leq k^t + \sum_{l=0,\dots,d}\frac{1}{n^{d-l}}\frac{c'_l}{t^{l/2}}$$

□

Corollary 1: Given $\{P_n\}_{n\geq\delta}$ a Cayley or a grid matrix family associated with $p(z_1,\dots,z_d)$, there exists constants $n_0\in\mathbb{N}$, $k_1,k_2>0$ such that, for all $n\geq n_0$,

$$k_1\max\left\{\frac{1}{N},\frac{1}{t^{d/2}}\right\} \leq J_e\,(P_n,t) \leq k_2\max\left\{\frac{1}{N},\frac{1}{t^{d/2}}\right\}$$

□

Notice that if the average of all sensors' measurements was performed in a centralized way, the exact average thus obtained would be the best possible estimate of the measured value (under the simple model we are considering), but it would still have error variance $1/N$. Thus, it is not surprising to find

a term $1/N$ in the asymptotic behavior of the decentralized algorithm running on grids. What is more interesting is to look at the dependence on N and t, which shows that a sensible design of the number of nodes should take into account also the computational time allowed for communication and computation. In fact, Corollary 1 clearly shows that when both t and N grow there are two very different regimes, a first one with $N \ll t^{d/2}$, where the error decays as $1/N$, and a second regime with $N \gg t^{d/2}$ where regardless the number of nodes the cost is dominated by a term scaling as $1/t^{d/2}$. Finally, it is interesting to notice that, despite $\rho_{\text{ess}} \to 1$ for $N \to \infty$ would suggest that these families of graphs have decreasing performance for growing number of agents, indeed it is clear that a bigger number of measurements can improve the quality of the estimate, and in fact $\lim_{t,N\to\infty} J_e(P_n, t) = 0$.

B. Random geometric graphs

In this section we focus on the other relevant family of graphs we are dealing with in this paper, the random geometric graph. While several probabilistic results are known about the number of components in the graph as a function of the threshold r and the number of vertices N (see e.g. the monograph [23]), no comprehensive theoretical characterization has been provided yet for the behavior of the eigenvalues of doubly-stochastic matrices associated to random geometric graphs. In this direction only few results are present so far. It is worth citing them briefly. In [3], the authors first prove some regularity properties on the degrees of the nodes of random geometric graphs; then, relying on these results, they find a lower bound for the mixing time of random walks on random geometric graphs with the mixing time of random walks on regular grids on torus (the mixing time is related to the essential spectral radius of the transition matrices associated). In [24] an asymptotic spectral concentration result is presented. Families of random geometric graphs of increasing size N are considered; they are built on $[0,1]^d$ with a threshold r which is assumed to tend to 0 as $N \to \infty$, but in such a way that the graphs so obtained have and increasing degree and are almost surely connected. Under this assumption it is shown that the spectrums of the transition matrices of the random walks on these families of graphs converge, as $N \to \infty$, to those of the graphs on deterministic grids.

In this section we provide some numerical results characterizing the behavior of the cost J_Δ and J_e for random geometric graphs. Interestingly, we will bring to light further similarities between the random geometric graphs and the deterministic grids.

Figures 2, 3, 4 and 5 refer to J_Δ. Precisely, in Figure 2 we depicted the behavior of J_Δ/N for $d = 1$, in Figure 3 the behavior of $J_\Delta/\log N$ for $d = 2$ and in Figures 4 and 5 the behavior of J for $d = 3$ and $d = 4$, respectively. For each value of d we run simulations from $N = 50$ up to 600. We consider families of connected graphs of increasing size obtained with a decreasing threshold r in a such a way that the average size of the neighborhood of the nodes is kept almost constant independently from the value of N (in this specific simulations close to 12 for any value of d). For each value of N we calculated the value of the plotted variable as the mean of 50 trials.

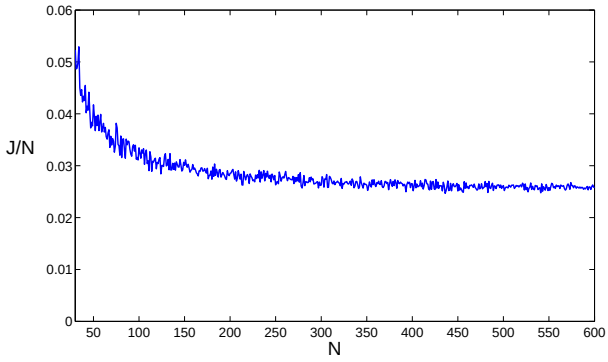

Fig. 2. Behavior of $\frac{J_\Delta}{N}$ for $d = 1$.

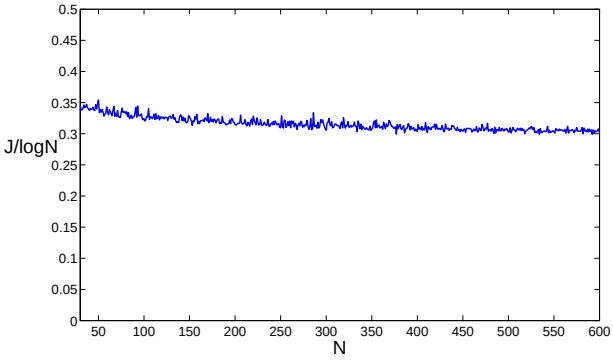

Fig. 3. Behavior of $\frac{J_\Delta}{\log N}$ for $d = 2$.

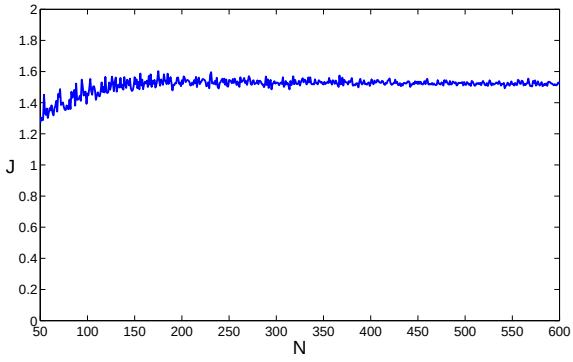

Fig. 4. Behavior of J_Δ for $d = 3$.

From Figures 2, 3, 4, 5, we can infer that J_Δ increases linearly for $d = 1$, logarithmically for $d = 2$, whereas it becomes asymptotically constant for $d = 3$ and $d = 4$. The analogy with Proposition 1 is evident.

Figures 6, 7, 8 and 9 provide numerical results for J_e. Again we run simulations for $N = 50$ up to 600 by considering families of connected graphs of increasing size built with a decreasing threshold r as in the previous set of simulations;

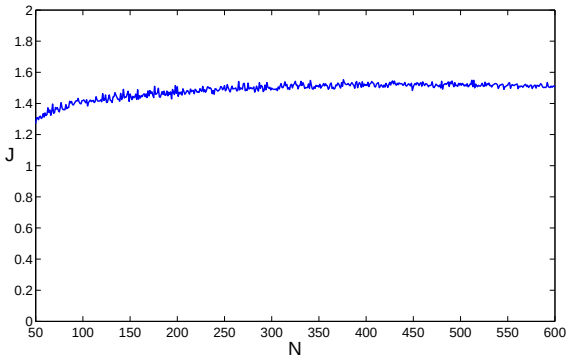

Fig. 5. Behavior of J_Δ for $d = 4$.

however, in this case all the figures refer to the 2-dimensional case $d = 2$. Our aim is to underline the different scaling when both t and N grow, with t being different functions of N: t constant, $t = \sqrt{N}$, $t = N$, $t = N^{3/2}$. In Figure 6 we plotted the behavior of $J_e(P_N, 20)$, in Figure 7 the behavior of $\sqrt{N}\, J_e\left(P_N, \sqrt{N}\right)$, in Figure 8 the behavior of $N\, J_e\left(P_N, N\right)$ and in Figure 9 the behavior of $N\, J_e\left(P_N, N^{3/2}\right)$. From the drawn plots, we can deduce that J_e evaluated for the random geometric graphs, exhibits a behavior very similar to the one stated in Corollary 1.

Finally, these numerical results emphasize further evident similarities between the spectral behavior of the transition matrices built on random geometric graphs and the spectral behavior of the transition matrices associated to the deterministic grids.

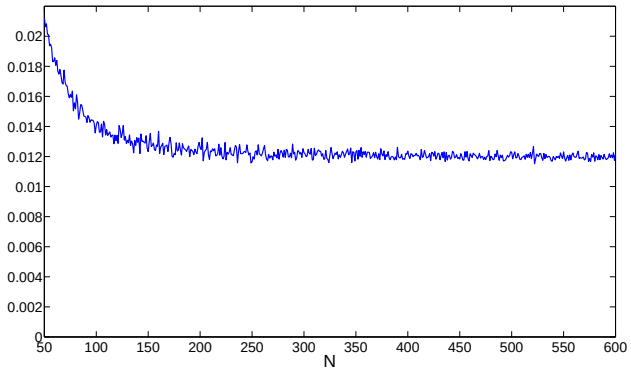

Fig. 6. Behavior of $J_e(P_N, 20)$.

APPENDIX

We give here the main steps of the proof of our main results (Propositions 1 and 2). For notation and definitions, we refer the reader to Sect. III-A and Sect. III-B.

A. Preliminary remarks

Our bounds for both costs are based on a simple but effective technique for getting a bound for the eigenvalues. In

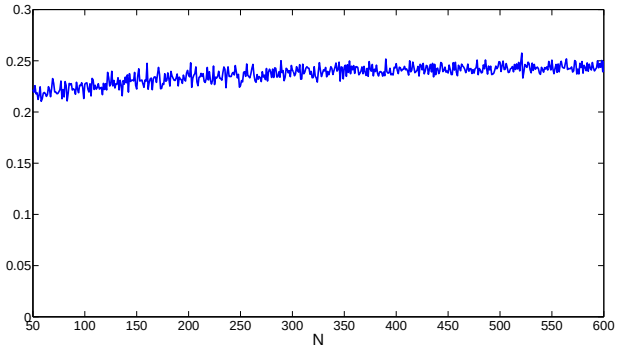

Fig. 7. Behavior of $\sqrt{N}\, J_e(P_N, \sqrt{N})$.

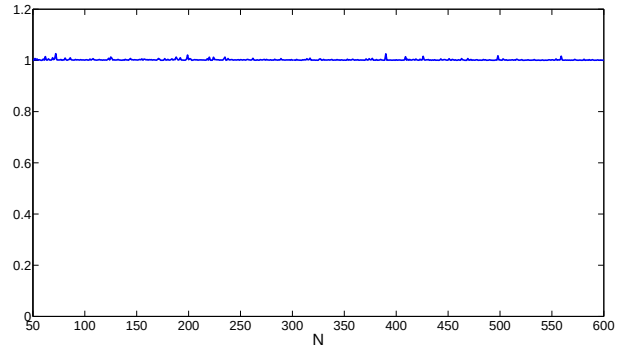

Fig. 8. Behavior of $N\, J_e(P_N, N)$.

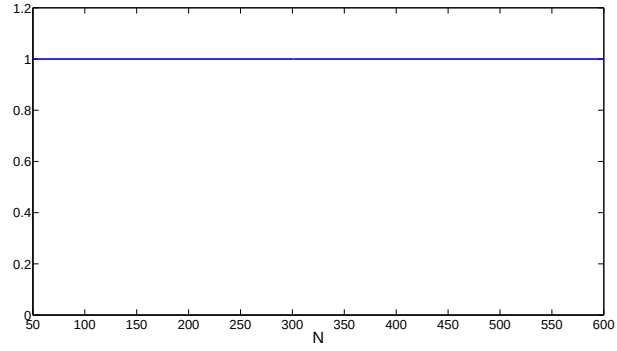

Fig. 9. Behavior of $N\, J_e(P_N, N^{3/2})$.

fact, both for grids on toruses and grids in $\mathbb{R}^d_n$, the essential object in the definition of the eigenvalues is the function $f : \mathbb{R}^n \to [0, +\infty)$ defined by

$$f(\boldsymbol{x}) = |p(e^{ix_1}, \ldots, e^{ix_d})|^2 .$$

Notice that

$$\begin{aligned} f(\boldsymbol{x}) &= p(e^{ix_1}, \ldots, e^{ix_d}) p(e^{-ix_1}, \ldots, e^{-ix_d}) \\ &= \sum_{\boldsymbol{l} \in D_{2\delta}} q_{\boldsymbol{l}} \cos\left(l_1 x_1 + \cdots + l_d x_d\right) \end{aligned}$$

where

$$q_{\boldsymbol{l}} = \sum_{\substack{\boldsymbol{h},\boldsymbol{k}\in D_\delta:\\ \boldsymbol{h}-\boldsymbol{k}=\boldsymbol{l}}} p_{\boldsymbol{h}} p_{\boldsymbol{k}}\,.$$

Clearly f is a trigonometric polynomial, with $f(\mathbf{0}) = 1$ and $0 \le f(\boldsymbol{x}) < 1$ for all $\boldsymbol{x} \in [-\pi,\pi]^d \setminus \{\mathbf{0}\}$. Notice that $\nabla f(\mathbf{0}) = \mathbf{0}$, and the Hessian matrix in $\boldsymbol{x} = 0$, $H_f(\mathbf{0})$, is given by $\frac{\partial^2 f}{\partial x_r \partial x_s} = -\sum_{\boldsymbol{h}\in D_{2\delta}} q_{\boldsymbol{h}} h_r h_s$. It is easy to see that $H_f(\mathbf{0})$ is negative semi-definite. Using the assumption that the infinite Cayley graph on $\mathbb{Z}^d$ associated with $p(z_1,\dots,z_d)$ is connected and has self-loops allows to prove that it is also non-singular (the proof is omitted here).

As a consequence, there exists $a \in (0,\pi)$, $\alpha,\beta > 0$, $c \in (0,1)$ such that, for all $\boldsymbol{x} \in [-\pi,\pi]^d$, $f_L(\boldsymbol{x}) \le f(\boldsymbol{x}) \le f_U(\boldsymbol{x})$ and $g_L(\boldsymbol{x}) \le f(\boldsymbol{x}) \le g_U(\boldsymbol{x})$, where the functions f_L, f_U, g_L, and g_U are defined as:

$$f_L(\boldsymbol{x}) = \begin{cases} 1-\beta\,\boldsymbol{x}^T\boldsymbol{x} & \text{for } x\in(-a,a)^d \\ 0 & \text{otherwise}\,, \end{cases}$$

$$f_U(\boldsymbol{x}) = \begin{cases} 1-\alpha\,\boldsymbol{x}^T\boldsymbol{x} & \text{for } x\in(-a,a)^d \\ c & \text{otherwise}\,, \end{cases}$$

$$g_L(\boldsymbol{x}) = \begin{cases} e^{-\beta\,\boldsymbol{x}^T\boldsymbol{x}} & \text{for } x\in(-a,a)^d \\ 0 & \text{otherwise}\,, \end{cases}$$

$$g_U(\boldsymbol{x}) = \begin{cases} e^{-\alpha\,\boldsymbol{x}^T\boldsymbol{x}} & \text{for } x\in(-a,a)^d \\ c & \text{otherwise}\,. \end{cases}$$

B. LQ cost

We will use the explicit expressions for the eigenvalues given in Sect. III-A and Sect. III-B and the bounds above. Let's first consider a family $\{P_n\}$ of Cayley graphs:

$$\begin{aligned} J_\Delta(P_n) &= \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in V_n\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{1-|\lambda_{\boldsymbol{h}}|^2} \\ &= \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in V_n\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{1-f(e^{i\frac{2\pi}{n}h_1},\dots,e^{i\frac{2\pi}{n}h_d})} \end{aligned}$$

Define $V_n' = \{-\lfloor\frac{n}{2}\rfloor,\dots,0,\dots,+\lfloor\frac{n}{2}\rfloor\}^d$. Clearly, $f(\boldsymbol{x})$ has period 2π in each of its variables, so that

$$\begin{aligned} J_\Delta(P_n) &\le \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in V_n'\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{1-f(e^{i\frac{2\pi}{n}h_1},\dots,e^{i\frac{2\pi}{n}h_d})} \\ &\le \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in V_n'\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{1-f_U(e^{i\frac{2\pi}{n}h_1},\dots,e^{i\frac{2\pi}{n}h_d})} \end{aligned}$$

Thus, if you also define $V_n'' = \{-\lfloor\frac{an}{2\pi}\rfloor,\dots,0,\dots,+\lfloor\frac{an}{2\pi}\rfloor\}^d$, you have:

$$J_\Delta(P_n) \le \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in V_n''\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{\alpha(\frac{2\pi}{n})^2[h_1^2+\dots+h_d^2]} + (2\pi)^d c$$

Now you can conclude using the following Lemma (which will be useful also for the upper bound)

Lemma 1: Given a constant $c \in \mathbb{N}$, there exists constants $K_d, K_d' > 0$ (depending on d and c only) such that

- if $d = 1$,

$$K_1 \le \sum_{\substack{\boldsymbol{h}\in\{-cn,\dots,cn\}^d\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{h_1^2+\dots+h_d^2} \le K_1'$$

- if $d = 2$,

$$K_2\log n \le \sum_{\substack{\boldsymbol{h}\in\{-cn,\dots,cn\}^d\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{h_1^2+\dots+h_d^2} \le K_2'\log n$$

- if $d \ge 3$,

$$K_d n^{d-2} \le \sum_{\substack{\boldsymbol{h}\in\{-cn,\dots,cn\}^d\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{h_1^2+\dots+h_d^2} \le K_d' n^{d-2}$$

□

The proof is omitted.

For the lower bound, the proof is very similar. Define $\tilde{V}_n' = \{-\lfloor\frac{n-1}{2}\rfloor,\dots,0,\dots,+\lfloor\frac{n-1}{2}\rfloor\}^d$, so that

$$\begin{aligned} J_\Delta(P_n) &\ge \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in\tilde{V}_n'\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{1-f(e^{i\frac{2\pi}{n}h_1},\dots,e^{i\frac{2\pi}{n}h_d})} \\ &\ge \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in\tilde{V}_n'\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{1-f_L(e^{i\frac{2\pi}{n}h_1},\dots,e^{i\frac{2\pi}{n}h_d})} \end{aligned}$$

Then, with V_n'' as in the upper bound,

$$J_\Delta(P_n) \ge \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in V_n''\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{\beta(\frac{2\pi}{n})^2[h_1^2+\dots+h_d^2]}$$

Finally, you conclude using Lemma 1

When you consider grids in $\mathbb{R}^d$, the proof is very similar. In this case,

$$J_\Delta(P_n) = \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in V_n\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{1-f(e^{i\frac{\pi}{n}h_1},\dots,e^{i\frac{\pi}{n}h_d})}$$

If you define $\bar{V}_n = \{0,\dots,\lfloor\frac{an}{\pi}\rfloor\}^d$ you get the following bounds.

$$\begin{aligned} J_\Delta(P_n) &\le \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in\bar{V}_n\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{\alpha(\frac{2\pi}{n})^2[h_1^2+\dots+h_d^2]} + \pi^d c \\ &\le \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in V_n''\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{\alpha(\frac{2\pi}{n})^2[h_1^2+\dots+h_d^2]} + \pi^d c \end{aligned}$$

and

$$\begin{aligned} J_\Delta(P_n) &\ge \frac{1}{N}\sum_{\substack{\boldsymbol{h}\in\bar{V}_n\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{\beta(\frac{2\pi}{n})^2[h_1^2+\dots+h_d^2]} \\ &\ge \frac{1}{N}\frac{1}{2^d}\sum_{\substack{\boldsymbol{h}\in V_n''\\ \boldsymbol{h}\neq\mathbf{0}}} \frac{1}{\beta(\frac{2\pi}{n})^2[h_1^2+\dots+h_d^2]} \end{aligned}$$

Then again you conclude using Lemma 1.

C. Quadratic estimation error

This proof is very similar, but uses the bounds g_L, g_U instead of f_L, f_U. We use the same sets of indexes $V_n', V_n'', \bar{V}_n$ defined while proving previous proposition. For the Cayley family $\{P_n\}$,

$$\begin{aligned} J_e(P,t) &= \tfrac{1}{N} \sum_{\boldsymbol{h} \in V_n} [f(\tfrac{2\pi}{n}h_1, \ldots, \tfrac{2\pi}{n}h_d)]^t \\ &\leq \tfrac{1}{N} \sum_{\boldsymbol{h} \in V_n'} [f_U(\tfrac{2\pi}{n}h_1, \ldots, \tfrac{2\pi}{n}h_d)]^t \\ &\leq \tfrac{1}{N} \sum_{\boldsymbol{h} \in V_n''} e^{-\alpha(\frac{2\pi}{n})^2(h_1^2+\cdots+h_d^2)t} + (2\pi)^d c^t \end{aligned}$$

The you conclude using the following lemma.

Lemma 2: For any constants $c \in (0, \frac{1}{2})$, $\gamma > 0$, there exists constants $c_0 = c_0' = 1$, $c_1, \ldots, c_d, c_1', \ldots, c_d' > 0$ (depending on c, γ and d only) such that

$$\tfrac{1}{N} \sum_{\boldsymbol{h} \in \{-\lfloor cn \rfloor, \ldots, \lfloor cn \rfloor\}^d} e^{-\gamma(\frac{2\pi}{n})^2(h_1^2+\cdots+h_d^2)t} \leq \sum_{l=0,\ldots,d} \frac{1}{n^{d-l}} \frac{c_l}{t^{l/2}}$$

and

$$\tfrac{1}{N} \sum_{\boldsymbol{h} \in \{-\lfloor cn \rfloor, \ldots, \lfloor cn \rfloor\}^d} e^{-\gamma(\frac{2\pi}{n})^2(h_1^2+\cdots+h_d^2)t} \geq \max_{l=0,\ldots,d} \frac{1}{n^{d-l}} \frac{c_l'}{t^{l/2}}$$

□

The proof is omitted.

For the lower bound,

$$\begin{aligned} J_\Delta(P_n) &\geq \frac{1}{N} \sum_{\boldsymbol{h} \in \bar{V}_n'} \frac{1}{1 - f_L(e^{i\frac{2\pi}{n}h_1}, \ldots, e^{i\frac{2\pi}{n}h_d})} \\ &\geq \tfrac{1}{N} \sum_{\boldsymbol{h} \in V_n''} e^{-\alpha(\frac{2\pi}{n})^2(h_1^2+\cdots+h_d^2)t} \end{aligned}$$

from which you conclude using Lemma 2.

The proof for the grid matrix family is very similar to the one reported in previous section, using g_L, g_U instead of f_L, f_U and using Lemma 2 instead of Lemma 1.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the European Community's Seventh Framework Programme under agreement n. FP7-ICT-223866-FeedNetBack.

REFERENCES

[1] T. C. Aysal, M. Coates, and M. Rabbat. Distributed average consensus using probabilistic quantization. In *IEEE Workshop on Statistical Signal Processing*, pages 640–644, Maddison, Wisconsin, Aug. 2007.

[2] S. Boyd, P. Diaconis, P. Parrilo, and L. Xiao. Symmetry analysis of reversible markov chains. *Internet Mathematics*, 2:31–71, 2005.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE/ACM Trans. Netw.*, 14(SI):2508–2530, 2006.

[4] F. Bullo, J. Cortes, and S. Martinez. *Distributed Control of Robotic Networks*. Available at http://coordinationbook.info, 2008.

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

[6] J. Cortes, S. Martinez, and F. Bullo. Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions. *IEEE Trans. Automatic Control*, 51(8):1289–1298, 2006.

[7] G. Cybenko. Dynamic load balancing for distributed memory multiprocessors. *J. Parallel Distrib. Comput.*, 7(2):279–301, 1989.

[8] F. Fagnani and S. Zampieri. Average consensus with packet drop communication. 2008. Special issue on "Control and Optimization in Cooperative Networks" Accepted.

[9] F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. *IEEE J. on Selected Areas of Communications*, 26(4):634–649, 2008.

[10] M. Franceschetti and R. Meester. *Random networks for communication*. Cambridge University Press, Cambridge, 2007.

[11] P. Frasca, R. Carli, F. Fagnani, and S. Zampieri. Average consensus on networks with quantized communication. *Intern. Journ. on Non-linear and Robust Control*, to appear, 2008.

[12] S. Ganeriwal, R. Kumar, and M. B. Srivastava. Timing-sync protocol for sensor networks. In *SenSys '03: Proceedings of the 1st international conference on Embedded networked sensor systems*, pages 138–149, New York, NY, USA, 2003. ACM.

[13] E. N. Gilbert. Random plane networks. *Journal of SIAM*, 9:533–543, 1961.

[14] P. Gupta and P. Kumar. The capacity of wireless networks. *Information Theory, IEEE Transactions on*, 46(2):388–404, Mar 2000.

[15] M. Huang and J. H. Manton. Stochastic approximation for consensus seeking: mean square and almost sure convergence. In *Conference on Decision and Control (CDC)*, pages 306–311, New Orleans, USA, Dec. 2007.

[16] A. Kashyap, T. Başar, and R. Srikant. Quantized consensus. *Automatica*, 43(7):1192–1203, 2007.

[17] L. Moreau. Stability of multi-agent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(2):169–182, 2005.

[18] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis. On distributed averaging algorithms and quantization effects. Technical Report 2778, LIDS-MIT, Nov. 2007.

[19] R. Olfati-Saber. Distributed kalman filter with embedded consensus filters. *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pages 8179–8184, Dec. 2005.

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

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

[22] S. Patterson, B. Bamieh, and A. E. Abbadi. Distributed average consensus with stochastic communication failures. In *IEEE Conference on Decision and Control (CDC)*, pages 4215–4220, New Orleans, LA, Dec. 2007.

[23] M. Penrose. *Random Geometric Graphs*. Oxford University Press, 2003.

[24] S. Rai. The spectrum of a random geometric graph is concentrated. *J. Theoret. Probab.*, 20(2):119–132, 2007.

[25] W. Ren and R. W. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control*. Springer, 2007.

[26] J. N. Tsitsiklis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, MIT, Nov. 1984. Technical Report LIDS-TH-1424, Laboratory for Information and Decision Systems.

[27] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Systems and Control letters*, 53(1):65–78, 2004.

[28] L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. *Journal of Parallel and Distributed Computing*, 67(1):33–46, 2007.

[29] L. Xiao, S. Boyd, and S. Lall. A scheme for asynchronous distributed sensor fusion based on average consensus. In *International Conference on Information Processing in Sensor Networks (IPSN'05)*, pages 63–70, Los Angeles, CA, Apr. 2005.