Average consensus on networks with quantized communication

Paolo Frasca, Ruggero Carli, Fabio Fagnani and Sandro Zampieri

Abstract

This work presents a contribution to the solution of the average agreement problem on a network with quantized links. Starting from the well-known linear diffusion algorithm, we propose a simple and effective adaptation which is able to preserve the average of states and to drive the system near to the consensus value, when a uniform quantization is applied to communication between agents. The properties of this algorithm are investigated both by a worst case analysis and by a probabilistic analysis, and are shown to depend on the spectral properties of the evolution matrix. A special attention is devoted to the issue of the dependence of the performance on the number of agents, and several examples are given.

I. INTRODUCTION AND PROBLEM STATEMENT

In the last years, we have witnessed an increasing interest for studying control, estimation and algorithmic problems over networks. A common feature of these problems is the fact that there is a fundamental constraint on the information flow: data are distributed among a large number of nodes communicating among each other through some communication network. A prototype of such problems is the so-called consensus problem.

Suppose we have a directed graph $\mathcal{G}$ with set of nodes $V = \{1, \ldots, N\}$ and a real quantity x_i for every node $i \in V$. The average consensus problem consists of computing the average $x_A = N^{-1} \sum_i x_i$ in an iterative and distributed way, exchanging information among nodes exclusively along the available edges in $\mathcal{G}$. This problem appears in a number of different contexts since the early 80's (decentralized computation [37], load balancing [10], [13], [14]) and, more recently, has attracted much attention for possible applications to sensor networks (data fusion problems [22], [40], [39], clock synchronization [27]) and to coordinated control of mobile autonomous agents [23], [28], [34], [30], [36], [8], [35]. Other places where consensus algorithms have been studied are [4], [3], [29], [15].

Different algorithms for average consensus have been proposed in the literature. They can be distinguished on the basis of the amount of communication and computation they require, of their scalability properties with respect to the number of nodes, of their adaptability to time-varying graphs, and, finally, on the basis of their deterministic or randomized operating protocol.

Deterministic (time-invariant and time-varying) consensus algorithms have been studied in many papers. Starting from the pioneering work [37], many variations can be found in the above cited literature. Most of the papers study the same kind of schemes: every node runs a first order linear dynamical system to update its estimation and the systems are coupled through the available communication edges. The overall system is a linear system whose state updating matrix is a stochastic matrix. The problems typically considered in the literature concern necessary and sufficient conditions for convergence, speed of convergence and optimization issues.

Most of the literature on the consensus problem assumes that the communication channel between the nodes allows to transfer real numbers with no errors. In many practical applications this is not a realistic assumption: if we think, for instance, of sensor networks communicating in a wireless fashion, it is evident that energy and bandwidth limitation yield a finite capacity channel. This suggests that in many situations the communication channel should be rather considered as digital, accepting messages taking values in some finite alphabet. This clearly forces a quantization on the real numbers that agents have to transmit.

The effects of quantization in feedback control problems have been widely studied in the past [31], mainly in the stabilization problem. Moreover granularity effects different from quantization in the consensus problems have been tackled in few papers, especially in the load balancing applications [14], [25]. Our setting is however different: this paper studies the consensus problem under the assumption that communications are quantized, and we assume an ideal noiseless digital channel. The consensus problem under quantization transmission has first appeared in the final section of [39]. In principle, we could simply try to take the usual linear algorithm which yields average consensus and try to apply it with quantization transmission. However, this turns out not to be a good idea since in the new setting the algorithm leads in general to an approximate consensus which may be far though from the average. This was noted in [5], [6], [7] where a new algorithm was proposed and partially analyzed.

Other works where other quantized algorithms have been considered are [25], [32], [1], [24]. In [25], the authors study systems having (and transmitting) integer-valued states and propose a class of gossip[1] algorithms which preserve the average

Fabio Fagnani and Paolo Frasca are with the Dipartimento di Matematica del Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy. {fabio.fagnani,paolo.frasca}@polito.it

Ruggero Carli and Sandro Zampieri are with the DEI, University of Padova, Via Gradenigo 6/a, 35131 Padova, Italy. {carlirug,zampi}@dei.unipd.it

[1]A gossip algorithm is a consensus algorithm in which pairs of agents, randomly selected, communicate. See [3] or [16] for a treatment without quantization, and [25] and [17] for quantized versions.

of states and are guaranteed to converge up to one quantization bin. The recently appeared technical report [32] studies a deterministic quantization scheme, not preserving the average, in which the agents can only store quantized data. Several results are obtained on the error between the convergence value and the initial average, and on the speed of convergence. The work [1] actually deals with quantized communication, and proposes to use a specific randomized quantizer together with a time invariant communication topology. Such a scheme achieves almost surely a consensus at a quantization level, but only the expectation of the average is preserved. Also [24] proposes to use a randomized quantizer, together with time dependent weights.

This paper contains an extended analysis of the algorithm appeared in [5], [6], [7], which uses a deterministic uniform quantization and a fixed time invariant communication topology.

In discrete time, if we have an ideal exchange of information, the typical approach is to set as dynamics the following equations

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

where $x_i(t) \in \mathbb{R}$ is the state of the i-th agent at the time t and P_{ij} are coefficients belonging to a doubly stochastic matrix $P \in \mathbb{R}^{N\times N}$. More compactly we can write

$$x(t+1) = Px(t),$$

where $x(t)$ is the column vector whose entries $x_i(t)$ represent the agents states. In [33], [4] are provided conditions on P, ensuring that the previous algorithm solves the *average consensus problem*, namely

$$\lim_{t\to+\infty} x_i(t) = \frac{1}{N}\left(\sum_{j=1}^{N} x_j(0)\right) \qquad \forall\, i \in V. \tag{2}$$

We will detail them in Section II.

In this work, we assume that the information exchanged between the agents is quantized so that the real numbers $x_i(t)$ can not be exactly transmitted. Information is quantized by a uniform quantizer which is defined as follows.

Let $q : \mathbb{R} \to \mathbb{Z}$ be the mapping sending z to its nearest integer, namely,

$$q(z) = n \in \mathbb{Z} \quad \Leftrightarrow \quad \begin{array}{ll} z \in [n-1/2, n+1/2[, & \text{if } z \geq 0 \\ z \in]n-1/2, n+1/2], & \text{if } z < 0. \end{array} \tag{3}$$

In a first naive approach, one could assume that the agents apply to received data the same linear algorithm as in the ideal case, so that the states evolve following

$$x_i(t+1) = P_{ii}\, x_i(t) + \sum_{j\neq i} P_{ij}\, q(x_j(t)), \tag{4}$$

This dynamical system is much more difficult to study than the original linear one, because quantization brings nonlinearity into the evolution law. Moreover, there is an important shortcoming. Let $x_a(t) = N^{-1}\sum_{i=1}^{N} x_i(t)$ denote the average of states. Then, for the map (4) it's not always true that $x_a(t+1) = x_a(t)$, that is the average of the agents' states is not necessarily preserved. This problem, which has been remarked in [39], is due to the loss of symmetry between neighbors in the use of information. Moreover, the dynamical system (4), besides not preserving the average of states, shows in simulations poor convergence properties, in many cases of practical interest [5].

Because of these facts, we propose a variation of this algorithm which is able to overcome this problem. By remarking that (1) can be rewritten as

$$x_i(t+1) = x_i(t) \; + \sum_{j=1}^{N} P_{ij}\, [x_j(t) - x_i(t)],$$

we propose the evolution scheme

$$x_i(t+1) = x_i(t) \; + \sum_{j=1}^{N} P_{ij}\, [q(x_j(t)) - q(x_i(t))]. \tag{5}$$

We will observe in the next section that the algorithm (5), differently from algorithm (4), preserves the average of the initial conditions, that is $x_a(t) = x_a(0), \;\forall\, t \geq 0$.

Of course also the algorithm (5), because of the quantization effects, is not expected to converge in the sense (2). What we can hope is that agents reach states which are close to each other and close to the average $x_a(0)$. To measure this asymptotic disagreement, we introduce the quantity $\Delta_i(t) := x_i(t) - x_a(t) = x_i(t) - x_a(0)$ which represents the distance, at time t, of the i-th agent from the average of the initials conditions. Defined $\Delta(t) = [\Delta_1(t), \ldots, \Delta_N(t)]^*$, we are interested in evaluating the quantity

$$d_\infty(P) := \sup_{x(0)} \limsup_{t\to\infty} \frac{1}{\sqrt{N}} \|\Delta(t)\|,$$

where $\|\cdot\|$ is the 2-norm. A special attention is given to the dependence of this quantity on the number of the nodes N.

As preliminary examples we consider here two particular communication topologies: the random geometric graph and the hypercube graph.

The random geometric graph is commonly used to model wireless networks [20]. It is constructed by randomly placing N nodes in the unit square, and joining them with edges whenever their distance is below a threshold $R = \Theta(\sqrt{\log N/N})$ for $N \to \infty$. Moreover we assume the weights of the matrix P are chosen by using the maximum-degree weights[2].

The n-dimensional hypercube graph is the graph obtained drawing the edges of a n-dimensional hypercube. It has $N = 2^n$ nodes which can be identified with the binary words of length n. Two nodes are neighbors if the corresponding binary words differ in only one component. Thus every node exchanges information with other n nodes. In this case a matrix P can be constructed by setting equal every non-zero entry of P. We will precise this in Section III.

In Figures 1 and 2 is depicted the behavior of $d_\infty(P)$, which is obtained as the average over several realizations of the initial conditions, for sequences of graphs of both topologies with increasing size. It is worth noting that in both cases, $d_\infty(P)$ appears to be bounded uniformly in N, the size of the communication graph.

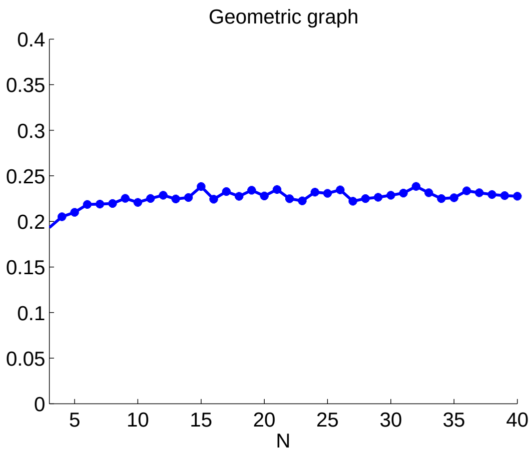

Fig. 1. Performance of the random geometric graph. Since the graph itself is random, the plot of $d_\infty(P)$ comes from averaging over realizations of both the graph and the initial condition.

Clearly $d_\infty(P)$ depends on the particular communication topology we are considering and on the weights assigned to the matrix P. In general, assigned the communication topology, what one would like to do is finding the matrix of weights P minimizing $d_\infty(P)$. In order to do so, it is important to provide some tools which permit to estimate $d_\infty(P)$.

In spite of the fact that the proposed algorithm is still intrinsically linear, the quantization effects introduce nonlinearities which make the exact asymptotic analysis of the algorithm quite hard. Indeed, we are able to carry on such an analysis only in very specific examples.

To overcome this limitations, in this paper we undertake a twofold analysis of the algorithm: worst case and probabilistic. The basic idea is to study the system considering the communication errors induced by quantization as (unknown) bounded disturbances. The worst case analysis is obtained by introducing a bounded error model and maximizing $d_\infty(P)$ with respect to all possible realizations of the communication errors: of course in this way we obtain an upper bound to the performance of our algorithm. This is done in Section III where we prove in general the *convergence to a neighborhood of the average*, obtaining bounds on its size. These bounds are independent of the initial condition but depend on the diffusion matrix. However the worst-case analysis is intrinsically conservative. In fact we will show, for the hypercube example cited above, that the worst-case displacement grows logarithmically in N, behavior which is in disagreement with the experimental evidence displayed by Figure 2.

For this reason, in Section IV we propose an alternative method and develop a probabilistic analysis, modeling the quantization error as additive random noise affecting the received data. A classical mean squared analysis for the asymptotic error can be

[2]According to the definitions given at the end of the introduction, we denote by d_i the in-degree of node i, and by d_M the maximal in-degree of the graph, or an upper bound of it. Then the maximum-degree weights are $P_{ij} = \begin{cases} \frac{1}{d_M+1} & \text{if } (i,j) \in E \\ 1 - \frac{d_i}{d_M} & \text{if i=j} \\ 0 & \text{otherwise.} \end{cases}$ See also [26].

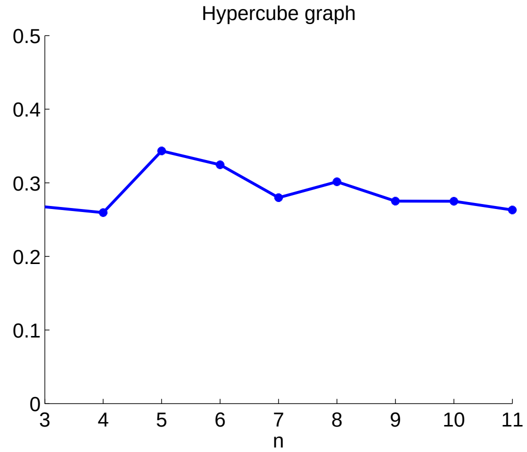

Fig. 2. Performance of the n-dimensional hypercube graph (of order $N = 2^n$).

carried on in this case. A similar analysis is done in [38]. It comes out that, under mild assumptions on the diffusion matrix, the expected behavior depends only on the assumed distribution of the errors and on the spectrum of the evolution matrix. Of course, this probabilistic analysis, in principle, does not offer any rigorous bound on our system. However, simulations clearly show that the probabilistic analysis is very close to the experimental evidence, contrarily to the worst case analysis which instead is quite conservative. Finally in Section V we gather out our conclusions and potential developments of this research.

Mathematical Preliminaries

Before proceeding, we collect some definitions and notations which are used through the paper. The reader can refer to [19] [18] for further readings.

The communications between agents are modeled by a *directed graph* $\mathcal{G} = (V, E)$ where $V = \{1, \ldots, N\}$ is the set of vertices and E is the set of (directed) edges, i.e. a subset of $V \times V$. If $(j, i) \in E$, it means that j can transmit information about its state to i. Any $(i, i) \in E$ is called a *self loop*. The *adjacency matrix* A of $\mathcal{G}$ is a $\{0, 1\}$-valued square matrix indexed by the elements in V defined by letting $A_{ij} = 1$ if and only if $(j, i) \in E$ and $j \neq i$. Define the *in-degree* of a vertex i as $\sum_j A_{ij}$ and the *out-degree* of a vertex j as $\sum_i A_{ij}$. A graph is said to be *undirected* (or *symmetric*) if $(i, j) \in E$ implies that $(j, i) \in E$. A graph is *strongly connected* if for any given pair of vertices (i, j) there exists a path which connects i to j. A *path* in $\mathcal{G}$ consists in a sequence of vertices $(i_1, i_2, \ldots, i_r)$ such that $(i_j, i_{j+1}) \in E$ for every $j \in \{1, \ldots, r-1\}$. A graph is said to be *fully connected* or *complete* if $E = V \times V$.

A matrix M is said to be *nonnegative* if $M_{ij} \geq 0$ for all i and j. A matrix M is said to be *stochastic* if it is nonnegative and the sums along each row are equal to 1. Moreover a matrix M is said to be *doubly stochastic* if it is stochastic and also the sums along each column are equal to 1. Given a nonnegative matrix $M \in \mathbb{R}^{N \times N}$, we can define an 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, M is said to be *adapted* or *compatible* with $\mathcal{G}$ if $\mathcal{G}_M \subset \mathcal{G}$.

An interesting class of graphs are the *Cayley graphs* defined on Abelian groups [2]. Let G be any finite Abelian group of order $|G| = N$, whose operation is denoted by $+$, and let S be a subset of G. The Cayley graph $\mathcal{G}(G, S)$ is the directed graph with vertex set G and arc set

$$E = \{(g, h) \in G \times G : h - g \in S\}.$$

Most examples we are going to deal with belong to this class. If $G = \mathbb{Z}_N$, the cyclic group of order N, then the graph is said to be *circulant*.

A notion of Cayley structure can also be introduced for matrices. Let G be any finite Abelian group of order $|G| = N$. A matrix $M \in \mathbb{R}^{G \times G}$ is said to be a *Cayley* matrix over the group G if

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

This means that a generating vector $\pi \in \mathbb{R}^N$ exists such that $M_{i,j} = \pi(i - j)$. If $G = \mathbb{Z}_N$ then P is said *circulant* [11]. Cayley matrices enjoy many properties: among others, they are *normal*, that is, denoting with the star the conjugate transpose, $M^* M = M M^*$. Moreover, it is possible to write their spectrum explicitly.

Now we give some notational conventions. Given $v \in \mathbb{R}^N$, we denote by $\|v\|$ the Euclidean norm of v and by $\|v\|_\infty$ the sup norm of v. Given a matrix M, we denote by $\|M\|$ its induced Euclidean norm. Given a set A, $|A|$ denotes its cardinality.

Given a matrix $M \in \mathbb{R}^{N\times N}$, let $\sigma(M)$ denote the set of eigenvalues of M and $\rho(M)$ the spectral radius of M: $\rho(M) = \max\{|\lambda| : \lambda \in \sigma(M)\}$. When the matrix is stochastic, it is also worth to define the *essential* spectral radius as

$$\rho_{ess}(M) = \begin{cases} 1 & \text{if } \dim \ker(M-I) > 1 \\ \max\{|\lambda| : \lambda \in \sigma(M) \setminus \{1\}\} & \text{if } \dim \ker(M-I) = 1\,, \end{cases} \tag{6}$$

Finally, for $f, g : \mathbb{N} \to \mathbb{R}$, we say that $f = o(g)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$ and that $f = \Theta(g)$ if there exist $N_0 \in \mathbb{N}$ and k_1, k_2 positive real numbers, such that $k_1|g(n)| \leq |f(n)| \leq k_2|g(n)|$ for all $n \geq N_0$.

II. Algorithms with uniform quantized communications

Assume that we have a set V of N agents 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 state of agent i at time t. In the sequel we study the algorithm (5),

$$x_i(t+1) = x_i(t) + \sum_{j=1}^{N} P_{ij}\left[q(x_j(t)) - q(x_i(t))\right] \tag{7}$$

where P_{ij} are the entries of a doubly stochastic matrix P compatible with $\mathcal{G}$, and where $q(\cdot)$ denotes the quantizer defined in (3). A more general uniform quantizer, with quantization step ϵ, can be defined from q by $q_\epsilon(x) = \epsilon q(x/\epsilon)$. For the sake of notational simplicity we will work only with the case $\epsilon = 1$. The general case can be simply recovered by a suitable scaling. If we have a vector $x \in \mathbb{R}^N$, with a slight abuse of notation, we use the notation $q(x) \in \mathbb{R}^N$ to denote the vector such that $q(x) = [q(x_1), \dots, q(x_N)]^*$. Hence more compactly we can write

$$x(t+1) = x(t) + (P-I)\, q(x(t)). \tag{8}$$

where $x(t)$ is the column vector whose entries $x_i(t)$ represent the agents states. It is well known in the literature [33], [4] that, if P is a doubly stochastic matrix with positive diagonal entries and such that $\mathcal{G}_P$ strongly connected, then the algorithm with exchange of perfect information (1) solves the *average consensus problem*, namely

$$\lim_{t\to+\infty} x(t) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i(0)\right)\mathbf{1}, \tag{9}$$

where $\mathbf{1}$ is the (column) vector of all ones. From now on we assume the following.

Assumption 1. *P is a doubly stochastic matrix such that $P_{ii} > 0$, $i = 1, \dots, N$, and $\mathcal{G}_P$ is strongly connected.*

We collect here some facts about such matrix P, which are consequences of the Perron-Frobenius theorem or straightforward consequences [18], [4].

Lemma 1. *Let P satisfy Assumption 1 and let $Y = I - N^{-1}\mathbf{1}\mathbf{1}^*$. Then the following facts hold.*

- *The eigenvalues of P are bounded in modulus by 1.*
- *There is only one eigenvalue of modulus 1. It is simple (namely it has algebraic multiplicity 1), and it is equal to 1. The corresponding eigenvector is $\mathbf{1}$.*
- $\lim_{t\to\infty} P^t = N^{-1}\mathbf{1}\mathbf{1}^*$.
- *YP has the same spectrum and eigenvectors system as P, except that the eigenvalue 1 is replaced by 0. This implies that*

$$\lim_{t\to\infty} YP^t = 0.$$

- *$PY = YP$ and $Y^2 = Y$.*
- *$\|PY\| = \sqrt{\rho\left((PY)^*PY\right)}$. Since $(PY)^*PY = P^*PY$, $(P^*P)_{ii} > 0$ and $\mathcal{G}_{P^*P}$ is strongly connected. Therefore $\sigma(P^*P) = \{1, \lambda_1, \dots, \lambda_{N-1}\}$, where $|\lambda_i| < 1,\ 1 \leq i \leq N-1$. Observe that $\sigma(P^*PY) = \{\sigma(P^*P) - \{1\}\} \cup \{0\}$. Hence $\|PY\| < 1$.*

As already mentioned, we observe that the algorithm (8) preserves the average of the initial conditions, that is $x_a(t) = x_a(0)$ for all $t \geq 0$. Indeed, since P is doubly stochastic we have that $x_a(t+1) = N^{-1}\mathbf{1}^*x(t+1) = N^{-1}\mathbf{1}^*x(t) + N^{-1}\mathbf{1}^*(P - I)\, q(x(t)) = N^{-1}\mathbf{1}^*x(t) = x_a(t)$.

However, also the algorithm (8), because of the quantization effects, is not expected to converge in the sense (9). What we can hope is for the agents to reach states which are close to each other and close to the average $x_a(0)$. To measure this asymptotic disagreement, we introduce the following quantity $\Delta_i(t) := x_i(t) - x_a(t) = x_i(t) - x_a(0)$ which represents the distance, at time t, of the i-th agent from the average of the initials conditions. Let now $Y = I - N^{-1}\mathbf{1}\mathbf{1}^*$ and $\Delta(t) = [\Delta_1(t), \dots, \Delta_N(t)]^*$. Then, $\Delta(t) = Y\, x(t)$. Finally we define the performance index

$$d_\infty(P, x(0)) = \limsup_{t\to\infty} \frac{1}{\sqrt{N}}||\Delta(t)||. \tag{10}$$

We can get rid of the initial condition by considering $d_\infty(P) = \sup_{x(0)} d(P, x(0))$.

The problem we would like to address is, given a matrix P, to evaluate how big d_∞ is. In particular, we are interested in evaluating how, for sequences of graphs of increasing size, this quantity depends on the number of the nodes N.
We start with a remark about the best achievable performance. It is clear that, when all states lie in the same quantization interval, namely $q(x_i(T)) = Q$ for all i, differences are not perceivable and states do not evolve. Therefore the best the algorithm can assure is that the system reaches such an equilibrium in which $q(x_i(t)) = Q$ for all i and for all $t \geq T$. This implies both that $|\Delta_i(t)| \leq 1$ for all i and for all $t \geq T$, and that $d_\infty \leq 1/2$. This is the best result which can be guaranteed by any algorithm. Obtaining such a performance is not trivial: simulations show that the error from the agreement can actually be bigger. However, it is worth noting that the states of the agents subject to (8) are bounded. In particular one can see, by convexity arguments, that for any node $i \in V$ and for all $t \geq 0$,

$$x_i(t) \in \left[\min_{j\in V}\{q(x_j(0))\} - \frac{1}{2},\ \max_{j\in V}\{q(x_j(0))\} + \frac{1}{2}\right]. \tag{11}$$

Of course (11) is a very weak result since we can hope that in general the disagreement decreases as time goes on and that the asymptotic disagreement does not depend on the initial conditions, but just possibly on P. We consider now two examples in which the evolution of the system can be studied explicitly, the complete graph and the directed circuit graph.

Example 1 (Complete graph)**.** If the communication graph is complete and the communication is exact, i.e. not quantized, the average consensus problem can be solved in one step taking $P = \frac{1}{N}\mathbf{1}\mathbf{1}^*$. We now compute the exact performance degradation due to quantization. The system is in this case

$$x(t+1) = x(t) - q(x(t)) + N^{-1}\mathbf{1}\mathbf{1}^* q(x(t)). \tag{12}$$

We have that, for $t \geq 1$,

$$\begin{aligned}
|\Delta_i(t)| &= \left| x_i(t) - \frac{1}{N}\sum_{j=1}^N x_j(0) \right| = \left| x_i(t) - \frac{1}{N}\sum_{j=1}^N x_j(t-1) \right| \\
&= \left| x_i(t-1) - q\left(x_i(t-1)\right) + \frac{1}{N}\sum_{j=1}^N q\left(x_j(t-1\right) - \frac{1}{N}\sum_{j=1}^N x_j(t-1) \right| \\
&\leq \left| x_i(t-1) - q\left(x_i(t-1)\right)\right| + \frac{1}{N}\sum_{j=1}^N \left| q\left(x_j(t-1\right) - x_j(t-1)\right| \\
&= \frac{1}{2} + \frac{1}{N}\frac{N}{2} = 1.
\end{aligned}$$

This implies that

$$d_\infty(N^{-1}\mathbf{1}\mathbf{1}^*) \leq 1. \tag{13}$$

Example 2 (Directed circuit)**.** Now we consider a more interesting example, the directed circuit graph, which is described by the Cayley graph $\mathcal{G}(\mathbb{Z}_N, \{1\})$. In this case each agent communicates with only one neighbor and evolves following

$$x_i(t+1) = x_i(t) + \frac{1}{2}[q(x_{i+1}(t)) - q(x_i(t))] \quad i = 1, \dots, N, \tag{14}$$

where summation of the indexes is to be intended $\mod N$. The evolution of (14) can be studied exactly by means of a *symbolic dynamics* approach. This analysis is definitively not trivial, but permits us to characterize precisely the evolution of (14) and to obtain a strong result. Indeed, it is possible to show (see Appendix) that there exists $T \in \mathbb{N}$ such that

$$|x_i(t) - x_j(t)| \leq 1 \quad \forall\, i, j \quad \forall\, t > T.$$

It follows that

$$d_\infty(P) \leq 1/2.$$

III. Bounded error model-Worst case analysis

An exact analysis of the dynamics of system (8), as we did in the previous two examples, is not feasible for general graphs. In this section we undertake a worst case analysis which can instead be applied to the general case. We start by observing that (8) can be rewritten in the following way

$$x(t+1) = Px(t) + (P - I)(q(x(t)) - x(t)), \tag{15}$$

where $q(x(t)) - x(t)$ is such that $\|q(x(t)) - x(t)\|_\infty \leq 1/2$. In order to carry out a worst-case analysis of (15), we introduce the following bounded error model

$$\begin{cases} x_w(t+1) = Px_w(t) + (P-I)e(t), & x_w(0) = x(0) \\ \Delta_w(t) = Yx_w(t) \end{cases} \tag{16}$$

where $e(t) \in \mathbb{R}^N$ is such that $\|e(t)\|_\infty \leq 1/2$ for all $t \geq 0$. Notice that in this case $e(t)$ is no more a quantization error, but instead represents an unknown bounded disturbance. Clearly, when $e(t) = q(x(t)) - x(t)$ it turns out that $x_w(t) = x(t)$ and $\Delta_w(t) = \Delta(t)$ for all $t \geq 0$.

We define now a performance index for (16), considering the worst asymptotic disagreement, worst with respect to all the possible choices of the time sequence of the vectors $e(t)$. To be more precise, let us introduce $\mathcal{E}^\infty = \left\{\{e(\cdot)\}_{t=0}^\infty \,|\|e(t)\|_\infty \leq \frac{1}{2},\ \forall\, t \geq 0\right\}$, namely the set of all the sequences of N-dimensional vectors having sup norm less than $1/2$. Then, for the system (16), we define

$$d_\infty^w(P, x_w(0)) = \sup_{\mathcal{E}^\infty} \limsup_{t\to\infty} \frac{1}{\sqrt{N}} ||\Delta_w(t)||. \tag{17}$$

Note that $\lim_{t\to\infty} YP^t = 0$. This implies that the asymptotic behavior of $\Delta_w(t)$ is independent of the initial condition $x_w(0)$ and hence this is the case also for the quantity $d_\infty^w(P, x_w(0))$. Thus, from now on we denote $d_\infty^w(P, x_w(0))$ simply by $d_\infty^w(P)$. As a preliminary remark, note that

$$d_\infty(P) \leq d_\infty^w(P).$$

We start our analysis of $d_\infty^w(P)$ by the following example.

Example 3. In this example we consider the hypercube graph. Precisely, we consider the group $\mathbb{Z}_2^n$ where $2^n = N$ and the Cayley graph $\mathcal{G}(\mathbb{Z}_2^n, S)$, where $S = \{e_0, e_1, \dots, e_n\}$, with $e_0 = [0, \dots, 0]^*$ and e_j the vector with all elements equal to 0 except a 1 in position j if $j \neq 0$. Clearly $|S| = n+1$. We assume that the matrix P has the following structure

$$P_{ij} = \begin{cases} \frac{1}{n+1} \text{ if } i - j = e_h \;\; \exists\, h \,:\; 0 \leq h \leq n, \\ 0 \text{ otherwise} \end{cases} \tag{18}$$

for all i and j belonging to $\mathbb{Z}_2^n$. In other words we have that $P = \frac{1}{n+1}(I + A)$ where A is the adjacency matrix of the Cayley graph $\mathcal{G}(\mathbb{Z}_2^n, S)$. We have the following result whose proof is presented in the Appendix.

Theorem 2. *Let the communication graph be the hypercube graph and let P be as in Example 3. Then*

$$d_\infty^w(P) = \frac{n}{2} = \frac{\log_2 N}{2}.$$

This result shows immediately that there exists a discrepancy between the simulative evidence on d_∞ and the estimate d_∞^w. Indeed, while d_∞ seems to be uniformly bounded on N (see Figure 2), $d_\infty^w \to \infty$ as $N \to \infty$.
Nevertheless, in the literature of the quantized control, the bounded error model is the only model which permits us to infer some theoretical analysis on (8) and thus to provide some bound on d_∞. There is an another model, the probabilistic model that we propose in the next section, that seems to be more in accordance with the experimental results. But the fact that the probabilistic model seems to capture the main features of (8) comes only from the experimental evidence: there is no theoretical justification motivating this agreement. For this reason we proceed with the analysis of d_∞^w.

We start from the following result that provides a general bound for d_∞^w, valid for all schemes achieving average consensus with perfect communication.

Proposition 3. *Let P be a matrix satisfying Assumption 1. Then $\|PY\| < 1$ and*

$$d_\infty^w(P) \leq \frac{1}{1 - \|PY\|}. \tag{19}$$

Proof:

We know that $\|PY\| < 1$ from Lemma 1. Consider now $\Delta_w(t)$. From standard algebraic tools, and using Lemma 1, we have that

$$\begin{aligned} \Delta_w(t) &= YP^t x(0) + Y \sum_{s=0}^{t-1} P^s (I-P) e(t-s-1) \\ &= (PY)^t \Delta(0) + \sum_{s=0}^{t-1} (PY)^s (I-P) e(t-s-1). \end{aligned}$$

Now we have that

$$\begin{aligned}\|\Delta_w(t)\| &= \|(PY)^t\Delta_w(0) + \sum_{s=0}^{t-1}(PY)^s(I-P)e(t-s-1)\| \\ &\leq \|(PY)^t\|\,\|\Delta_w(0)\| + \|I-P\|\sum_{s=0}^{t-1}\|(PY)\|^s\|e(t-s-1)\| \\ &= \|(PY)^t\|\,\|\Delta_w(0)\| + \sqrt{N}\frac{1-\|PY\|^t}{1-\|PY\|}\end{aligned}$$

where in the last inequality we used the facts that $\|I-P\| \leq 2$ and $\|e(t)\| \leq \sqrt{N}/2$ for all $t \geq 0$. By letting $t \to \infty$ we obtain (19). ■

Note that, if P is normal we have that $\|PY\| = \rho_{ess}(P)$ and hence (19) becomes

$$d^w(P) \leq \frac{1}{1-\rho_{ess}(P)}.$$

However, when P is a normal matrix the bound on $d^w_\infty(P)$ can be improved as stated in the next proposition.

Proposition 4. *If P is normal, then*

$$d^w_\infty(P) \leq \frac{1}{2}\sum_{s=0}^{\infty}\rho(P^s(I-P)). \tag{20}$$

Proof: Starting from the expression of $\Delta_w(t)$ provided along the proof of Proposition 3 we can write that

$$\begin{aligned}\|\Delta_w(t)\| &\leq \|(PY)^t\Delta_w(0)\| + \|\sum_{s=0}^{t-1}(PY)^s(I-P)e(t-s-1)\| \\ &\leq \|(PY)^t\Delta_w(0)\| + \frac{\sqrt{N}}{2}\|(PY)^s(I-P)\|\end{aligned}$$

Since P is normal we have that $\|(PY)^s(I-P)\| = \rho((PY)^s(I-P)) = \rho(P^s(I-P))$. By letting $t \to \infty$ in the last inequality, we obtain (20). ■

Remark 1. It is worth noting that, from the sub-multiplicative inequality $\|(PY)^s(I-P)\| \leq \|PY\|^s\,\|I-P\|$, it follows immediately that $\frac{1}{2}\sum_{s=0}^{\infty}\rho(P^s(I-P)) \leq \frac{1}{1-\rho_{ess}(P)}$ which shows that the bound (19) is indeed an improvement of the bound (20).

Example 4 (Complete graph)**.** We recall that, in this case, we set $P = \frac{1}{N}\mathbf{1}\mathbf{1}^*$. Hence $PY = 0$. Thus we have that $\frac{1}{1-\|PY\|} = 1$. This is an alternative way to prove (13). However, since $\left(\frac{1}{N}\mathbf{1}\mathbf{1}^*\right)^k = \frac{1}{N}\mathbf{1}\mathbf{1}^*$ for all $k > 0$, it follows immediately that $\frac{1}{2}\sum_{s=0}^{\infty}\rho(P^s(I-P)) = \frac{1}{2}$ and this represents a refinement of (13). We conclude that for the complete graph, with the above choice of weights, $d_\infty(P) \leq \frac{1}{2}$.

In general it is quite hard to evaluate (20). We provide two results which permit us to approximate (20) under some mild assumptions. First a notational definition. Given $c \in \mathbb{C}$ and $r \in \mathbb{R}$ such that $r \geq 0$, we denote

$$B_{c,r} := \{z \in \mathbb{C} \,|\, \|z-c\| \leq r\},$$

namely the closed ball of complex numbers of radius r and centered in c.

The first result is the most general, since it holds for normal matrices. Its proof is deferred to the Appendix.

Proposition 5. *Let P be a normal matrix satisfying the Assumption 1. Let R be such that $0 < R < 1$ and $\sigma(P) \subseteq B_{1-R,R}$ and let $\bar{\rho} = \rho_{ess}(P)$ denote the essential spectral radius of P. Then*

$$\sum_{s=0}^{\infty}\rho(P^s(I-P)) \leq \frac{1}{1-R} + \sqrt{\frac{8R}{(1-R)(1-\bar{\rho})}}. \tag{21}$$

Remark that the given bound depends on R and $\bar{\rho}$ only, which are functions of the spectrum of P. Let us illustrate it with an example.

Example 5 (Directed circuit)**.** Consider the directed circuit graph introduced in Example 2 and the evolution law given by (14). In this case P is the circulant matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ \frac{1}{2} & 0 & 0 & 0 & \cdots & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

We have theoretically proved (see Corollary 13 in Appendix) that $d_\infty \leq \frac{1}{2}$. Instead we could not evaluate d_∞^w in this case. Consider the bound introduced in Proposition 3. Since $\rho_{ess} = 1 - \frac{\pi^2}{2}\frac{1}{N^2} + o\left(\frac{1}{N^2}\right)$ (see [4]) and since each circulant matrix is a normal matrix we have that $\frac{1}{1-\|PY\|} = \frac{1}{1-\rho_{ess}(P)} = \Theta\left(N^2\right)$. Observe now that all the eigenvalues of P are inside the ball $B_{\frac{1}{2},\frac{1}{2}}$. Hence we obtain that $\frac{1}{1-R} + \sqrt{\frac{8R}{(1-R)(1-\bar{\rho})}} = \Theta(N)$. This means that the bound (21) improves the bound proposed in (19). Moreover, by numerical experiments one can see that $\frac{1}{2}\sum_{s=0}^{\infty} \rho(P^s(I-P)) = \Theta(N)$ meaning, that for $N \to \infty$, (21) behaves as (20): in this sense the bound is tight.

If P is symmetric we can provide a stronger result, whose proof is given in the Appendix.

Proposition 6. *Let P be a symmetric stochastic matrix satisfying Assumption 1. Let R be such that $0 < R < 1$ and $\sigma(P) \subseteq B_{1-R,R}$ and let $\bar{\rho} = \rho_{ess}(P)$ denote the essential spectral radius of P. Then,*

$$\sum_{s=0}^{+\infty} \rho\left(P^s(I-P)\right) \leq \frac{3}{2} + \frac{1}{1-R} + \frac{1}{2}\log\left(\frac{1}{1-\bar{\rho}}\right). \tag{22}$$

The significance of the result can be seen with an example.

Example 6 (Undirected circuit)**.** In this example we consider the circulant matrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & 0 & \cdots & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & \cdots & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ \frac{1}{3} & 0 & 0 & 0 & \cdots & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Simulations (analogous to Figure 1) suggest that d_∞ is uniformly bounded on N. Also in this case we are not able to evaluate d_∞^w. By considering the bound (19), since $\rho_{ess}(P) \cong 1 - \frac{4}{3}\frac{\pi^2}{N^2}$ (see [4]) we obtain that $\frac{1}{1-\|PY\|} = \frac{1}{1-\rho_{ess}} = \Theta(N^2)$. Observe that all the eigenvalues of P are greater than $-\frac{1}{3}$. Hence it results, letting $R = 2/3$, that $\frac{3}{2} + \frac{1}{1-R} + \frac{1}{2}\log\left(\frac{1}{1-\bar{\rho}}\right) = \Theta(\log N)$. Moreover, numerically it is possible to observe that also $\frac{1}{2}\sum_{s=0}^{\infty} \rho(P^s(I-P))$ grows logarithmically, meaning that asymptotically in N, we have that (22) behaves as (20).

Remark 2. It is worth noting that this improvement is more general. Consider a sequence of symmetric Cayley matrices P, having the elements on the diagonal uniformly lower bounded on N and supported on Cayley graphs having in-degree uniformly upper bounded on N. Then the above arguments can be applied to argue that, for this class of graphs, we can not obtain from (20) a bound for $d_\infty^w(P)$ stronger than a logarithmic dependence on N. In the next example we show that a logarithmic bound can be proved to be tight, that is $d_\infty^w(P) = \Theta(\log N)$.

Example 7 (Hypercube)**.** Consider the hypercube graph and the matrix P compatible with the hypercube graph as defined in Example 3. We have already seen, that, by simulations, $d_\infty(P)$ seems to be a uniformly bounded quantity on N, while we have analytically proved that $d_\infty^w = \frac{\log_2 N}{2}$ (see Example 3). It is possible to see that also $\frac{1}{2}\sum_{s=0}^{\infty} \rho(P^s(I-P)) = \frac{\log_2 N}{2}$. This fact is stated in Lemma 14 (see Appendix). Moreover, since the eigenvalues of P are $1 - \frac{2k}{n+1}$ for $k = 0, \ldots, n$ we have that all the eigenvalues of P different from 1 are inside the interval $\left[\frac{1-n}{n+1}, \frac{n-1}{n+1}\right]$. This implies that $\frac{3}{2} + \frac{1}{1-R} + \frac{1}{2}\log\left(\frac{1}{1-\bar{\rho}}\right) = \Theta(\log_2(N))$.

IV. Probabilistic model

In the previous section we have shown that the bounded error model does not seem to really capture the behavior of the quantized model. In particular the upper bound to the performance we have found seems to be quite conservative. In this section we undertake a probabilistic approach, modeling the quantization error as a random variable. We carry on a classical mean square analysis and we show that it gets quite close to simulations of the real quantized model. This suggests that the probabilistic approach is more appropriate to describe quantization errors even though we have no theoretical evidence of this fact: the probabilistic model, differently from the worst case analysis, does not provide any bound to the performance.

For $i \in V$ and $t \in \mathbb{N}$, let $n_i(t)$ be random variables of zero mean and variance σ^2, which have their supports inside $[-1/2, 1/2]$ and are uncorrelated and identically distributed in both i and t, i.e., $\mathbb{E}\left[n_i(t)n_j(\tau)\right] = 0$ if $i \neq j$ or $t \neq \tau$. Define $n(t)$ as the random vector whose components are $n_i(t)$ and consider the stochastic model

$$\begin{cases} x_r(t+1) = Px_r(t) + (P-I)n(t), \qquad x_r(0) = x(0) \\ \Delta_r(t) = Yx_r(t). \end{cases} \tag{23}$$

We define

$$d_\infty^r(P, x_r(0)) = \limsup_{t\to\infty} \sqrt{\frac{1}{N}\mathbb{E}[\|\Delta_r(t)\|^2]}.$$

Like for the model in Section III, we have that $d_\infty^r(P, x_r(0))$ is independent of the initial condition $x_r(0)$. Hence, in the sequel we denote $d_\infty^r(P, x_r(0))$ with the symbol $d_\infty^r(P)$. It is worth to point out that this index, as the previously defined $d_\infty^w(P)$, expresses the asymptotic error induced by quantization, but gives no information about diverse issues like speed of convergence or finite time behavior.

We start our analysis of the probabilistic model with the following result.

Theorem 7. *Let P be a matrix satisfying Assumption 1. Then*

$$[d_\infty^r(P)]^2 = \frac{\sigma^2}{N}\mathrm{tr}\,\left[(I-P)(I-\tilde{P}\tilde{P}^*)^{-1}(I-P)^*\right].$$

where $\tilde{P} = PY$. In particular, if P is normal, and $\sigma(P) = \{1, \lambda_1, \ldots, \lambda_{N-1}\}$ denotes the spectrum of P, we have that

$$[d_\infty^r(P)]^2 = \frac{\sigma^2}{N}\sum_{i=1}^{N-1}\frac{|1-\lambda_i|^2}{1-|\lambda_i|^2}. \tag{24}$$

Proof: Define $Q(t) = \mathbb{E}[\Delta_r(t)\Delta_r(t)^*]$, and remark that $\frac{1}{N}\mathbb{E}[||\Delta_r(t)||^2] = \frac{1}{N}\mathrm{tr}\; Q(t)$. Using the facts that $Y^k = Y$ for all positive integer k and $Y(P-I) = P-I$, it is easy to see that Δ_r satisfies the following recursive equation

$$\Delta_r(t+1) = \tilde{P}\Delta_r(t) + (P-I)n(t).$$

Now, thanks to the hypotheses on $n_i(t)$,

$$\begin{aligned} Q(t+1) &= \mathbb{E}[\Delta_r(t+1)\Delta_r(t+1)^*] \\ &= \mathbb{E}[\tilde{P}\Delta_r(t)\Delta_r(t)^*\tilde{P}^*] + (I-P)\mathbb{E}[n(t)n(t)^*](I-P)^* \\ &= \tilde{P}\,Q(t)\,\tilde{P}^* + \sigma^2(I-P)(I-P)^*, \end{aligned}$$

and by a simple recursion

$$Q(t) = \tilde{P}^t\,Q(0)\,(\tilde{P}^*)^t + \sigma^2\sum_{s=0}^{t-1}\tilde{P}^s(I-P)(I-P)^*(\tilde{P}^*)^s.$$

Since P satisfies Assumption 1, from Lemma 1, $\rho_{ess}(P^*P) < 1$. Moreover we have that $\rho(\tilde{P}) = \rho_{ess}(P) < 1$ and $\rho(\tilde{P}^*\tilde{P}) = \rho_{ess}(P^*P) < 1$. Using the linearity and cyclic property of the trace,

$$\begin{aligned} \mathrm{tr}\; Q(t) &= \mathrm{tr}\,\left[\tilde{P}^t\,Q(0)\,(\tilde{P}^*)^t\right] + \mathrm{tr}\,\left[\sigma^2\sum_{s=0}^{t-1}(PP^* - (P+P^*) + I)(\tilde{P}\tilde{P}^*)^s\right] \\ &= \mathrm{tr}\,\left[\tilde{P}^t\,Q(0)\,(\tilde{P}^*)^t\right] + \sigma^2\mathrm{tr}\,\left[(PP^* - (P+P^*) + I)(I - (\tilde{P}\tilde{P}^*)^t)(I-\tilde{P}\tilde{P}^*)^{-1}\right]. \end{aligned}$$

and hence

$$\lim_{t\to\infty}\mathrm{tr}\; Q(t) = \sigma^2\mathrm{tr}\,\left[(PP^* - (P+P^*) + I)(I-\tilde{P}\tilde{P}^*)^{-1}\right].$$

If moreover P is normal, we can find a unitary matrix O of eigenvectors and a diagonal matrix of eigenvalues Λ, such that $P = O\Lambda O^*$. This implies

$$\mathrm{tr}\,\left[(PP^* - (P+P^*) + I)(1-\tilde{P}\tilde{P}^*)^{-1}\right] = \sum_{i=1}^{N-1}\frac{|1-\lambda_i|^2}{1-|\lambda_i|^2}.$$

∎

From now on we restrict to the case in which P is normal. Note that the expression for the mean square error of formula (24) is the product of two terms, $d_\infty^r(P) = \sigma^2\Phi(P)$ where

$$\Phi(P) := \frac{1}{N}\sum_{i=1}^{N-1}\frac{|1-\lambda_i|^2}{1-|\lambda_i|^2},$$

is a functional[3] of the matrix P, depending only on its spectral structure.

As in the previous section, we are mainly interested in sequences of matrices of increasing size. We will see that the above functional scales well with N in the following examples.

Example 8 (Complete graph)**.** In this case it is easy to compute $\Phi(P)$. We have that $\Phi(P) = \frac{N-1}{N}$.

Example 9 (Directed circuit)**.** Consider the matrix P defines in Example 5. In this case $\Phi(P)$ can be exactly computed. We have

$$\begin{aligned}\Phi(P_N) &= \frac{1}{N}\sum_{h=1}^{N-1}\frac{(1-\lambda_h)(1-\lambda_h^*)}{1-\lambda_h\lambda_h^*}\\ &= \frac{1}{N}\sum_{h=1}^{N-1}\frac{(1-(1/2+1/2\,e^{i\frac{2\pi}{N}h}))(1-(1/2+1/2\,e^{-i\frac{2\pi}{N}h}))}{1-(1/2+1/2\,e^{i\frac{2\pi}{N}h})(1/2+1/2\,e^{-i\frac{2\pi}{N}h})}\\ &= \frac{1}{N}\sum_{h=1}^{N-1}\frac{1/2(1-\cos(\frac{2\pi}{N}h))}{1/2(1-\cos(\frac{2\pi}{N}h))} = \frac{N-1}{N}.\end{aligned}$$

Example 10 (Undirected graph)**.** Consider the undirected undirected circuit graph and the matrix P introduced in Example 6. The eigenvalues of P are

$$\lambda_h = \frac{1}{3}+\frac{2}{3}\cos\left(\frac{2\pi}{N}h\right) \quad h = 0,\dots,N-1,$$

and we have

$$\Phi(P_N) = \frac{1}{N}\sum_{h=1}^{N-1}\frac{1-\lambda_h}{1+\lambda_h} = \frac{1}{N}\sum_{h=1}^{N-1}\frac{1-\cos\left(\frac{2\pi}{N}h\right)}{2+\cos\left(\frac{2\pi}{N}h\right)}.$$

In this case it is difficult to work out the computation explicitly. However, it is possible to compute the limit for $N\to\infty$, since the summation can be interpreted as Riemann sum relative to the function $f(x) = \dfrac{1-\cos(x)}{2+\cos(x)}$. We thus obtain

$$\lim_{N\to\infty}\Phi(P_N) = \frac{1}{2\pi}\int_0^{2\pi}\frac{1-\cos(x)}{2+\cos(x)}dx = \sqrt{3}-1. \tag{25}$$

Example 11 (Hypercube)**.** Consider the hypercube graph and the matrix P defined in (18). We have that

$$\begin{aligned}\Phi(P) &= \frac{1}{N}\sum_{i=1}^{N-1}\frac{|1-\lambda_i|^2}{1-|\lambda_i|^2} = \frac{1}{2^n}\sum_{k=1}^{n}\frac{\left(\frac{2k}{n+1}\right)^2}{1-\left(\frac{n+1-2k}{n+1}\right)^2}\binom{n}{k} =\\ &= \frac{1}{2^n}\sum_{k=1}^{n}\frac{k}{n+1-k}\binom{n}{k} = \frac{1}{2^n}\sum_{k=1}^{n}\binom{n}{k-1} = \frac{2^n-1}{2^n} = \frac{N-1}{N}.\end{aligned}$$

Then

$$\Phi(P) = \frac{N-1}{N}.$$

While in the previous section the hypercube provided the negative example for the worst case behavior, the probabilistic analysis is in agreement with the evidence showed in the simulations. This highlights the differences between the bounded error model and the probabilistic model: with the same assumptions on P the two worst-case analysis and the mean-square analysis give different results.

Example 12 (Random geometric graph)**.** For the random geometric graph, as defined in the introduction, we have no explicit formula for the eigenvalues. However, d_∞^r can be numerically evaluated, and compared with d_∞. The results are shown in Figure 3 where d_∞^r seems to describe well the qualitative behavior of d_∞.

In all the above examples we have that $\Phi(P)$ scales well with N. We provide now some general conditions ensuring that the functional P for a given sequence of matrices of increasing size is uniformly bounded on N.

We start by observing that, if $\rho_{ess}(P) \le B < 1$ then $\Phi(P) \le \frac{N-1}{N}\frac{4}{1-B^2}$. This implies that, given a sequence of matrices of increasing size, if the essential spectral radius of the sequence is uniformly bounded away from 1 then the functional cost Φ is uniformly bounded in N. This fact is true also for d_∞^w as can be easily seen by recalling the expression of the bound provided

[3]Remarkably, the functional $\Phi(P)$ also arises, with a rather different meaning, in [12], as a cost functional describing the transient of the diffusion methods for average consensus over graphs with ideal communication.

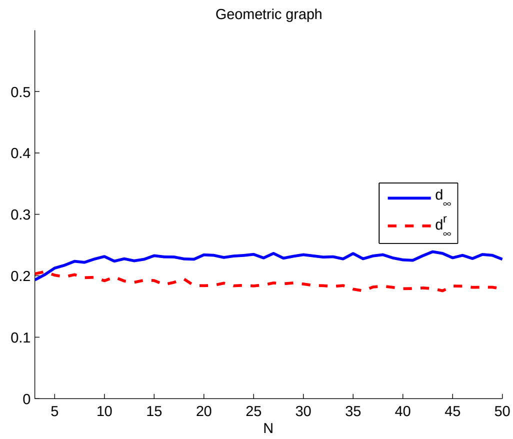

Fig. 3. d_∞^r and d_∞ for random geometric graphs, averaged over 25 realizations of the initial condition and 18 of the graphs, for each N. The plot assumes $\sigma^2 = 1/12$.

by Proposition 3. The interesting fact is that the performance index d_∞^r can exhibit the same behavior *even when the essential spectral radius is not bounded away from* 1, as the following proposition shows.

Proposition 8. *Let $B_{c,r} \subset \mathbb{C}$ denote the closed ball of complex numbers with center the point of coordinates $(c, 0)$ and radius r. If there exists $0 < R < 1$ such that $\sigma(P) \subseteq B_{1-R,R}$ then*

$$\Phi(P) \leq \frac{R}{1-R}. \tag{26}$$

Proof: The inequality $0 < R < 1$ is clear from Assumption 1. It means that the spectrum is contained in a disc of radius R internally tangent in 1 to the unit disc of the complex plane.

Then, we need to prove (26). For all i, the eigenvalue $\lambda_i \in B_{1-R,R}$, so

$$\lambda_i = (1-r) + r\, e^{i\theta}$$

with $\theta \in [0, 2\pi[$ and $0 \leq r \leq R$. Moreover, if $i \geq 1$, then $\theta > 0$. Hence,

$$\begin{aligned} \frac{|1-\lambda_i|^2}{1-|\lambda_i|^2} &= \frac{|r - re^{i\theta}|^2}{1 - |1 - r + r\, e^{i\theta}|^2} = \\ &= \frac{r^2 |1 - e^{i\theta}|^2}{1 - (1-r)^2 - 2r(1-r)\cos\theta - r^2} = \\ &= \frac{r^2 2(1-\cos\theta)}{2r(1-r)(1-\cos\theta)} = \frac{r}{1-r} \leq \frac{R}{1-R} \quad \forall i. \end{aligned}$$

This yields the result. ■

Note that the above bound depends only on R and does not depend on the essential spectral radius of the matrix P, while the worst case bounds provided in Proposition 21 and in Proposition 6 did. This is why (26) is bounded in all the cases we considered.

Many of them are covered by the following corollary.

Corollary 9. *Let $p = \min_i P_{ii}$ and R as above. Then, $R \leq 1 - p$, and*

$$\Phi(P) \leq \frac{1-p}{p}. \tag{27}$$

Proof: By Gershgorin theorem, $\sigma(P) \subseteq \bigcup_i B_{P_{ii}, 1-P_{ii}} \subseteq B_{p,1-p}$. ■

The above result has the following interpretation. If in a family of matrices P_N we have that $\min_i (P_N)_{ii}$ is lower bounded uniformly in N, then (27) gives a finite bound, uniform in N, on the asymptotic displacement. This is a useful hint to construct sequences of matrices whose performance scales well with N. It would be enough to prescribe that the agents assign a minimum weight to their own values.

V. Our contribution and future work

In this work we studied the effects of a uniform quantization on the average consensus problem, and, starting from the well-known linear diffusion algorithm, we proposed a simple and effective adaptation which is able to preserve the average of states and to drive the system reasonably near to the consensus.

The exact asymptotic analysis of the truly nonlinear quantized system has been performed in one special case, while general convergence results have been obtained by a worst case analysis and by a probabilistic model. An interesting open question is how and why the results of the probabilistic analysis are closer to experimental results.

A special attention has been given to the scalability in N of the performance, which is a crucial issue for applications in which the number of agents is huge. In this direction we obtained several favorable results, which we applied to sequences of Cayley graphs. In particular, we showed how the convergence properties of the algorithm depend on the spectrum of a matrix defining the system: this gives deep theoretical insights, and also leads to useful design suggestions, as in Corollary 9.

Potential further developments include the extension to random geometric graphs of the analysis we performed on Cayley examples, and the application of the scheme (8) to time varying topologies. One important case are randomly varying communication topologies, like the gossip algorithm discussed in [17].

References

[1] Tuncer C. Aysal, Mark Coates, and Michael Rabbat. Distributed average consensus using probabilistic quantization. *Statistical Signal Processing, 2007. SSP '07. IEEE/SP 14th Workshop on*, pages 640–644, 26-29 Aug. 2007.

[2] László Babai. Spectra of Cayley graphs. *J. Combin. Theory Ser. B*, 27(2):180–189, 1979.

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

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

[5] Ruggero Carli, Fabio Fagnani, Paolo Frasca, Thomas Taylor, and Sandro Zampieri. Average consensus on networks with transmission noise or quantization. In *ECC '07, Kos*, pages 1852–1857, 2007.

[6] Ruggero Carli, Fabio Fagnani, Paolo Frasca, and Sandro Zampieri. A probabilistic analysis of the average consensus algorithm with quantized communication. In *IFAC 08, Seoul*, 2008.

[7] Ruggero Carli, Fabio Fagnani, Paolo Frasca, and Sandro Zampieri. The quantization error in the average consensus problem. In *Mediterranean Conference on Control and Automation*, July 2008.

[8] Jorge Cortés, Sonia Martínez, and Francesco Bullo. Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions. *IEEE Trans. Automat. Control*, 51(8):1289–1298, 2006.

[9] D. Cvetovic, M. Doob, and H. Sachs. *Spectra of graphs*. Academic Press, 1980.

[10] G. Cybenko. Dynamic load balancing for distributed memory multiprocessors. *Journal of parallel and distributed computing*, 7:279–301, 1989.

[11] P. J. Davis. *Circulant matrices*. A Wiley-Interscience Publication, Pure and Applied Mathematics. John Wiley & Sons, New York-Chichester-Brisbane, 1979.

[12] J. C. Delvenne, R. Carli, and S. Zampieri. Fast strategies in the consensus problem. In *NeCST 2007, Nancy*, 2007.

[13] R. Diekmann, A. Frommer, and B. Monien. Efficient schemes for nearest neighbor load balancing. *Parallel computing*, 25:789–812, 1999.

[14] Robert Elsässer, Burkhard Monien, and Stefan Schamberger. Distributing unit size workload packages in heterogeneous networks. *J. Graph Algorithms Appl.*, 10(1):51–68 (electronic), 2006.

[15] F. Fagnani and S. Zampieri. Average consensus with packet drop communication. *SIAM J. Control Optim.*, 2006. to appear.

[16] Fabio Fagnani and Sandro Zampieri. Randomized consensus algorithms over large scale networks. *IEEE Journal on Selected Areas in Communications*, 26(4):634–649, May 2008.

[17] Paolo Frasca, Ruggero Carli, Fabio Fagnani, and Sandro Zampieri. Average consensus by gossip algorithms with quantized communication. In *Conf. Decision and Control*, Dec 2008. to be presented.

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

[19] Chris Godsil and Gordon Royle. *Algebraic graph theory*, volume 207 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2001.

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

[21] Frank Harary, John P. Hayes, and Horng-Jyh Wu. A survey of the theory of hypercube graphs. *Comput. Math. Applic.*, 15(4):277–289, 1988.

[22] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion: a scalable and robust communication paradigm for sensor networks. In *Proceedings ACM/IEEE Conf. Mobile Computing and Networking*, pages 56–67, 2000.

[23] 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.

[24] Soummya Kar and Jose M. F. Moura. Distributed consensus algorithms in sensor networks: Quantized data. submitted, Dec 2007.

[25] A. Kashyap, T. Basar, and R Srikant. Quantized consensus. *Automatica*, 43:1192–1203, 2007.

[26] S. Lall L. Xiao, S. Boyd. A scheme for robust distributed sensor fusion based on average consensus. In *Proc. Int. Conf. Information Processing in Sensor Networks*, 2005.

[27] Q. Li and D. Rus. Global clock syncronization in sensor networks. *IEEE Transaction on Computers*, 55(2):214– 226, 2006.

[28] J. Lin, A. S. Morse, and B. D. O. Anderson. The multi-agent rendezvous problem: an extended summary. In V. Kumar, N. E. Leonard, and A. S. Morse, editors, *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.

[29] C.C. Moallemi and B. Van Roy. Consensus propagation. *IEEE Transactions on Information Theory*, 52(11):4753–4766, 2006.

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

[31] G.N. Nair, F. Fagnani, S. Zampieri, and R.J. Evans. Feedback control under data rate constraints: An overview. *Proceedings of the IEEE*, 95(1):108–137, Jan. 2007.

[32] Angelia Nedic, Alex Olshevsky, Asuman Ozdaglar, and John N. Tsitsiklis. On distributed averaging algorithms and quantization effects. Technical Report 2778, LIDS-MIT, Nov. 2007.

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

[34] 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.

[35] Wei Ren and Ella Atkins. Distributed multi-vehicle coordinated control via local information exchange. *Internat. J. Robust Nonlinear Control*, 17(10-11):1002–1033, 2007.

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

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

[38] 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.
[39] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *IPSN'05, Los Angeles, pp. 63-70*, 2005.
[40] J. Zhao, R. Govindan, and D. Estrin. Computing aggregates for monitoring wireless sensor networks. In *Proc. Int. Workshop on Sensor Net Protocols and Applications*, pages 139–148, 2003.

APPENDIX
THEORETICAL ANALYSIS OF THE EXAMPLE 2

Here we construct and study the symbolic dynamics underlying the system (14). To start, we need the following technical lemma. Let $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote the floor and ceiling operators from $\mathbb{R}$ to $\mathbb{Z}$.

Lemma 10. *Given $\alpha, \beta \in \mathbb{N}$ and $x \in \mathbb{R}$, it holds*

$$\lfloor x \rfloor = \left\lfloor \frac{\lfloor \alpha x \rfloor}{\alpha} \right\rfloor \tag{28}$$

$$q(x) = \lfloor x + 1/2 \rfloor = \left\lceil \frac{1}{2} \left\lfloor \frac{\lfloor 2\beta x \rfloor}{\beta} \right\rfloor \right\rceil . \tag{29}$$

Proof: We first prove (28). Let $m = \lfloor x \rfloor$. So

$$m \leq x < m + 1$$
$$\alpha m \leq \alpha x < \alpha m + \alpha .$$

Hence, we can find $s \in \mathbb{N}, 0 \leq s \leq \alpha - 1$ such that $\alpha m + s \leq \alpha x < \alpha m + s + 1$. This yields $\lfloor \alpha x \rfloor = \alpha m + s$ and $\left\lfloor \frac{\lfloor \alpha x \rfloor}{\alpha} \right\rfloor = m$.

Then we prove equation (29). The equality $q(x) = \lfloor x + 1/2 \rfloor$ is clear from the definition of $q(x)$. To prove the second equality, let $h = \lfloor 2x \rfloor$. Then $h \leq 2x < h + 1$, from which follows that

$$\frac{h}{2} + \frac{1}{2} = \frac{h+1}{2} \leq x + 1/2 < \frac{h+2}{2} = \frac{h}{2} + 1.$$

From this inequality it follows that $\lfloor x + 1/2 \rfloor = \lceil \frac{h}{2} \rceil$. This, with (28), implies (29).

■

We define $n_i(t) = \lfloor 2x_i(t) \rfloor$. Simple properties of floor and ceiling operators, together with the above lemma, allow us to remark that $q(x_i(t)) = \left\lceil \frac{n_i(t)}{2} \right\rceil$ and to derive from (14) that

$$\lfloor 2x_i(t+1) \rfloor = \lfloor 2x_i(t) \rfloor + q(x_{i+1}(t)) - q(x_i(t))$$
$$n_i(t+1) = n_i(t) + \left\lceil \frac{n_{i+1}(t)}{2} \right\rceil - \left\lceil \frac{n_i(t)}{2} \right\rceil$$
$$n_i(t+1) = \left\lfloor \frac{n_i(t)}{2} \right\rfloor + \left\lceil \frac{n_{i+1}(t)}{2} \right\rceil .$$

We have thus found an iterative system involving only the symbolic signals $n_i(t)$:

$$n_i(t+1) = g(n_i(t), n_{i+1}(t)) \tag{30}$$
$$\text{where } g(h,k) = \left\lfloor \frac{h}{2} \right\rfloor + \left\lceil \frac{k}{2} \right\rceil .$$

The asymptotic analysis of (30) will then allow us to obtain information about the asymptotic of $x_i(t)$ (since $n_i(t) = \lfloor 2x_i(t) \rfloor$) up to quantization errors equal to 1.

We now start the analysis of system (30). Define the following quantities: $m(t) = \min_{1 \leq i \leq N} n_i(t)$, $M(t) = \max_{1 \leq i \leq N} n_i(t)$, $D(t) = M(t) - m(t)$. From the form of (30) one can easily remark that $m(t)$ can not decrease and $M(t)$ can not increase. Hence $D(t)$ is not increasing. A much stronger result about the monotonicity of $D(t)$ is the content of the following lemma.

Lemma 11. *If $D(t_0) \geq 2$, there exists $T \in \mathbb{N}$ such that $D(t_0 + T) < D(t_0)$.*

Proof: Let $I_m(t) = \{j \in \mathbb{Z}_N \text{ s.t. } n_j(t) = m(t)\}$. The idea of the proof is to show that the set $I_m(t)$ eventually decreases if we are in the range $D(t) \geq 2$.

Notice first that, for $h, k \in \mathbb{Z}$, $g(h+2, k+2) = g(h,k) + 2$. Hence, by an appropriate translation of the initial condition, we can always restrict ourselves to the case $m(0) \in \{0, 1\}$.

Case $m(t_0) = 0$. Notice that

$$g(h,k) > 0 \ \forall h \geq 0,\ k > 0, \quad g(h,0) > 0 \ \forall h \geq 2 . \tag{31}$$

This easily implies that $I_m(t)$ is not increasing. Now, since $D(t_0) \geq 2$, we can find $j_0 \in I_m(t_0)$ and two integers $U > 0$ and $W \geq 0$ such that

$$\begin{aligned}
&n_{j_0-W-1}(t_0) > 1\\
&n_{j_0-v}(t_0) = 1 \quad 0 < v \leq W\\
&n_{j_0+u}(t_0) = 0 \quad 0 \leq u < U\\
&n_{j_0+U}(t_0) > 0.
\end{aligned}$$

After W instants step, we then obtain,

$$\begin{aligned}
&n_{j_0-W-1}(t_0+W) > 1\\
&n_{j_0-u}(t_0+W) = 0 \quad W-U+1 \leq u \leq W\\
&n_{j_0+W-U}(t_0+W) > 0.
\end{aligned}$$

At the following time step, one 0 will then disappear

$$\begin{aligned}
&n_{j_0-W-1}(t_0+W+1) > 0\\
&n_{j_0-u}(t_0+W+1) = 0 \quad W-U+2 \leq u \leq W\\
&n_{j_0+W-U-1}(t_0+W+1) > 0.
\end{aligned}$$

This implies that $|I_m(t_0+W+1)| < |I_m(t_0)|$.

Case $m(t_0) = 1$. Notice that

$$g(h,k) > 1 \;\forall h \geq 2,\; k \geq 1, \quad g(1,k) > 1 \;\forall k \geq 3\,. \tag{32}$$

This easily implies that $I_m(t)$ is again not increasing. Now, since $D(t_0) \geq 2$, we can find $j_0 \in I_m(t_0)$ and an integer $W \geq 0$ such that

$$\begin{aligned}
&n_{j_0}(t_0) = 1\\
&n_{j_0+w}(t_0) = 2 \quad 1 \leq w \leq W\\
&n_{j_0+W+1}(t_0) > 2.
\end{aligned}$$

The evolution of the above configuration yields, after W instant steps

$$\begin{aligned}
&n_{j_0}(t_0+W) = 1\\
&n_{j_0+1}(t_0+W) > 2.
\end{aligned}$$

The next step, we obtain $n_{j_0}(t_0+W+1) > 1$. Therefore, $|I_m(t_0+W+1)| < |I_m(t_0)|$.

In both cases we have proven that $|I_m(t)|$ strictly decreases in finite number of steps. A straightforward induction principle then implies that a finite $T \in \mathbb{N}$ exists such that $m(t_0+T) > m(t_0)$. This proves the result. ∎

The interesting consequence of this lemma is the following result.

Theorem 12. *There exist $T \in \mathbb{N}$ and $h \in \mathbb{Z}$ such that, for all $t > T$, $D(t) < 2$ and, moreover, one of the following condition holds*

1) $n_i(t) = h$, $\forall i$;
2) $\{n_i(t) \,:\, i = 0, \dots N\} = \{h, h+1\}$ *and each* $n_i(t)$ *is constant in time;*
3) $\{n_i(t) \,:\, i = 0, \dots N\} = \{h, h+1\}$ *and each* $n_i(t)$ *is periodic in time of period* N.

Proof: From Lemma 11, it follows that a finite $T \in \mathbb{N}$ can be found, such that $D(t) < 2$ for all $t > T$. Once we reach his condition, there are two possibilities: either the $n_i(t)$ are all equal or they differ by 1. In the first case, the system remains constant (condition 1.). In the second case, it follows from the way g is defined that if the lowest state is odd, the evolution is constant (condition 2.), while if the lowest one is even, the state evolution is a leftward shift (condition 3.). Such evolution is periodic of period N (and possibly of some divisor of N). ∎

We can now go back to the original system

Corollary 13. *For system (14), there exists $T \in \mathbb{N}$ such that*

$$|x_i(t) - x_j(t)| \leq 1 \quad \forall\, i, j \quad \forall\, t > T, \tag{33}$$

and hence $d_\infty(P) \leq 1/2$.

Proof: Immediate consequence of Theorem 12, considering the relation $n_i(t) = \lfloor 2x_i(t) \rfloor$. ∎

Remark that (33) is satisfied while the system has reached a non consensus fixed point or a periodic point.

PROOF OF THEOREM 2

Consider the hypercube graph and the matrix P as defined in Example 3.
First we give the following preliminary result.

Lemma 14. *Let P be as above. Then*

$$\sum_{s=0}^{\infty} \rho(P^s(I-P)) = n = \log_2 N. \tag{34}$$

Proof: The eigenvalues of P are $\lambda_k = 1 - \frac{2k}{n+1} \quad k = 0 \dots n$, with multiplicities $p_k = \binom{n}{k}$ (see [9]). Then,

$$\sum_{s=0}^{\infty} \rho(P^s(I-P)) = \sum_{s=0}^{\infty} \rho_{ess}(P^s)\rho(I-P) = \sum_{s=0}^{\infty} (1 - \frac{2}{n+1})^s (2 - \frac{2}{n+1}) = n.$$

■

We are able now to provide the proof of Theorem 2.

Proof: First we rewrite the expression of $d_{\infty}^{w}(P)$. Since P is symmetric, it is diagonalizable by an orthogonal matrix. We can write that $P = \sum_{h=0}^{N-1} \lambda_h q_h q_h^*$ where q_h are orthonormal. These facts are true also for $P^s(I-P)$. Moreover we have that $\rho\left(P^s(I-P)\right) = ||P^s(I-P)||$. Let

$$\Delta_w^{(f)}(t) := \sum_{s=0}^{t-1} P^s(I-P)e(t-s-1).$$

Then,

$$\begin{aligned} \|\Delta_w^{(f)}(t)\|^2 &= \| \sum_{s=0}^{t-1} P^s(I-P)e(t-s-1)\|^2 = \| \sum_{s=0}^{t-1}\sum_{h=0}^{N-1} \lambda_h^s (1-\lambda_h) q_h q_h^* e(t-s-1)\|^2 \\ &= \| \sum_{h=0}^{N-1} q_h (1-\lambda_h) \sum_{s=0}^{t-1} \lambda_h^s q_h^* e(t-s-1)\|^2 = \sum_{h=0}^{N-1} \left[(1-\lambda_h) \sum_{s=0}^{t-1} \lambda_h^s q_h^* e(t-s-1) \right]^2. \end{aligned}$$

Hence, $[d_{\infty}^{w}(P)]^2 = \max_{\mathcal{E}^\infty} \limsup_{t\to\infty} \frac{1}{N} \|\Delta_w^{(f)}(t)\|^2$.

Now we start using combinatorial tools. Indeed the vertices of the hypercube, as well as the eigenvalues and eigenvectors of P, can be indexed by the subsets of $\{1 \dots, n\}$ (see [21]). With this indexing, for each $I \subseteq \{1 \dots n\}$ the corresponding eigenvalue is $\lambda_I = 1 - \frac{2|I|}{n+1}$ and the eigenvector is the 2^n-dimensional vector $q^{(I)}$, such that its $J-$th component is equal to $q_J^{(I)} = 2^{-n/2}(-1)^{|I\cap J|}$.

Let T be any positive integer and consider the sequence of vectors $e(0), e(1), \dots, e(t), \dots$ such that J-th component of the vector $e(t)$ is equal to $\frac{1}{2}(-1)^{T-1-r}(-1)^{|J|}$, where r is the remainder in the Euclidean division of t over T. Observe that $e(t+T) = e(t)$ for all $t \geq 0$. Observe, moreover, that $e(t)$ is an eigenvector of P corresponding to the eigenvalue $\frac{1-n}{1+n}$ for all $t \geq 0$. Hence we have that

$$\begin{aligned} \frac{1}{N}\|\Delta_w^{(f)}(T)\|^2 &= \frac{1}{N} \sum_{h=0}^{N-1} \left[(1-\lambda_h) \sum_{s=0}^{T-1} \lambda_h^s q_h^* e(T-s-1) \right]^2 = \\ &= \frac{1}{2^n} \left[(1 - \frac{1-n}{n+1}) \sum_{s=0}^{T-1} \left(\frac{1-n}{n+1} \right)^s 2^{-\frac{n}{2}} \sum_{J \subseteq \{1,\dots,n\}} (-1)^{|J|} \frac{1}{2} (-1)^s (-1)^{|J|} \right]^2 \\ &= \frac{1}{4^n} \left[\frac{n}{n+1} \sum_{s=0}^{T-1} \left(\frac{n-1}{n+1} \right)^s 2^n \right]^2 \\ &= \frac{n^2}{(n+1)^2} \left[\frac{1 - \left(\frac{n-1}{n+1}\right)^T}{1 - \frac{n-1}{n+1}} \right]^2 = \frac{n^2}{4} \left[1 - \left(\frac{n-1}{n+1} \right)^T \right]^2. \end{aligned}$$

Assume now that T is an even positive integer. By recalling that $e(t+T) = e(t)$ for all $t \geq 0$, for $t = kT$ where $k \in \mathbb{N}$ it turns out that

$$\frac{1}{N}\|\Delta_w^{(f)}(kT)\|^2 = \frac{n^2}{4}\left[1-\left(\frac{n-1}{n+1}\right)^T\right]^2 \sum_{u=0}^{k-1}\left(\frac{1-n}{n+1}\right)^{uT} =$$

$$= \frac{n^2}{4}\left[1-\left(\frac{n-1}{n+1}\right)^T\right]^2 \left[\frac{1-\left(\frac{n-1}{n+1}\right)^{kT}}{1-\left(\frac{n-1}{n+1}\right)^T}\right]^2 = \frac{n^2}{4}\left[1-\left(\frac{n-1}{n+1}\right)^{kT}\right]^2.$$

Letting $k \to \infty$ we obtain that, for the particular sequence considered

$$\lim_{k\to\infty} \frac{1}{N}\|\Delta_w^{(f)}(kT)\|^2 = \frac{n^2}{4}$$

Therefore we have proved that

$$\limsup_{t\to\infty} \frac{1}{N}\|\Delta_w^{(f)}(kT)\|^2 \geq \frac{n^2}{4}$$

and hence $[d_\infty^w(P)]^2 \geq \frac{n^2}{4}$. Now, Lemma 14 implies that $[d_\infty^w(P)]^2 \leq \frac{n^2}{4}$, and then the claim follows. ■

PROOF OF PROPOSITION 5

Proof: We want to upper bound $\rho(P^s(I-P)) = \max_{k=1}^{N-1} |\lambda_k^s(1-\lambda_k)|$. In order to do so we consider the function $f : \mathbb{C} \to \mathbb{R}$ defined as $f(z) = z^s(1-z)$. Let us consider the closed balls $B_{1-R,R}$ and $B_{0,\bar{\rho}}$. By Gershgorin's Theorem we have that $\sigma(P) \subseteq B_{1-R,R}$. By the definition of essential spectral radius it holds that $\sigma(P)\setminus\{1\} \subseteq B_{0,\bar{\rho}}$. Hence $\sigma(P)\setminus\{1\} \subseteq B_{0,\bar{\rho}} \cap B_{1-R,R}$. Let $A := B_{1-R,R} \cap B_{0,\bar{\rho}}$. Clearly

$$\max_{k=1}^{N-1} |\lambda_k^s(1-\lambda_k)| \leq \max_{z\,\in\, A} |f(z)|.$$

Since f is an analytic function and A is a compact set, from the Maximum Modulus Principle it follows that

$$\max_{k=1}^{N-1} |\lambda_k^s(1-\lambda_k)| \leq \max_{z\,\in\,\partial A} |f(z)|,$$

where ∂A denotes the boundary of A.
Consider now the curves $\gamma : [0, 2\pi] \to \mathbb{C}$,

$$\gamma(t) = 1 - R + Re^{jt},$$

and $\theta : [0, 2\pi] \to \mathbb{C}$,

$$\theta(t) = \bar{\rho}e^{jt},$$

which represent, respectively, the boundaries of $B_{1-R,R}$ and of $B_{0,\bar{\rho}}$. In the following, since $|f(z)| = |f(z^*)|$, will consider γ and θ only on the interval $[0, \pi]$.
By calculating the intersection between γ and θ one can see that $\partial A = \tilde{\gamma} \cup \tilde{\theta}$ where

$$\tilde{\gamma} := \left\{ z = z_x + iz_y \,\in\, \gamma \,:\, z_x \leq \frac{1-2R+\bar{\rho}^2}{2(1-R)} \right\}$$

and

$$\tilde{\theta} := \left\{ z = z_x + iz_y \,\in\, \theta \,:\, z_x \geq \frac{1-2R+\bar{\rho}^2}{2(1-R)} \right\}.$$

We consider now $|f(z)|$ on $\tilde{\gamma}$. By straightforward calculations one can show that

$$|f(\gamma(t))|^2 = 2R^2(1-\cos t)\left[1 - 2R(1-R)(1-\cos t)\right]^s.$$

Now let $x = R\cos t + 1 - R$. In order to analyze the behavior of $|f(\gamma(t))|^2$ we introduce the following auxiliary function

$$F(x) = 2R(1-x)\left[1 - 2(1-R)(1-x)\right]^s.$$

A straightforward calculation shows that studying $|f(z)|^2$ on $\tilde{\gamma}$ is equivalent to study F on $\left[1-2R, \frac{1-2R+\bar{\rho}^2}{2(1-R)}\right]$. By taking the first derivative of F we obtain

$$\frac{\partial F}{\partial x} = 2R\left[1 - 2(1-R)(1-x)\right]^{s-1}\left[-1 + 2(1-R)(s+1)(1-x)\right]$$

We have that $\frac{\partial F}{\partial x} = 0$ for $x := x_1 = 1 - \frac{1}{2(1-R)}$ and $x := x_2 = 1 - \frac{1}{2(1-R)(s+1)}$. Note that $1 - \frac{1}{2(1-R)} \leq 1 - 2R$ for all $R > 0$. Moreover note that F is monotone increasing in $[x_1, x_2]$ and monotone decreasing for $[x_2, +\infty)$. Hence F reached its maximum value inside the interval $\left[1-2R, \frac{1-2R+\bar{\rho}^2}{2(1-R)}\right]$ on $1-2R$ if $x_2 \leq 1-2R$, on x_2 if $1-2R \leq x_2 \leq \frac{1-2R+\bar{\rho}^2}{2(1-R)}$, on $\frac{1-\bar{\rho}^2}{2R(1-R)}$ if $x_2 \geq \frac{1-\bar{\rho}^2}{2R(1-R)}$. We have that $x_2 \leq 1-2R \Leftrightarrow s \leq \frac{(1-2R)^2}{4R(1-R)}$, $1-2R \leq x_2 \leq \frac{1-2R+\bar{\rho}^2}{2(1-R)} \Leftrightarrow \frac{(1-2R)^2}{4R(1-R)} < s < \frac{\bar{\rho}^2}{1-\bar{\rho}^2}$, $x_2 \geq \frac{1-\bar{\rho}^2}{2R(1-R)} \Leftrightarrow s \geq \frac{\bar{\rho}^2}{1-\bar{\rho}^2}$. Let $\bar{s} = \lfloor \frac{(1-2R)^2}{4R(1-R)} \rfloor$ and $s^* = \lfloor \frac{\bar{\rho}^2}{1-\bar{\rho}^2} \rfloor$. Therefore

$$\max_{1-2R \leq x \leq \frac{1-2R+\bar{\rho}^2}{2(1-R)}} F(x) = \begin{cases} 4R^2\left(1-2R\right)^{2s} & \text{if} \quad s \leq \bar{s} \\ \frac{R}{1-R}\frac{s^s}{(s+1)^{s+1}} & \text{if} \quad \bar{s}+1 \leq s \leq s^* \\ \frac{R}{1-R}\bar{\rho}^{2s}\left(1-\bar{\rho}^2\right) & \text{if} \quad s \geq s^*+1. \end{cases}$$

Consider now $|f(z)|^2$ on $\tilde{\theta}$. By simple algebraic manipulations one can see that

$$|f(\theta(t))|^2 = \bar{\rho}^{2s}\left(1+\bar{\rho}^2-2\bar{\rho}\cos t\right)$$

Note that $|f(\theta(t))|^2$ is monotone increasing for $t \in [0,\pi]$ and hence it reaches its maximum on $\tilde{\theta}$ when $\cos t = \frac{1-2R+\bar{\rho}^2}{2(1-R)}$ Therefore we can conclude that

$$\max_{k=1}^{N-1} |\lambda_k^s(1-\lambda_k)| \leq \begin{cases} 2R\left|1-2R\right|^s & \text{if} \quad s \leq \bar{s} \\ \sqrt{\frac{R}{1-R}\frac{s^s}{(s+1)^{s+1}}} & \text{if} \quad \bar{s}+1 \leq s \leq s^* \\ \sqrt{\frac{R}{1-R}\bar{\rho}^{2s}\left(1-\bar{\rho}^2\right)} & \text{if} \quad s \geq s^*+1. \end{cases}$$

Hence we can write

$$\sum_{s=0}^{t-1}\rho(P^s(I-P)) \leq \sum_{s=0}^{\bar{s}} 2R\left|1-2R\right|^s + \sum_{s=\bar{s}+1}^{s^*}\sqrt{\frac{R}{1-R}\frac{s^s}{(s+1)^{s+1}}} + \sum_{s=s^*+1}^{t-1}\sqrt{\frac{R}{1-R}\bar{\rho}^{2s}\left(1-\bar{\rho}^2\right)}.$$

Notice now that

$$\sum_{s=0}^{\bar{s}} 2R\left|1-2R\right|^s \leq \frac{2R}{1-|1-2R|} \leq \frac{1}{1-R}.$$

and that

$$\sum_{s=s^*+1}^{t-1}\sqrt{\frac{R}{1-R}\bar{\rho}^{2s}\left(1-\bar{\rho}^2\right)} \leq \sqrt{\frac{R}{1-R}}\sqrt{1-\bar{\rho}^2}\sum_{s=0}^{\infty}\bar{\rho}^s$$
$$= \sqrt{\frac{R(1-\bar{\rho}^2)}{(1-R)(1-\bar{\rho})^2}} \leq \sqrt{\frac{2R}{(1-R)(1-\bar{\rho})}}.$$

Notice finally that, since $\sum_{i=1}^m \sqrt{\frac{1}{i+1}} \leq 2\sqrt{m+1}$, we can argue that

$$\sum_{s=\bar{s}+1}^{s^*}\sqrt{\frac{R}{1-R}\frac{s^s}{(s+1)^{s+1}}} \leq \sqrt{\frac{R}{1-R}}\sum_{s=1}^{s^*}\sqrt{\frac{1}{2}}\sqrt{\frac{1}{1+s}}$$
$$\leq \sqrt{\frac{4R}{2(1-R)(1-\bar{\rho}^2)}} \leq \sqrt{\frac{2R}{(1-R)(1-\bar{\rho})}}.$$

Putting together these three inequalities we obtain (21). ∎

PROOF OF PROPOSITION 6

Proof: Assume that $\sigma(P) = \{\lambda_0 = 1, \lambda_1, \ldots, \lambda_{N-1}\}$. Note that $\sigma(P) \setminus \{1\} \subseteq [1-2R, \bar{\rho}]$. We want to upper bound $\rho(P^s(I-P)) = \max_{k=1}^{N-1}|\lambda_k^s(1-\lambda_k)|$. To do this, consider the function $f(x) = |x^s(1-x)|$. It is continuous in $[-1,1]$, positive and decreasing in $[-1,0]$, it vanishes in $x=0$ and in $x=1$ and has a local maximum in $x = x_M = \frac{s}{1+s}$, with $f(x_M) = (\frac{s}{1+s})^s(\frac{1}{1+s})$. We need to evaluate $\max_{x\in[1-2R,\bar{\rho}]} f(x)$.

First observe that there exists $\bar{s}$, only depending on the value $1-2R$, such that for all $s > \bar{s}$ we have $f(1-2R) < (\frac{s}{1+s})^s(\frac{1}{1+s})$ and then the global maximum of $f(x)$ is assumed at $x = x_M$ if $x_M \leq \bar{\rho}$. Since x_M tends to 1 as s goes to infinity, it happens that, when $s > \frac{\bar{\rho}}{1-\bar{\rho}}$, $x_M(s) > \bar{\rho}$. In conclusion we have that

$$\max_{x\in[1-2R,\bar{\rho}]} f(x) = \begin{cases} 2R|1-2R|^s, & \text{if } 0 \leq s \leq \bar{s}; \\ (\frac{s}{1+s})^s(\frac{1}{1+s}), & \text{if } \bar{s} < s \leq s^*; \\ \bar{\rho}^s(1-\bar{\rho}), & \text{if } s^* < s < \infty. \end{cases}$$

where $s^* := \lfloor \frac{\bar{\rho}}{1-\bar{\rho}} \rfloor$. Hence we can write

$$\sum_{s=0}^{t-1} \rho(P^s(I-P)) \leq \sum_{s=0}^{\bar{s}} 2R|1-2R|^s + \sum_{s=\bar{s}+1}^{s^*} \left(\frac{s}{1+s}\right)^s \left(\frac{1}{1+s}\right) + \sum_{s=s^*+1}^{t-1} \bar{\rho}^s(1-\bar{\rho}).$$

Notice now that

$$\sum_{s=0}^{\bar{s}} 2R|1-2R|^s \leq \frac{1}{1-R},$$

and that

$$\sum_{s=s^*+1}^{t-1} \bar{\rho}^s(1-\bar{\rho}) = \bar{\rho}^{s^*+1} - \bar{\rho}^t \leq 1.$$

Notice finally that, since $\sum_{i=1}^{m} 1/i \leq 1 + \ln m$ we can argue that

$$\begin{aligned}
\sum_{s=\bar{s}+1}^{s^*} \left(\frac{s}{1+s}\right)^s \left(\frac{1}{1+s}\right) &\leq \sum_{s=0}^{s^*} \left(\frac{1}{2}\right) \left(\frac{1}{1+s}\right) \leq \frac{1}{2}(1+\log(s^*+1)) \\
&\leq \frac{1}{2} + \frac{1}{2}\log\left(\frac{\bar{\rho}}{1-\bar{\rho}} + 1\right) = \frac{1}{2} + \frac{1}{2}\log\left(\frac{1}{1-\bar{\rho}}\right).
\end{aligned}$$

Putting together these three inequalities we obtain (22). ■