Geometric Bounds for Convergence Rates of Averaging Algorithms

Bernadette Charron-Bost

Geometric Bounds for Convergence Rates of Averaging Algorithms

Bernadette Charron-Bost

CNRS, École polytechnique, 91128 Palaiseau, France

July 10, 2020

Abstract

We develop a generic method for bounding the convergence rate of an averaging algorithm running in a multi-agent system with a time-varying network, where the associated stochastic matrices have a time-independent Perron vector. This method provides bounds on convergence rates that unify and refine most of the previously known bounds. They depend on geometric parameters of the dynamic communication graph such as the normalized diameter or the bottleneck measure.

As corollaries of these geometric bounds, we show that the convergence rate of the Metropolis algorithm in a system of n agents is less than $1-1/4n^2$ with any communication graph that may vary in time, but is permanently connected and bidirectional. We prove a similar upper bound for the EqualNeighbor algorithm under the additional assumptions that the number of neighbors of each agent is constant and that the communication graph is not too irregular. Moreover our bounds offer improved convergence rates for several averaging algorithms and specific families of communication graphs.

Finally we extend our methodology to a time-varying Perron vector and show how convergence times may dramatically degrade with even limited variations of Perron vectors.

1 Introduction

Motivated by the applications of the Internet and the development of mobile devices with communication capabilities, the design of distributed algorithms for networks with a swarm of agents and time-varying connectivity has been the subject of much recent work. The algorithms implemented in such dynamic networks ought to be decentralized, using local information, and resilient to mobility and link failures while remaining efficient.

One of the basic problems arising in multi-agent networked systems is an agreement problem, called *asymptotic consensus*, or just *consensus*, in which agents are required to compute values that become infinitely close to each other. For example, in clock synchronization, agents attempt to maintain a common time scale; or sensors may try to agree on estimates of a certain variable; or vehicles may attempt to align their direction of motions with their neighbors in coordination of UAV's and control formation.

1.1 Network model and averaging algorithms

Let us consider a fixed set of agents that operate synchronously and communicate by exchanging values over an underlying time-varying communication network. In the consensus problem, the objective is to design distributed algorithms in which the agents start with different initial values and reach agreement on one value that lies in the range of the initial values. The term of *constrained consensus* is used when the goal is to compute a specific value in this range (e.g., the average of the initial values).

Natural candidates for solving the consensus problem are the *averaging algorithms* in which each agent maintains a scalar variable that it repeatedly updates to a convex combination of its own value and of the values it has just received from its neighbors. The weights used by an agent can only depend on local informations available to this agent. The matrix formed with the weights at each time step of an averaging algorithm is a stochastic matrix, and the graph associated to the stochastic matrix coincides with the communication graph. Hence, in the discrete-time model, every execution of an averaging algorithm determines a sequence of stochastic matrices.

Every averaging algorithm corresponds to a specific rule for computing the weights. Three averaging algorithms are of particular interest, namely the *EqualNeighbor* algorithm with weights equal to the inverse of the degrees in the communication graph, its space-symmetric version called *Metropolis*, and the *FixedWeight* algorithm which is a time-uniformization of the EqualNeighbor algorithm in the sense that each agent uses some bound on its degree instead of its (possibly time-varying) degree. A specific feature of the Metropolis algorithm is to address the constrained consensus problem with convergence on the average of the initial values.

The convergence of averaging algorithms has been proved under various assumptions on the connectivity of the communication graph, in particular when it is time-varying but permanently connected [13, 2]. The goal in this paper is to establish novel and tight bounds on the convergence rates of averaging algorithms that depend on geometric parameters of the communication graph. As demonstrated in the simple case of a fixed communication graph and fixed weights, the convergence rate involves the second largest singular values of the corresponding stochastic matrices. Thus a primary step is to develop geometric bounds of these singular values and to get some control on the successive associated eigenspaces.

1.2 Contribution

In this paper, our first contribution concerns upper bounds on the second largest eigenvalue of a reversible stochastic matrix. We start with an analytic bound and then develop a geometric bound. This second bound compares well with previous geometric bounds derived through Cheeger-like inequalities or Poincaré inequalities, and is often much easier to compute. We derive geometric bounds on the second largest singular value of a reversible stochastic matrix. We also obtain an analytic bound on the second largest singular value that is weaker, but still holds when the matrix is not reversible.

Our second contribution is a generic method for bounding the convergence rate of an execution of an averaging algorithm when the associated stochastic matrices have all the same Perron vector. Combined with the above bounds on the second largest singular value of stochastic matrices, this method provides bounds on convergence rates that unify and refine most of the previously known bounds. Basically, the approach consists in masking time fluctuations of the network topology by a constant Perron vector. Two typical examples implementing this strategy for coping with time-varying topologies are the *Metropolis* algorithm and the *FixedWeight* algorithm. Using the geometric bounds developed herein, our method offers improved convergence rates of these algorithms for large classes of communication graphs.

We show that for any time-varying topology that is permanently connected and bidirectional, the convergence rate of the Metropolis algorithm is at most $1 - 1/4n^2$, where n is the number of agents. As a byproduct, we obtain that the second largest eigenvalue of the random walk on a connected regular bidirectional graph is in $1-O(n^{-2})$. A similar result holds for the EqualNeighbor algorithm with limited degree fluctuations over both time and space: the convergence rate is less than $1 - 1/\left(3 + d_{\max} - d_{\min}\right) n^2$ if each agent has a constant number of neighbors in the range $[d_{\min}, d_{\max}]$. These two quadratic bounds exemplify the performance of the Poincaré inequality developed by Diaconis and Stroock [7].

Finally, we extend our methodology to a time-varying Perron vector: we provide a heuristic analysis of the convergence rates of averaging algorithms that demonstrates how time-fluctuations of Perron vectors may lead to exponential degradation of convergence times. Our approach consists in replacing the Euclidean norm associated to the Perron vector by the generic semi-norm $\mathrm{N}(x) = \max(x_i) - \min(x_i)$ defined on $\mathbb{R}^n$, which does not depend on Perron vectors anymore.

Related work. Several geometric bounds on the second largest eigenvalue and the second largest singular value of a reversible stochastic matrix have been previously developed (e.g., see [22, 21, 7, 12]). Our geometric bound expressed in terms of the *normalized diameter* of the associated graph is novel to the best of our knowledge. The analytic bound presented in this paper is a generalization of the bound developed by Nedić et al. for doubly stochastic matrices [16].

Concerning the convergence rate of averaging algorithms, there is also considerable literature. Let us cite the bound established by Xiao and Boyd for the Metropolis algorithm on a fixed topology [23], the one developed by Cucker and Smale for modelling formation of flocks in a complete graph [5], the bound by Olshevsky and Tsitsiklis which concerns the EqualNeighbor algorithm with fixed degrees [19, 20], the analytic bound developed by Nedić et al. [16] in the case of doubly stochastic matrices (and hence, with the typical application to the Metropolis algorithm), and the one developed by Chazelle [4] for the FixedWeight algorithm.

From the quadratic bound on the hitting time of Metropolis walks established by Nonaka et al. [17], Olshevsky [18] deduced that the convergence rate of the *Lazy Metropolis* algorithm in any system of n agents connected by a fixed bidirectional communication graph is less than $1-1/71n^2$. Our general quadratic bound for the Metropolis algorithm is obtained with a different approach based on the discrete analog of the Poincaré inequality developed by Diaconis and Strook [7]. Applied to Lazy Metropolis, our approach gives the improved bound of $1 - 1/8n^2$. It also proves that the quadratic time complexity result in [18] extends to the case of time-varying topologies.

The case of time-varying Perron vectors is addressed by Nedić and Liu [14] with a different method than ours: instead of dealing with the sequence of Perron vectors and using the non-Euclidean norm N, they consider the *absolute probability sequence* associated with the sequence of stochastic matrices [10] and the sequence of associated Euclidean norms.

2 Preliminaries on stochastic matrices

2.1 Notation

Let n be a positive integer and let $[n] = \{1, \dots, n\}$. For every positive probability vector $\pi \in \mathbb{R}^n$, we define

$$< x, y >_\pi = \sum_{i \in [n]} \pi_i \, x_i \, y_i ,$$

that is a positive definite inner product on $\mathbb{R}^n$. The associated Euclidean norm is denoted by $\|.\|_\pi$.

For any $n \times n$ square matrix P, $P^{\dagger_\pi}$ denotes the adjoint of P with respect to the inner product $< .\,, . >_\pi$. We easily check that

$$P^{\dagger_\pi}_{ij} = \frac{\pi_j}{\pi_i} P_{ji}.$$

Equivalently,

$$P^{\dagger_\pi} = \delta_\pi^{-1} P^{\mathrm{T}} \delta_\pi$$

where $\delta_\pi = \mathrm{diag}(\pi_1, \ldots, \pi_{\mathrm{n}})$ and P^{T} is P's transpose.

The real vector space generated by $\mathbf{1} = (1, \ldots, 1)^T$ is denoted by $\Delta = \mathbb{R}.\mathbf{1}$, and $\Delta^{\perp_\pi}$ is the orthogonal complement of Δ in $\mathbb{R}^n$ for the inner product $< .,. >_\pi$. Clearly, $\|\mathbf{1}\|_\pi = 1$.

Another norm on $\Delta^{\perp_\pi}$ is provided by the restriction to $\Delta^{\perp_\pi}$ of the semi-norm N on $\mathbb{R}^n$ defined by

$$\mathrm{N}(x) = \max_{i \in [n]}(x_i) - \min_{i \in [n]}(x_i).$$

2.2 Reversible stochastic matrices

Let P be a stochastic matrix of size n, and let G_P denote the directed graph associated to P. We assume throughout that P is *irreducible*, i.e., G_P is strongly connected. The Perron-Frobenius theorem shows that the spectral radius of P, namely 1, is an eigenvalue of P of geometric multiplicity one. Then P has a unique Perron vector, that is, there is a unique positive probability vector π_P such that $P^{\mathrm{T}}\,\pi_P = \pi_P$. The matrix $P^{\dagger_{\pi_P}}$, simply denoted $P^\dagger$, is stochastic. Indeed,

$$\left(\delta_{\pi_P}^{-1} P^{\mathrm{T}} \delta_{\pi_P}\right) \mathbf{1} = \left(\delta_{\pi_P}^{-1} P^{\mathrm{T}}\right) \pi_P = \delta_{\pi_P}^{-1} \pi_P = \mathbf{1}.$$

Therefore, $\Delta^{\perp_{\pi_P}}$, denoted $\Delta^{\perp_P}$ for short, is stable under the action of P. Moreover the two matrices P and $P^\dagger$ share the same Perron vector.

The matrix P is said to be π-*self-adjoint* if $P^{\dagger_\pi} = P$. A simple argument based on the unicity of the Perron vector of an irreducible matrix shows that if P is π-self-adjoint, then π is P's Perron vector, i.e., $\pi = \pi_P$. In this case, the matrix P is said to be *reversible*.

2.3 A formula *à la Green*

We start with an equality that is a generalization of Green's formula.

Proposition 1. *Let π be any positive probability vector in $\mathbb{R}^n$, and let L be a square matrix of size n. If L is π-self-adjoint and $\mathbf{1} \in \ker(L)$, then for all vector $x \in \mathbb{R}^n$, it holds that*

$$\langle x, L.x \rangle_\pi = -\frac{1}{2} \sum_{i \in [n]} \sum_{j \in [n]} \pi_i \, L_{i,j} \, (x_i - x_j)^2 \,.$$

Proof. First we observe that

$$\begin{aligned}
\textstyle\sum_{i\,,\,j} \pi_i \, L_{i\,j} \, (x_i - x_j)^2 &= \textstyle\sum_{i \neq j} \pi_i \, L_{i\,j} \, (x_i - x_j)^2 \\
&= \textstyle\sum_{i \neq j} \pi_i \, L_{i\,j} \, x_i^2 + \sum_{i \neq j} \pi_i \, L_{i\,j} \, x_j^2 - 2 \sum_{i \neq j} \pi_i \, L_{i\,j} \, x_i \, x_j .
\end{aligned}$$

Because of the assumptions on L, the first two terms are both equal to $-\sum_{i\in[n]} \pi_i \, L_{i\,i} \, x_i^2$ and so

$$\sum_{i,j} \pi_i \, L_{i\,j} \, (x_i - x_j)^2 = -2 \left(\sum_{i} \pi_i \, L_{i\,i} \, x_i^2 + \sum_{i \neq j} \pi_i \, L_{i\,j} \, x_i \, x_j \right) .$$

Besides, we have

$$\langle x, Lx \rangle_\pi = \sum_{i,j} \pi_i \, L_{i\,j} \, x_i \, x_j = \sum_{i} \pi_i \, L_{i\,i} \, x_i^2 + \sum_{i \neq j} \pi_i \, L_{i\,j} \, x_i \, x_j$$

and the lemma follows. □

2.4 Norms on $\Delta^{\perp_\pi}$

As an immediate consequence of the above proposition, we obtain that if P is a reversible stochastic matrix, then the quadratic form

$$\mathcal{Q}_P(x) = \langle x, x - Px \rangle_{\pi_P}$$

is non-negative and its restriction to $\Delta^{\perp_P}$ is positive definite. Moreover, P has n real eigenvalues $\lambda_1(P), \ldots, \lambda_n(P)$ that satisfy

$$-1 \leqslant \lambda_n(P) \leqslant \ldots \leqslant \lambda_2(P) \leqslant \lambda_1(P) = 1.$$

The Perron-Frobenius theorem shows that if, in addition, P has a positive diagonal entry, then the first and the last inequalities are strict.

Besides, we obtain the classical minmax characterization of the eigenvalues of reversible stochastic matrices.

Lemma 2. *Let P be any reversible stochastic matrix, and let π be its Perron vector. For any positive real number γ, the two following assertions are equivalent*

1. $\lambda_2(P) \leqslant 1 - \gamma$;
2. $\forall x \in \Delta^{\perp_P}, \ \mathcal{Q}_P(x) \geqslant \gamma \, \|x\|_\pi^2$.

In other words, $\lambda_2(P) = 1 - \inf_{x \in \Delta^{\perp_P} \setminus \{\mathbf{0}\}} \frac{\mathcal{Q}_P(x)}{\|x\|_\pi^2}$.

Proof. Let $\{\varepsilon_1, \ldots, \varepsilon_n\}$ be an orthonormal basis for the inner product $\langle ., . \rangle_\pi$ such that $\varepsilon_1 = \mathbf{1}$ and for each index $i \in [n]$,

$$P\varepsilon_i = \lambda_i(P) \, \varepsilon_i.$$

Let $z_1, \ldots, z_n$ the components of x in this basis, namely,

$$x = z_1\varepsilon_1 + \cdots + z_n\varepsilon_n \, .$$

Hence,

$$\mathcal{Q}_P(x) = \sum_{i \in [n]} \big(1 - \lambda_i(P)\big) z_i^2$$

which shows the equivalence of the two assertions in the lemma. □

Another corollary of Proposition 1 is the following inequality between the two norms $\|.\|_\pi$ and N on $\Delta^{\perp_\pi}$, where π is any positive probability vector.

Corollary 3. *If π is a positive probability vector, then the Euclidean norm $\|.\|_\pi$ is bounded above on $\Delta^{\perp_\pi}$ by the semi-norm $\mathrm{N}/\sqrt{2}$, i.e.,*

$$\forall x \in \Delta^{\perp_\pi},\ \mathrm{N}(x) \geqslant \sqrt{2}\,\|x\|_\pi.$$

Proof. Let us consider the orthogonal projector $\pi.\mathbf{1}^{\mathrm{T}}$ on Δ, where π is P's Perron vector. Thus, for any vector in $x \in \Delta^{\perp_P}$, we have

$$\|x\|_\pi^2 = \langle x, x - \pi.\mathbf{1}^{\mathrm{T}}.\, x\rangle_\pi.$$

Since $\pi.\mathbf{1}^{\mathrm{T}}$ is stochastic and reversible, Proposition 1 gives

$$\|x\|_\pi^2 = \frac{1}{2}\sum_{i\in[n]}\sum_{j\in[n]}(x_i - x_j)^2\pi_i\,\pi_j \tag{1}$$

and the inequality $\mathrm{N}(x) \geqslant \sqrt{2}\,\|x\|_\pi$ immediately follows. □

3 The spectral gap of a reversible stochastic matrix

3.1 An analytic bound

We start by introducing the following notation: given a stochastic matrix P and its Perron vector π, we set

$$\mu(P) = \min_{\emptyset\subsetneq S\subsetneq[n]}\left(\sum_{i\in S}\sum_{j\notin S}\pi_i\,P_{i\,j}\right).$$

Lemma 4 (Lemma 8 in [16]). *If P is a reversible stochastic matrix, then for every vector $x \in \mathbb{R}^n$,*

$$\mathcal{Q}_P(x) \geqslant \frac{\mu(P)}{n-1}\big(\mathrm{N}(x)\big)^2.$$

Proof. Using index permutation, we assume that $x_1 \leqslant \ldots \leqslant x_n$. Since for any nonnegative numbers $v_1, \ldots, v_k$, we have

$$(v_1 + \cdots + v_k)^2 \geqslant v_1^2 + \cdots + v_k^2\,,$$

it follows that

$$\sum_{i<j}\pi_i P_{ij}(x_i - x_j)^2 \geqslant \sum_{i<j}\left(\pi_i P_{ij}\sum_{d=i}^{j-1}(x_{d+1} - x_d)^2\right).$$

By reordering the terms in the last sum, we obtain

$$\sum_{i<j}\pi_i P_{ij}(x_i - x_j)^2 \geqslant \sum_{d=1}^{n-1}\sum_{i=1}^{d}\sum_{j=d+1}^{n}\pi_i P_{ij}(x_{d+1} - x_d)^2\,.$$

Then Proposition 1 shows that

$$\mathcal{Q}_P(x) \geqslant \mu(P) \sum_{d=1}^{n-1} (x_{d+1} - x_d)^2 .$$

By Cauchy-Schwarz, we have

$$\sum_{d=1}^{n-1} (x_{d+1} - x_d)^2 \geqslant \frac{1}{n-1} (x_n - x_1)^2 ,$$

which completes the proof. □

That leads us to introduce

$$\eta(P) = \frac{n-1}{2\,\mu(P)}. \tag{2}$$

Combining Corollary 3 with Lemmas 2 and 4, we obtain the following lower bound on the spectral gap of a reversible stochastic matrix.

Proposition 5. *If P is a reversible stochastic matrix, then*

$$\lambda_2(P) \leqslant 1 - \frac{1}{\eta(P)}$$

with $\eta(P)$ defined by (2).

The quantity $\mu(P)$ is related to the *Cheeger constant*

$$h(P) = \min_{\pi(S) \leqslant 1/2} \frac{\sum_{i \in S} \sum_{j \notin S} \pi_i \, P_{i\,j}}{\pi(S)}$$

and satisfies $\mu(P) \leqslant h(P)/2$. Cheeger's inequalities

$$1 - 2h(P) \leqslant \lambda_2(P) \leqslant 1 - \frac{h(P)^2}{2} \tag{3}$$

give an estimate of the second eigenvalue of P. The bound $1 - 1/\eta(P)$ in Proposition 5 is incomparable with $1 - h(P)^2/2$, but turns out to be worse in most cases[1]. Moreover, computing $\mu(P)$, or equivalently $\eta(P)$, is as difficult as computing $h(P)$ in general – so why presenting the bound $1 - 1/\eta(P)$? In fact, our primary motivation here is developed in Section 4: the latter bound gives a simple estimate on the singular values of even non-reversible stochastic matrices.

3.2 A geometric bound

Following [7], we define the *P-length* of a path $\gamma = u_1, \ldots, u_{\ell+1}$ in the graph G_P by

$$|\gamma|_P = \sum_{k \in [\ell]} \left(\pi_{u_k} P_{u_k\, u_{k+1}} \right)^{-1}.$$

[1] If $\mu(P) \geqslant 1/(n-1)$, then $1 - h(P)^2/2 \leqslant 1 - 1/\eta(P)$. This inequality also holds in all the examples in Section 6.

For our geometric bound, we consider a family of paths in the graph G_P defined as follows: for each pair of nodes i, j, let $\Gamma_{i,j}$ be a non empty set of edge-disjoint paths from i to j. Since P is irreducible, such a set exists. Moreover, Menger's theorem shows that $\Gamma_{i,j}$ may be chosen with cardinality equal to any integer in $[\kappa]$, where κ is the edge-connectivity of G_P[2]. As will become clear, the quality of our estimate depends on making a judicious choice for the path sets $\Gamma_{i,j}$.

The geometric quantity that appears in our bound is

$$\kappa(P) = \max_{i \neq j} \left(\sum_{\gamma \in \Gamma_{i,j}} |\gamma|_P^{-1} \right)^{-1} . \tag{4}$$

Proposition 6. *If P is a reversible stochastic matrix, then*

$$\lambda_2(P) \leqslant 1 - \frac{1}{\kappa(P)}$$

where $\kappa(P)$ is defined by (4).

Proof. Let i and j be any pair of distinct nodes. Proposition 1 shows that

$$\mathcal{Q}_P(x) \geqslant \frac{1}{2} \sum_{\gamma \in \Gamma_{i,j}} \sum_{(u,v) \in \gamma} \pi_u \, P_{uv} (x_u - x_v)^2,$$

where π denotes the Perron vector of P. By convexity of the square function, we have

$$\left(\sum_{(u,v) \in \gamma} (x_u - x_v) \right)^2 \leqslant \sum_{(u,v) \in \gamma} \pi_u \, P_{uv} (x_u - x_v)^2 \sum_{(u,v) \in \gamma} \frac{1}{\pi_u \, P_{uv}},$$

which implies

$$\mathcal{Q}_P(x) \geqslant \left(\sum_{\gamma \in \Gamma_{i,j}} \frac{1}{|\gamma|_P} \right) \frac{(x_i - x_j)^2}{2} \geqslant \frac{(x_i - x_j)^2}{2 \, \kappa(P)}.$$

Hence

$$\mathcal{Q}_P(x) = \sum_{i \in [n]} \sum_{j \in [n]} \mathcal{Q}_P(x) \pi_i \pi_j \geqslant \frac{1}{\kappa(P)} \left(\frac{1}{2} \sum_{i \in [n]} \sum_{j \in [n]} (x_i - x_j)^2 \pi_i \, \pi_j \right) = \frac{1}{\kappa(P)} \, \|x\|_\pi^2,$$

and the result follows. The first equality holds because the sum of π's entries is 1 and the second one is the formula (1). ☐

Let us now recall some notions from graph theory (see, e.g., [8]). First, define the *depth* of a set of paths in a directed graph G as the maximum length of all its paths. For every positive integer k and every pair of nodes (i, j), the *k-distance from i to j*, denoted $d_k(i, j)$, is the minimum depth of the sets of pairwise disjoint-edge paths from i to j of cardinality k, if there is any; otherwise,

[2] The *edge-connectivity* of a directed graph G is defined to be the minimum number of edges in G whose removal results in a directed graph that is not strongly connected.

the k-distance from i to j is infinite. Then the *k-diameter of G*, denoted $\delta_k(G)$, is the maximum k-distance between any pair of nodes. The 1-diameter of G thus coincides with its diameter.

The parameter that naturally emerges when one looks for estimates of $\kappa(P)$ is the *normalized diameter* of G, denoted $\delta_*(G)$, defined by

$$\delta_*(G) = \min_{k \geqslant 1} \frac{\delta_k(G)}{k}. \tag{5}$$

It clearly satisfies $\delta_*(G) \leqslant \delta(G)$. Moreover, Menger's theorem shows that $\delta_k(G)$ is finite if and only if k is less or equal to the edge-connectivity of G, denoted $\tau_{\mathrm{e}}(G)$, thus providing the upper bound $\delta_*(G) \leqslant (n-1)/\tau_{\mathrm{e}}(G)$.

Let k be any integer such that $1 \leqslant k \leqslant \tau_{\mathrm{e}}(G)$. For every set $\Gamma_{i,j}$ of k edge-disjoint paths from i to j, we have

$$\sum_{\gamma \in \Gamma_{i,j}} \frac{1}{|\gamma|} \geqslant \frac{k}{d_k(i,j)} \geqslant \frac{k}{\delta_k(G_P)}.$$

It follows that if k realizes the minimum in (5), then

$$\sum_{\gamma \in \Gamma_{i,j}} \frac{1}{|\gamma|} \geqslant \frac{1}{\delta_*(G_P)}.$$

By setting

$$\alpha(P) = \min_{(i,j) \in E(G_P)} \pi_i P_{ij}, \tag{6}$$

we have $|\gamma|_P \leqslant |\gamma|/\alpha(P)$, and hence

$$\kappa(P) \leqslant \frac{\delta_* G_P)}{\alpha(P)}.$$

Thus, we obtain the following corollary to Proposition 6.

Corollary 7. *The eigenvalues of a reversible stochastic matrix smaller than 1 are bounded above by*

$$\beta_b(P) = 1 - \frac{\alpha(P)}{\delta_*(G_P)},$$

where $\alpha(P)$ is defined by (6) and $\delta_*(G_P)$ is the normalized diameter of the graph associated to P.

3.3 Diaconis and Stroock's geometric bound

We now present another geometric bound on the spectral gap of a reversible stochastic matrix, which has been developed by Diaconis and Stroock [7]. It depends on the choice of a set of paths in the directed graph G_P, one for each ordered pair of distinct nodes: for every pair i, j of nodes, let $\gamma_{i\,j}$ be a path from i to j, and let Γ be the set of all these paths.

The geometric quantity that appears in their bound is

$$\tilde{\kappa}(P) = \max_e \sum_{e \in \gamma_{i\,j}} |\gamma_{i\,j}|_P \pi_i \, \pi_j, \tag{7}$$

where the maximum is over edges in the directed graph G_P and the sum is over all paths in Γ that traverse e.

Diaconis and Stroock [7] developed a discrete analog of the Poincaré's inequality for estimating the spectral gap of the Laplacian on a domain:

Proposition 8 (Proposition 1 in [7]). *If P is a reversible stochastic matrix, then*

$$\lambda_2(P) \leqslant 1 - \frac{1}{\tilde{\kappa}(P)}$$

where $\tilde{\kappa}(P)$ is defined by (7).

As for our bound which depends on the choice of the path sets Γ_{ij}, the quality of their estimate depends on the choice for the paths $\gamma_{i,j}$: the lower bound $\tilde{\kappa}(P)$ is all the better if selected paths do not traverse any one edge too often. Following [7], every path $\gamma_{i,j}$ is chosen to be a geodesic. The geometric quantity that arises here is a measure of bottlenecks in G_P defined as

$$b(G_P) = \min_{\Gamma} \max_{e} |\{\gamma \in \Gamma : e \in \gamma\}| \,, \tag{8}$$

where the minimum is over the sets of paths Γ containing only geodesics, and the maximum is over all the edges of G_P. It can be shown that

$$\frac{n-1}{\tau_{\mathrm{e}}(G_P)} \leqslant b(G_P) \leqslant n^2, \tag{9}$$

where $\tau_{\mathrm{e}}(G_P)$ is the edge-connectivity of G_P. (The second inequality is straightforward; the first one may be proved by considering the partitioning of G_P into two strongly connected components when removing a certain set of $\tau_{\mathrm{e}}(G_P)$ edges.) Hence, $\delta_*(G_P) \leqslant b(G_P)$.

Like the first geometric bound $1 - 1/\kappa(P)$, the bound $1 - 1/\tilde{\kappa}(P)$ can be usefully approximated as follows.

Corollary 9. *The eigenvalues of a reversible stochastic matrix P other than 1 are upper-bounded by*

$$\beta_{DS}(P) = 1 - \frac{\alpha(P)}{(\pi_{\max})^2 \, \delta(G_P) \, b(G_P)}$$

where $\alpha(P)$ is defined by (6), $\pi_{\max}$ is the largest entry of the Perron vector of P, $\delta(G_P)$ and $b(G_P)$ are the diameter and the bottleneck measure of the graph associated to P, respectively.

4 Upper bounds on the second singular value of a stochastic matrix

Let A be any irreducible stochastic matrix of size n with positive diagonal entries. If π is the Perron vector of A, then the matrix $A^\dagger A$ is also stochastic, and the three stochastic matrices A, $A^\dagger$, and $A^\dagger A$ share the same Perron vector π. Moreover, $A^\dagger A$ is reversible and has n non negative eigenvalues.

Propositions 5, 6, and 8 provide lower bounds on the spectral gap of $A^\dagger A$, which involve the positive coefficients $\pi_i (A^\dagger A)_{ij}$ when positive. These coefficients are roughly bounded below by $\alpha(A)^2/\pi_{\max}$ with $\pi_{\max} = \max_{i \in [n]} \pi_i$ and $\alpha(A)$ defined by (6).

Interestingly, a generalization of a result in [16] combined with Proposition 5 gives an analytic bound on the spectral gap that is linear in the coefficient $\alpha(A)$ and that holds even when A is non reversible. In the case the matrix A is reversible, a lower bound on the spectral gap of A easily provides a lower bound on the spectral gap of $A^\dagger A$.

4.1 Analytic bound

We start with a lemma that has been established in [16] under the more restrictive assumption of doubly stochastic matrices.

Lemma 10. *If A is an irreducible stochastic matrix, then*

$$\mu(A^\dagger A) \geqslant \alpha(A)/2\,.$$

Proof. Let S be any non empty subset of $[n]$. Since A is a stochastic matrix, for every index $k \in [n]$, either $\sum_{i\in S} A_{ki} > 1/2$ or $\sum_{j\notin S} A_{kj} \geqslant 1/2$, and the two cases are exclusive, that is, the two subsets of $[n]$ defined by

$$S^+ = \{k \in [n] : \sum_{i\in S} A_{ki} > 1/2\} \text{ and } S^- = \{k \in [n] : \sum_{j\notin S} A_{kj} > 1/2\}$$

satisfy $S^- = [n] \setminus S^+$. Hence,

$$\sum_{i\in S}\sum_{j\notin S} \pi_i \left(A^\dagger A\right)_{i\,j} = \sum_{k\in[n]}\sum_{i\in S}\sum_{j\notin S} \pi_k A_{ki} A_{kj} \geqslant \frac{1}{2}\left(\sum_{k\in S^+}\sum_{j\notin S} \pi_k A_{kj} + \sum_{k\in S^-}\sum_{i\in S} \pi_k A_{ki}\right).$$

Then we consider the two following cases:

1. Either $S^- \cap S \neq \emptyset$ or $S^+ \cap ([n] \setminus S) \neq \emptyset$. If ℓ is in one of these two sets, then we obtain that

$$\sum_{i\in S}\sum_{j\notin S} \pi_i \left(A^\dagger A\right)_{i\,j} \geqslant \frac{\pi_\ell A_{\ell\ell}}{2}.$$

2. Otherwise, $S^+ = S$. Since A is irreducible, the non-empty set S has an outgoing edge (k_1, j) and an incoming edge (k_2, i) in G_A. It follows that

$$\sum_{i\in S}\sum_{j\notin S} \pi_i \left(A^\dagger A\right)_{i\,j} \geqslant \frac{1}{2}\left(\pi_{k_1} A_{k_1 j} + \pi_{k_2} A_{k_2 i}\right).$$

In both cases, we arrive at $\sum_{i\in S}\sum_{j\notin S} \pi_i \left(A^\dagger A\right)_{ij} \geqslant \alpha(A)/2$. □

Applied to the stochastic matrix $A^\dagger A$, Proposition 5 takes the form:

Proposition 11. *Let A be an irreducible stochastic matrix with a positive diagonal. The matrix $A^\dagger A$ has n real eigenvalues that satisfy*

$$0 \leqslant \lambda_n(A^\dagger A) \leqslant \ldots \leqslant \lambda_2(A^\dagger A) \leqslant 1 - \frac{\alpha(A)}{n-1} < \lambda_1(A^\dagger A) = 1.$$

4.2 The reversible case

If the stochastic matrix A with positive diagonal is reversible, then the n eigenvalues of A are all real and the Perron-Frobenius theorem implies that

$$-1 < \lambda_n(A) \leqslant \ldots \leqslant \lambda_2(A) < \lambda_1(A) = 1.$$

Similarly, the stochastic matrix $A^\dagger A = A^2$ has n real eigenvalues which, written in decreasing order, satisfy

$$0 \leqslant \lambda_n(A^\dagger A) \leqslant \ldots \leqslant \lambda_2(A^\dagger A) < \lambda_1(A^\dagger A) = 1.$$

Hence $\lambda_2(A^\dagger A) = \max(|\lambda_n(A)|^2, |\lambda_2(A)|^2)$.

Propositions 5, 6, and 8 show that

$$\lambda_2(A) \leqslant 1 - \frac{1}{\min(\eta(A), \kappa(A), \tilde{\kappa}(A))}.$$

Computing $\eta(A)$ is difficult in general and thus we keep on just with the two geometric bounds $\kappa(A)$ and $\tilde{\kappa}(A)$.

Every eigenvalue of A lies within at least one Gershgorin disc $D(A_{ii}, 1 - A_{ii})$, and thus

$$-1 + 2\,a(A) \leqslant \lambda_n(A) \tag{10}$$

where $a(A) = \min_{i \in [n]} A_{i\,i}$.

Proposition 12. *Let A be a irreducible stochastic matrix with a positive diagonal. If A is reversible, then the stochastic matrix $A^\dagger A$ has n real eigenvalues that satisfy*

$$0 \leqslant \lambda_n(A^\dagger A) \leqslant \ldots \leqslant \lambda_2(A^\dagger A) \leqslant \left(1 - \min\left(2\,a(A)\,,\ \frac{1}{\min(\kappa(A), \tilde{\kappa}(A))}\right)\right)^2 < \lambda_1(A^\dagger A) = 1$$

with $a(A) = \min_{i \in [n]} A_{i\,i}$, $\kappa(A)$, and $\tilde{\kappa}(A)$ defined by (4) and (7), respectively.

Every path γ in G_A satisfies

$$|\gamma|_A \leqslant \frac{|\gamma|}{\alpha(A)}$$

where $|\gamma|$ denotes γ's length and $\alpha(A)$ is defined by (6). If each path set Γ_{ij} is reduced to one shortest path from i to j, then

$$\kappa(A) \leqslant \frac{n-1}{\alpha(A)}.$$

Proposition 12 thus improves the general bound of $\alpha(A)/(n-1)$ in Proposition 11 when A is reversible.

5 Averaging algorithms and convergence rates

5.1 Averaging algorithms, stochastic matrices and asymptotic consensus

We consider a discrete time system of n autonomous agents, denoted $1, \ldots, n$, connected via a network that may change over time. Communications at time t are modelled by a directed graph $\mathbb{G}(t) = ([n], E(t))$. Since an agent can communicate with itself instantaneously, there is a self-loop at each node in every graph $\mathbb{G}(t)$. The sets of incoming and outgoing neighbors of the agent i in $\mathbb{G}(t)$ are denoted by $\mathrm{In}_i(t)$ and $\mathrm{Out}_i(t)$, respectively. The sequence $\mathbb{G} = \big(\mathbb{G}(t)\big)_{t\geqslant 1}$ is called *the dynamic communication graph*, or just the *communication graph*.

In an averaging algorithm $\mathcal{A}$, each agent i maintains a local variable x_i, initialized to some scalar value $x_i(0)$, and applies an update rule of the form

$$x_i(t) = \sum_{k \in \mathrm{In}_i(\mathbb{G}(t))} A_{ik}(t)\, x_k(t-1) \tag{11}$$

with $A_{ik}(t)$ which are all positive and $\sum_{k\in In_i(\mathbb{G}(t))} A_{ik}(t) = 1$. The algorithm $\mathcal{A}$ precisely consists in the choice of the weights $A_{ik}(t)$; typical averaging algorithms are examined in Section 6. The update rule (11) corresponds to the equation

$$x(t) = A(t)\, x(t-1)$$

where $A(t)$ is the $n \times n$ stochastic matrix whose (i,k)-entry is the weight $A_{ik}(t)$ if (k,i) is an edge in $\mathbb{G}(t)$, and 0 otherwise. Hence, the directed graph associated to the matrix $A(t)$ is the reverse graph of $\mathbb{G}(t)$.

An execution of $\mathcal{A}$ is totally determined by the initial state $x(0) \in \mathbb{R}^n$ and the communication graph $\mathbb{G}$. We say that $\mathcal{A}$ *achieves asymptotic consensus* in an execution if the sequence $x(t)$ converges to a vector x^* that is colinear to $\mathbf{1} = (1,\ldots,1)^{\mathrm{T}}$. The *convergence rate* in this execution is defined as

$$\varrho = \lim_{t\to\infty} \|x(t) - x^*\|^{1/t}$$

where $\|.\|$ is any norm on $\mathbb{R}^n$.

The classes of averaging algorithms under consideration and their executions are restricted by the following assumptions.

A1: All the directed graphs $\mathbb{G}(t)$ have a self-loop at each node and are strongly connected.

A2: There exists some positive lower bound on the positive entries of the matrices $A(t)$.

Observe that A1 is equivalent to the fact that every matrix $A(t)$ has a positive diagonal and is ergodic. As an immediate consequence of the fundamental convergence results in [13, 2], we have that asymptotic consensus is achieved in every run of an averaging algorithm satisfying A1-2.

5.2 Case of a constant Perron vector

Our first results concern executions that satisfy the following assumption in addition to A1-2.

A3: All the matrices $A(t)$ share the same Perron vector π.

Observe that under the assumption A3, the limit vector x^*, if exists, is equal to $\sum_{i\in[n]} \pi_i x_i(0)\mathbf{1}$.

The assumption A3 holds for time-varying communication graphs that arise in diverse classical averaging algorithms (e.g., see Section 6). Besides, the validity of A2 and A3 allows us to introduce the two positive infima

$$a = \inf_{i\in[n]} A_{i\,i}(t) \quad \text{and} \quad \alpha = \inf_{(i,j)\in E(t)} \pi_i A_{i\,j}(t). \tag{12}$$

The inequality (10) shows that all the eigenvalues of the matrices $A(t)$ are uniformly bounded below by $-1+2a > -1$. Moreover, since the number n of agents is fixed, the quantities $\eta(A(t))$, $\kappa(A(t))$, and $\tilde{\kappa}(A(t))$ defined by (2), (4), and (7), respectively, are uniformly bounded from the above.

Theorem 13. *In any of its executions satisfying the assumptions A1-3, an averaging algorithm achieves asymptotic consensus with a convergence rate*

$$\varrho \leqslant \sup_{t\geqslant 1} \sqrt{\lambda_2\big(A(t)^\dagger A(t)\big)}.$$

Proof. Let $y(t)$ denote the π-orthogonal of $x(t)$ on $\Delta^{\perp_\pi}$. Since $A(t)^\dagger$ is stochastic, then

$$\langle x(t), \mathbf{1}\rangle_\pi = \langle A(t)\, x(t-1), \mathbf{1}\rangle_\pi = \langle x(t-1), \mathbf{1}\rangle_\pi.$$

Therefore, the orthogonal projection of $x(t)$ on Δ is constant and $y(t) = A(t)y(t-1)$. Let $\mathcal{V}(t)$ be the variance of $x(t)$, that is

$$\mathcal{V}(t) = \|x(t) - a\mathbf{1}\|_\pi^2 = \|y(t)\|_\pi^2$$

with $a = \langle x(0), \mathbf{1}\rangle_\pi$. Then

$$\mathcal{V}(t-1) - \mathcal{V}(t) = \langle y(t-1), y(t-1)\rangle_\pi - \langle A(t)y(t-1), A(t)y(t-1)\rangle_\pi = \mathcal{Q}_{A(t)^\dagger A(t)}\left(y(t-1)\right).$$

By Proposition 1, it follows that $\mathcal{V}$ is non-increasing. Moreover, the variational characterization in Lemma 2 shows that

$$\mathcal{V}(t) \leqslant \beta^t\, \mathcal{V}(0)\,,$$

where β is any uniform upper bound on the second largest eigenvalues of the matrices $A(t)^\dagger A(t)$. ☐

In addition to A1-3, we may assume permanent reversibility.

A4: All the matrices $A(t)$ are reversible.

Corollary 14. *In any of its executions satisfying the assumptions A1-4, an averaging algorithm achieves asymptotic consensus with a convergence rate*

$$\varrho \leqslant 1 - \min\left(2\,a\,,\ \frac{1}{\min(\kappa, \tilde{\kappa})}\right),$$

where a is defined by (12), and κ and $\tilde{\kappa}$ are uniform upper bounds on $\kappa(A(t))$ and $\tilde{\kappa}(A(t))$.

Proof. Let κ and $\tilde{\kappa}$ be uniform upper bounds on $\kappa(A(t))$ and $\tilde{\kappa}(A(t))$, respectively (assumptions A2-3). Proposition 12 shows that for any positive integer t,

$$\lambda_2\big(A(t)^\dagger A(t)\big) \leqslant \left(1 - \min\left(2\,a\,,\ \frac{1}{\min(\kappa, \tilde{\kappa})}\right)\right)^2.$$

The result immediately follows from Theorem 13. ☐

If at every time t, the matrix $A(t)$ is symmetric and $\mathbb{G}(t)$ is the complete graph, then Corollary 14 gives the bound

$$\varrho \leqslant 1 - \frac{\inf_{i,j\in[n]^2,\ t\geqslant 1} A_{ij}(t)}{n}.$$

This is the bound developed by Cucker and Smale [5] to analyze the formation of flocks in a population of autonomous agents which move together.

5.3 Small variations of the Perron vector

Theorem 13 shows that in any execution of the *EqualNeighbor* algorithm – where the weights and the entries of Perron vectors are bounded below by $1/n$ and $1/n^2$, respectively (cf. Section 6) – the convergence rate is in $1-O(n^{-3})$ if the Perron vector is constant. With time-varying Perron vectors, no polynomial bound holds. Indeed, Olshevsky and Tsitsiklis [19] proved that the convergence time of this averaging algorithm is exponentially large in an execution where the support of the communication graph is fixed but agents move from one node to another node: in the $n/2$-periodic communication graph formed with bidirectional 2-stars of size n, the convergence rate is larger than $1 - 2^{3-n/2}$ while entries of each Perron vector is greater than $1/6$ for the two centers and greater than $2/3n$ for the other agents.

Our next result, which consists in an extension of Theorem 13 to the case of a time-varying Perron vector, provides a heuristic analysis of convergence rates: an exponential convergence time as in the above example may occur only if the Perron vector of the matrices $A(t)$ significantly vary over time.

We start by weakening the assumption A3.

A3b: Entries of the Perron vectors are uniformly lower bounded by some positive real number.

Under the assumption A3b, the infima a and α defined by (12) are still positive. Moreover, the quantity

$$\nu = \sup_{i\in[n],\, t>0} \sqrt{\frac{\pi_i(t+1)}{\pi_i(t)}} \tag{13}$$

is finite.

Theorem 15. *In any of its executions satisfying the assumptions A1-2 and A3b, an averaging algorithm achieves asymptotic consensus with a convergence rate*

$$\varrho \leqslant \nu \sup_{t\geqslant 1} \sqrt{\lambda_2\big(A(t)^{\dagger}A(t)\big)}.$$

Proof. For any norm $\|.\|$ on $\mathbb{R}^n$, let $\|.\|_{\mathbb{R}^n/\Delta}$ denote the quotient norm on the quotient vector space $\mathbb{R}^n/\Delta$, given by

$$\|\,[x]\,\|_{\mathbb{R}^n/\Delta} = \inf_{y\in[x]} \|y\|$$

where $[x] = x + \Delta$. It will be simply denoted $\|\,[x]\,\|$, as no confusion can arise. In the case of the Euclidean norm $\|.\|_{\pi}$, we have

$$\|\,[x]\,\|_{\pi} = \|y\|_{\pi},$$

where y is the orthogonal projection of x onto $\Delta^{\perp_{\pi}}$.

If Δ is an invariant subspace of the linear operator $A : \mathbb{R}^n \to \mathbb{R}^n$, then let $[A] : \mathbb{R}^n/\Delta \to \mathbb{R}^n/\Delta$ denote the corresponding quotient operator. The operator norm of $[A]$ associated to quotient norm $\|.\|_{\pi}$ is defined as $\|\,[A]\,\|_{\pi} = \sup_{[x]\neq 0} (\|\,[A]\,[x]\,\|_{\pi}/\|\,[x]\,\|_{\pi})$. One can easily check that

$$\|\,[A]\,\|_{\pi} = \sup_{y\in\Delta^{\perp_{\pi}}\setminus\{0\}} \frac{\|A\,y\|_{\pi}}{\|y\|_{\pi}},$$

i.e., $\|\,[A]\,\|_{\pi} = \|A/\Delta^{\perp_{\pi}}\|_{\pi}$. Hence $\|\,[A]\,\|_{\pi} = \sqrt{\lambda_2(A^{\dagger}A)}$.

Let $x \in \mathbb{R}^n$, and let π and π' be two positive probability vector. We easily get that

$$\|x\|^2_{\pi'} \leqslant \|x\|^2_{\pi} \max_{i\in[n]} \frac{\pi'_i}{\pi_i},$$

which implies that

$$\| \, [x] \, \|^2_{\pi'} \leqslant \| \, [x] \, \|^2_{\pi} \max_{i\in[n]} \frac{\pi'_i}{\pi_i}.$$

Let us now introduce the quotient form of $\mathcal{V}(t)$ defined as

$$\mathcal{W}(t) = \| \, [x(t)] \, \|^2_{\pi(t)}.$$

Then we have

$$\mathcal{W}(t) = \| \, [A(t)] \; [x(t-1)] \, \|^2_{\pi(t)} \leqslant \| \, [A(t)] \, \|^2_{\pi(t)} \, \| \, [x(t-1)] \, \|^2_{\pi(t)} \tag{14}$$

and thus

$$\mathcal{W}(t) \leqslant \| \, [A(t)] \, \|^2_{\pi(t)} \max_{i\in[n]} \frac{\pi_i(t)}{\pi_i(t-1)} \, \mathcal{W}(t-1),$$

which completes the proof. □

The bound in Theorem 15 allows for a qualitative analysis of convergence time, but is quantitatively trivial in most cases. However, the recurring inequality (14) on which it is based may be also helpful from a quantitative viewpoint, e.g., for controlling convergence times in case the Perron vector eventually stabilizes [3].

6 Metropolis, EqualNeighbor, and FixedWeight algorithms

We now examine three fundamental averaging algorithms, classically called *Metropolis*, *EqualNeighbor*, and *FixedWeight*, which all achieve asymptotic consensus if the (time-varying) topology is permanently strongly connected. While the EqualNeighbor algorithm is directly implementable in a distributed setting, FixedWeight requires the agents to have knowledge over time: the topology may vary, but each agent is supposed to know an upper bound on its in-degrees. As for Metropolis, it requires the agents to have knowledge at distance one: each agent is supposed to know the current in-degree of its neighbors.

For each of these algorithms, the Perron vectors are constant for large classes of time-varying topologies: when the communication graph is permanently bidirectional this holds for the Metropolis algorithm, when it is permanently Eulerian, for the FixedWeight algorithm, and when it is permanently Eulerian with constant (in time or in space) degrees, for EqualNeighbor. In each of these cases, the corresponding stochastic matrices are all reversible and thus Corollary 14 applies.

6.1 Algorithms and simplified bounds

First, let us fix some notation. If $p(G)$ denotes any parameter of a directed graph G, let $p(\mathbb{G})$ denote the associated parameter for the dynamic graph $\mathbb{G}$ defined as

$$p(\mathbb{G}) = \sup_{t\geqslant 1} p(\mathbb{G}(t)).$$

For instance, $\delta(\mathbb{G})$ denotes the diameter of $\mathbb{G}$, $\delta_*(\mathbb{G})$ its normalized diameter, and $b(\mathbb{G})$ its bottleneck measure. Similarly, if $d_i(t)$ denotes the in-degree of i in $\mathbb{G}(t)$ and $d_{\max}(t)$ the maximum in-degree in this graph (i.e., $d_{\max}(t) = \max_{i\in[n]} d_i(t)$), then

$$d_{\max}(\mathbb{G}) = \max_{i\in[n],\ t\geqslant 1} d_i(t).$$

Metropolis algorithm with a time-varying bidirectional topology. Weights in the the Metropolis algorithm are given by

$$M_{ij}(t) = \begin{cases} \frac{1}{\max(d_i(t),d_j(t))} & \text{if } \ j \in \mathrm{In}_i(t) \setminus \{i\} \\ 1 - \sum_{j\in\mathcal{N}_i(t)\setminus\{i\}} \frac{1}{\max(d_i(t),d_j(t))} & \text{if } \ j = i \\ 0 & \text{otherwise .} \end{cases}$$

If $\mathbb{G}(t)$ is bidirectional, then the matrix $M(t)$ is symmetric, and so doubly stochastic. Its Perron vector is $(\frac{1}{n}, \ldots, \frac{1}{n})^{\mathrm{T}}$. In any execution of Metropolis with a communication graph that is permanently bidirectional, the Perron vector is therefore constant. Furthermore, the quantities a and α in (12) satisfy $a \geqslant 1/d_{\max}$ and $\alpha \geqslant 1/(n\, d_{\max})$. Therefore Corollary 14 takes the form:

Corollary 16. *In any execution of the Metropolis algorithm with a communication graph $\mathbb{G}$ that is permanently bidirectional, the convergence rate ϱ satisfies*

$$\varrho \leqslant 1 - \min\left(\frac{2}{d_{\max}}, \max\left(\frac{1}{n\,\delta_*\, d_{\max}}, \frac{n}{\delta\, b\, d_{\max}}\right)\right)$$

where $b = b(\mathbb{G})$, $\delta = \delta(\mathbb{G})$, $\delta_* = \delta_*(\mathbb{G})$, and $d_{\max} = d_{\max}(\mathbb{G})$.

EqualNeighbor algorithm with an Eulerian topology and constant degrees. Weights in the *EqualNeighbor* algorithm are given by

$$N_{ij}(t) = \begin{cases} \frac{1}{d_i(t)} & \text{if } \ j \in In_i(t) \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathbb{G}(t)$ is Eulerian, then the i-th entry of the Perron vector of the matrix $N(t)$ is equal to

$$\pi_i(t) = \frac{d_i(t)}{|E(t)|}\,,$$

where $|E(t)| = \sum_{i=1}^{n} d_i(t)$ is the number of edges in $\mathbb{G}(t)$. Hence in every execution of the EqualNeighbor algorithm with a communication graph $\mathbb{G}$ that is permanently Eulerian, the matrices $N(t)$ share the same Perron vector if (a) every directed graph $\mathbb{G}(t)$ is regular or (b) each node i has a constant degree d_i. In case (a), the EqualNeighbor and Metropolis algorithms coincide and Corollary 16 applies. Thus we focus on case (b).

The coefficient a defined in (12) is equal to

$$a = \frac{1}{d_{\max}}.$$

With Corollaries 7 and 9, the bound in Corollary 14 simplifies into:

Corollary 17. *Let $\mathbb{G}$ be a dynamic graph that is permanently Eulerian and such that each node i has a constant degree d_i. In any execution of the EqualNeighbor algorithm with the communication graph $\mathbb{G}$, the convergence rate ϱ satisfies*

$$\varrho \leqslant 1 - \min\left(\frac{2}{d_{\max}}, \max\left(\frac{1}{\delta_* \, |E|}, \frac{|E|}{\delta \, d_{\max}^2 \, b}\right)\right)$$

where $b = b(\mathbb{G})$, $\delta = \delta(\mathbb{G})$, $\delta_* = \delta_*(\mathbb{G})$, $d_{\max} = d_{\max}(\mathbb{G})$, and $|E| = \sum_{i\in[n]} d_i$.

FixedWeight algorithm with an Eulerian topology. For each agent i, let q_i denote an upper bound on the number of in-neighbors of i in a given dynamic graph $\mathbb{G}$. Weights in the FixedWeight algorithm are given by

$$W_{ij}(t) = \begin{cases} 1/q_i & \text{if } \; j \in \mathrm{In}_i(t) \setminus \{i\} \\ 1 - (d_i(t) - 1)\,/q_i & \text{if } \; j = i \\ 0 & \text{otherwise } \; . \end{cases}$$

We easily check that if $\mathbb{G}(t)$ is Eulerian, then the i-th entry of the $W(t)$'s Perron vector is equal to

$$\pi_i\big(W(t)\big) = \frac{q_i}{Q} \, ,$$

where $Q = \sum_{i\in[n]} q_i$. It follows that with a communication graph that is permanently Eulerian, the Perron vector is constant and each matrix $W(t)$ is reversible. Furthermore, the quantities a and α in (12) satisfy $a \geqslant 1/q$ and $\alpha = 1/Q$. Using Corollaries 7 and 9, Corollary 14 specializes to the following corollary.

Corollary 18. *In any execution of the FixedWeight algorithm with a communication graph $\mathbb{G}$ that is permanently Eulerian, the convergence rate ϱ satisfies*

$$\varrho \leqslant 1 - \min\left(\frac{2}{q}, \max\left(\frac{1}{\delta_* \, Q}, \frac{Q}{\delta \, q^2 \, b}\right)\right)$$

where $b = b(\mathbb{G})$, $\delta = \delta(\mathbb{G})$, $\delta_* = \delta_*(\mathbb{G})$, $q = \max_{i\in[n]} q_i$ and $Q = \sum_{i\in[n]} q_i$.

The quantities $1/\delta_* Q$ and $Q/\delta \, q^2 \, b$ in the above bound depend not only on the geometric parameters of $\mathbb{G}$, but also on the parameters q and Q of the FixedWeight algorithm, and hence cannot be compared in general.

6.2 Quadratic bounds on convergence rates

Under the conditions specified in Corollaries 16 and 17, the convergence rate is bounded above by $1 - 1/n^3$ for both the EqualNeighbor and the Metropolis algorithms. We show that the original Poincaré's inequality in Proposition 8 yields a convergence rate in $1 - O(1/n^2)$ for Metropolis, and prove that this bound also holds for EqualNeighbor when the communication graph is not too irregular.

First observe that the *Metropolis-length* of any path $\gamma = (i_1, \ldots, i_{\ell+1})$ in $\mathbb{G}(t)$ of length $|\gamma| = \ell$ is given by

$$|\gamma|_{M(t)} = n \sum_{k\in[\ell]} \max(d_{i_k}(t), d_{i_{k+1}}(t)),$$

while the *EqualNeighbor-length* for a communication graph with constant degrees is

$$|\gamma|_{N(t)} = |E|\,|\gamma|.$$

Our general quadratic bound for Metropolis is based on a simple combinatorial lemma inspired by a nice idea in [9].

Lemma 19. *Let G be any bidirectional graph with n nodes, and let $i_1, \ldots, i_{\ell+1}$ be any geodesic in G. Then*

$$\max(d_{i_1}, d_{i_2}) + \cdots + \max(d_{i_\ell}, d_{i_{\ell+1}}) \leqslant 4n.$$

Proof. Let $\mathcal{N}_k$ denote the set of (incoming or outgoing) neighbors of i_k, and for each $k \leqslant \ell$, let

$$\mathcal{N}_k^* = \begin{cases} \mathcal{N}_k. & \text{if } d_k \geqslant d_{k+1} \\ \mathcal{N}_{k+1} & \text{otherwise.} \end{cases}$$

Since $i_1, \ldots, i_{\ell+1}$ is a geodesic, $\mathcal{N}_k$ and $\mathcal{N}_{k'}$ are disjoint if $k' \geqslant k+3$. Hence, $\mathcal{N}_k^*$ and $\mathcal{N}_{k'}^*$ are disjoint if $k' \geqslant k+4$. The lemma follows from the pigeonhole principle. □

Proposition 20. *The Metropolis algorithm with dynamic communication graphs that are permanently bidirectional and connected achieves asymptotic consensus with a convergence rate*

$$\varrho \leqslant 1 - \frac{1}{4n^2}.$$

Proof. Lemma 19 shows that every geodesic γ in $\mathbb{G}(t)$ satisfies

$$|\gamma|_{M(t)} \leqslant 4n^2.$$

Hence, for every bound $\tilde{\kappa}(M(t))$ defined by (7), it holds that

$$\tilde{\kappa}(M(t)) \leqslant 4\,b(\mathbb{G}). \tag{15}$$

Since $b(\mathbb{G}) \leqslant n^2$, we obtain $\tilde{\kappa} \leqslant 4n^2$. The result follows from Corollary 14 and $a \geqslant 1/n$. □

The same approach applies to the *Lazy Metropolis* algorithm where weights are defined by

$$L_{ij}(t) = \begin{cases} \frac{1}{2\max(d_i(t)-1,\, d_j(t)-1)} & \text{if } j \in \mathrm{In}_i(t) \setminus \{i\} \\ 1 - \sum_{j \in \mathcal{N}_i(t) \setminus \{i\}} \frac{1}{2\max(d_i(t)-1,\, d_j(t)-1)} & \text{if } j = i \\ 0 & \text{otherwise .} \end{cases}$$

Therefore,

$$\forall i \in [n],\ \forall t \geqslant 1, \quad L_{ii}(t) \geqslant \frac{1}{2}$$

and

$$|\gamma|_{L(t)} = 2n \sum_{k \in [\ell]} \max(d_{i_k}(t), d_{i_{k+1}}(t)).$$

Corollary 14 and Lemma 19 give the following result for the Lazy Metropolis algorithm.

Proposition 21. *The Lazy Metropolis algorithm with dynamic communication graphs that are permanently bidirectional and connected achieves asymptotic consensus with a convergence rate*

$$\varrho \leqslant 1 - \frac{1}{8n^2}.$$

From the quadratic bound on the hitting times of Metropolis walks proved by Nonaka et al. [17], Olshevsky [18] showed that the convergence rate of the Lazy Metropolis algorithm on any fixed graph that is connected and bidirectional is bounded from the above by $-1/71n^2$. Proposition 21 improves this result and extends it to the case of a time-varying topology.

The Metropolis and EqualNeighbor algorithms coincide in the case of communication graphs that are permanently regular. Proposition 20 shows that the convergence rate is bounded above by $1-1/4n^2$ for such topologies, thus extending the quadratic upper bound in [7] for distance transitive graphs to any regular graphs. With moderate irregularity [1], a close method for bounding $\tilde{\kappa}$ in the EqualNeighbor algorithm gives the following quadratic bound.

Proposition 22. *In any execution of the EqualNeighbor algorithm with a communication graph* $\mathbb{G}$ *that is permanently Eulerian and with a constant degree* d_i *at each node* i*, asymptotic consensus is achieved with a convergence rate*

$$\varrho \leqslant 1 - \frac{1}{(3 + d_{\max} - d_{\min})\, n^2}$$

where $d_{\min}$ *and* $d_{\max}$ *denote the minimum and maximum degree in each graph* $\mathbb{G}(t)$*.*

Proof. The EqualNeighbor-length of any path in the directed graph $\mathbb{G}(t)$ gives

$$\tilde{\kappa}(A(t)) = \frac{1}{|E|} \max_{e} \sum_{e \in \gamma_{i\,j}} |\gamma_{i\,j}|\, d_i\, d_j.$$

Hence

$$\tilde{\kappa}(A(t)) \leqslant \frac{1}{|E|} \sum_{i \neq j} |\gamma_{i\,j}|\, d_i\, d_j \leqslant \max_{j \in [n]} \sum_{i \in [n] \setminus j} |\gamma_{i\,j}|\, d_i.$$

The second inequality is due to the fact that $|E| = \sum_{k \in [n]} d_k$. An argument analog to Lemma 19 shows that the sum of the degrees along any geodesic is less than $3n$, and thus each term in the above sum is bounded above by

$$|\gamma_{i\,j}|\, d_i(t) \leqslant 3n + (d_i - d_{\min})|\gamma_{i\,j}|.$$

The result immediately follows from Corollary 14 and $a \geqslant 1/d_{\max}$. □

The example of the *barbell* graph developed by Landau and Odlyzko [11] shows that the convergence rate of the EqualNeighbor algorithm is greater than $1 - 32/n^3$ with a specific set of initial values (see also below). Thus the general quadratic bound for Metropolis in Proposition 20 does not hold for EqualNeighbor because of degree fluctuations in space. In the light of this example and of the striking result by Olshevsky and Tsitsiklis [20], demonstrating that the EqualNeighbor algorithm may experience an exponential convergence rate with degree fluctuations in time, the Metropolis algorithm appears as a powerful and efficient method for masking graph irregularities, requiring only bidirectional communication links and limited knowledge at each agent.

7 Bounds for specific communication graphs

We now examine some typical examples where the bounds presented above are easy to compute. For the FixedWeight algorithm, we just give the bound derived from the simple geometric bound β_b, while we present detailed comparisons of the various bounds for the EqualNeighbor and Metropolis algorithms (cf. Figure 3). For Metropolis and FixedWeight, the communication graph is time-varying, but it is supposed to belong to one of the listed classes of directed graphs. In other words, the support is fixed but node labelling may change over time. For the EqualNeighbor algorithm, the communication graph is supposed to be fixed if the directed graphs in the class under consideration are not regular. This section is completed with the case of the EqualNeighbor algorithm and the fixed *Butterfly* graph, which allows us to compare the various methods for bounding convergence rate in the case of non-reversible stochastic matrices.

Ring. Let $G = (V, E)$ be a bidirectional ring[3] with an odd number $n = 2m + 1$ of nodes. Here $|E| = 3n$, $d_{\max}(G) = 3$, and $\delta(G) = m$. We easily check that $\delta_*(G) = m$ and $b(G) = m(m + 1)/2$.

Since G is regular, the EqualNeighbor and Metropolis algorithms coincide, and we obtain

$$\beta_b = 1 - \frac{2}{3n^2} + O\left(\frac{1}{n^3}\right),\ \beta_{DS} = 1 - \frac{16}{3n^2} + O\left(\frac{1}{n^3}\right).$$

The two bounds are of the same order of magnitude with $\beta_{DS} < \beta_b$. Corollary 17 gives a convergence rate

$$\varrho \leqslant 1 - \frac{16}{3n^2},$$

which is the right order for n large

Hypercube. Let $G = (V, E)$ be the p-dimensional cube with $n = 2^p$ nodes. Here $|E| = (p+1)2^p$, $d(G) = p + 1$, $\delta(G) = p$, and $\delta_*(G) = 1$. Diaconis and Stroock [7] showed that $b(G) = 2^{p-1}$. The EqualNeighbor and Metropolis algorithms coincide, and we obtain

$$\beta_b = 1 - \frac{1}{(p+1)2^p} \quad \text{and} \quad \beta_{DS} = 1 - \frac{2}{p\,(p+1)}.$$

The bound β_{DS} is far better than β_b. Corollary 17 gives a convergence rate

$$\varrho \leqslant 1 - \frac{2}{p\,(p+1)} \leqslant 1 - \frac{1}{(\log_2 n)^2},$$

which is the right order for n large (e.g., see [6]).

Star. The star graph with n nodes has $3n-2$ edges. The maximum degree is n, its diameter is 2, its edge-connectivity is 1, and so its normalized diameter is 2. The bottleneck measure is equal to the lower bound in (9), namely $n - 1$.

For the EqualNeighbor algorithm, we obtain

$$\beta_b = 1 - \frac{1}{6n} + O\left(\frac{1}{n^2}\right) \quad \text{and} \quad \beta_{DS} = 1 - \frac{3}{2n^2} + O\left(\frac{1}{n^3}\right).$$

[3] For a chain, graph parameters are of the same order and so leads to bounds of the same order of magnitude.

The bound β_{b} is far better then β_{DS}, and Corollary 17 gives a convergence rate

$$\varrho \leqslant 1 - \frac{1}{6n}.$$

As for Metropolis, we have

$$\beta_{b} = 1 - \frac{1}{2n^2} \quad \text{and} \quad \beta_{DS} = 1 - \frac{1}{3(n-1)}.$$

The bound β_{DS} is asymptotically better than β_{b} and improves the bound given in [15]. Observe that the inequality (15) in the proof of Proposition 20 directly gives $\varrho \leqslant 1 - 1/4n$.

Two-star. A two-star graph G is composed of two identical stars with an edge connecting their centers. It has an even number n of nodes and $3n-2$ edges. Here, $d_{\max}(G) = 1 + n/2$, $\delta(G) = \delta_*(G) = 3$, and $b(G) = n^2/4$.

For the Metropolis algorithm, we obtain

$$\beta_{b} = 1 - \frac{2}{3n^2} + O\left(\frac{1}{n^3}\right) \quad \text{and} \quad \beta_{DS} = 1 - \frac{8}{3n^2} + O\left(\frac{1}{n^3}\right).$$

The bounds β_{b} and β_{DS} are of the same order with $\beta_{DS} < \beta_{b}$. Corollary 16 gives a convergence rate

$$\varrho \leqslant 1 - \frac{8}{3n^2}.$$

As for EqualNeighbor, we have

$$\beta_{b} = 1 - \frac{1}{9n} + O\left(\frac{1}{n^2}\right) \quad \text{and} \quad \beta_{DS} = 1 - \frac{16}{n^3} + O\left(\frac{1}{n^4}\right).$$

The bound β_{b} is far better than β_{DS}, and Corollary 17 gives a convergence rate

$$\varrho \leqslant 1 - \frac{1}{9n}.$$

Binary tree. Consider the full binary tree of depth $p > 1$. It has $n = 2^{p+1} - 1$ nodes, $3n-2$ edges, and the maximum degree is 4. The results for the EqualNeighbor and Metropolis algorithms are thus of the same order. The diameter is $2\,p$ and the normalized diameter is $2\,p$. We easily check that the bottleneck measure is $2^p(2^p - 1)$.

For Metropolis, we have

$$\beta_{b} = 1 - \frac{1}{8\,p(2^{p+1}-1)} \quad \text{and} \quad \beta_{DS} = 1 - \frac{2^{p+1}-1}{p\,2^{p+3}(2^p-1)}.$$

The bounds β_{b} and β_{DS} are of the same order with $\beta_{DS} < \beta_{b}$. Corollary 16 gives a convergence rate

$$\varrho \leqslant 1 - \frac{1}{2n\log_2 n}.$$

The results for EqualNeighbor are similar with a convergence rate

$$\varrho \leqslant 1 - \frac{1}{4n\log_2 n}.$$

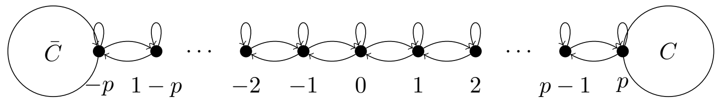

Figure 1: The barbell graph

Observe that, in the case of a general bidirectional tree, the number of edges remains equal to $3n-2$ while the diameter may be $n-1$, which leads to

$$\beta_b \leqslant 1 - \frac{1}{3\,n^2}$$

for the EqualNeighbor algorithm. Proposition 20 shows that a quadratic bound also holds for Metropolis.

Two-dimensional grid. Let p be an even positive integer, and let $G=(V,E)$ be the two-dimensional grid with $n=p^2$ nodes. Here $|E|=p(5p-4)$, $d(G)=5$, $\delta(G)=2(p-1)$, and $\delta_*(G)=p-1$. The results for EqualNeighbor and Metropolis are thus of the same order. Choosing paths $\gamma_{i,j}$ first with vertical edges and then with horizontal edges yields $b(G)\leqslant p^3(p+1)/8$.

For the Metropolis algorithm, we obtain

$$\beta_b = 1 - \frac{1}{5n^{3/2}} \quad \text{and} \quad \beta_{DS} \leqslant 1 - \frac{4}{5p(p-1)(p+1)} \leqslant 1 - \frac{2}{5n^{3/2}}.$$

The bounds β_b and β_{DS} are of the same order of magnitude. Corollary 16 gives a convergence rate

$$\varrho \leqslant 1 - \frac{2}{5n^{3/2}}.$$

Similarly, Corollary 17 implies that the convergence rate of the EqualNeighbor algorithm satisfies

$$\varrho \leqslant 1 - \frac{4}{5n^{3/2}}.$$

Barbell. The *barbell graph* $G=(V,E)$ of size $|V|=n=4p-1$ is composed of two cliques C and $\bar{C}$ with p nodes each, that are connected by a line of length $2p-1$; see Figure 1. The barbell graph is bidirectional with $|E|=2p^2+6p-1$ edges. The maximum degree is $p+1$. The diameter and the normalized diameter are equal to $2(p+1)$. Any geodesic connecting i to j with $i\leqslant 0$ and $j\geqslant 1$ crosses over the edge $(0,1)$, which is thus traversed by $2p(2p-1)$ geodesics. Clearly $(0,1)$ realizes the maximum in (8), and hence $b(G)=2p(2p-1)$).

For the Metropolis algorithm, the bounds β_b and β_{DS} are of the order of magnitude with $\beta_{DS}<\beta_b$, and β_{DS} is of the order of $1-32/p^3$. A better estimate on the convergence rate is obtained with (15) and gives

$$\varrho \leqslant 1 - \frac{1}{16p^2} = 1 - \frac{1}{(n+1)^2}.$$

The barbell graph thus exemplifies that the bound in Corollary 16 can be far from the original bound $1-1/\tilde{\kappa}$.

As for EqualNeighbor, the expression of $\tilde{\kappa}$ in (7) makes the barbell graph as a good candidate for a spectral gap that is cubic in $1/n$. Indeed, Landau and Odlyzko [11] consider the vector $v \in \mathbb{R}^n$ defined by

$$v_i = \begin{cases} -p & \text{if } i \in \bar{C} \\ i & \text{if } 1-p \leqslant i \leqslant p-1 \\ p & \text{if } i \in C. \end{cases}$$

Let N denote the stochastic matrix associated to the EqualNeighbor algorithm running on the barbell graph. Proposition 1 shows that

$$\mathcal{Q}_N(v) = \frac{1}{2|E|} \sum_{(i,j)\in E} (v_i - v_j)^2 \quad \text{and} \quad \|v\|_\pi^2 = \frac{1}{|E|} \sum_{i\in[n]} d_i v_i^2.$$

Hence

$$\mathcal{Q}_N(v) = \frac{p}{|E|} \quad \text{and} \quad \|v\|_\pi^2 = \frac{2}{|E|}\left(p^4 + \frac{4p^3}{3} + \frac{p^2}{2} + \frac{p}{6}\right).$$

Therefore

$$\lambda_2(N) \geqslant 1 - \frac{3}{6p^3 + 8p^2 + 3p + 6} \geqslant 1 - \frac{32}{n^3}.$$

The first inequality is Lemma 2 and the second one is because $n = 4p - 1$. In the execution with the initial values corresponding to one eigenvector associated to $\lambda_2(N)$, the convergence rate satisfies

$$\varrho = \lambda_2(N) \geqslant 1 - \frac{32}{n^3}.$$

Hence, as opposed to the Metropolis algorithm, no general quadratic bound holds for the convergence rate of EqualNeighbor on a fixed connected bidirectional graph.

Butterfly (and EqualNeighbor). The *Butterfly graph* has $n = 2m$ nodes and consists of two isomorphic parts that are connected by a bidirectional edge. We list the edges between the nodes $1, 2, \ldots, m$ which also determine the edges between the nodes $m+1, m+2, \ldots, 2m$ via the isomorphism $\overline{i} = n - i + 1$. The edges between the nodes $1, 2, \ldots, m$ are: (a) the edges $(i+1, i)$ for every $i \in [m-1]$, and (b) the edges $(1, i)$ for every $i \in [m]$. In addition, it contains a self-loop at each node and the two edges $(m, \overline{m})$ and $(\overline{m}, m)$. Hence, the butterfly graph is not bidirectional but it is strongly connected; see Figure 2.

We now consider the EqualNeighbor algorithm running on this fixed graph, yielding a fixed stochastic matrix B that is not reversible. Corollary 14 is not applicable, but the results in Section 4 give a convergence rate

$$\varrho \leqslant 1 - \max\left(\frac{\alpha(B)}{n-1}, \frac{1}{\kappa(B^\dagger B)}, \frac{1}{\tilde{\kappa}(B^\dagger B)}\right),$$

where $\alpha(B)$, $\kappa(B^\dagger B)$, and $\kappa(B^\dagger B)$ are defined by (6), (4), and (7), respectively.

We easily verify that the Perron vector of B, and thus of $B^\dagger B$, is given by

$$\pi_1 = \frac{1}{5},\ \pi_i = \frac{3}{5\,.\,2^i} \quad \text{for } i \in \{2, \ldots, m-1\} \text{ and } \frac{1}{5}.$$

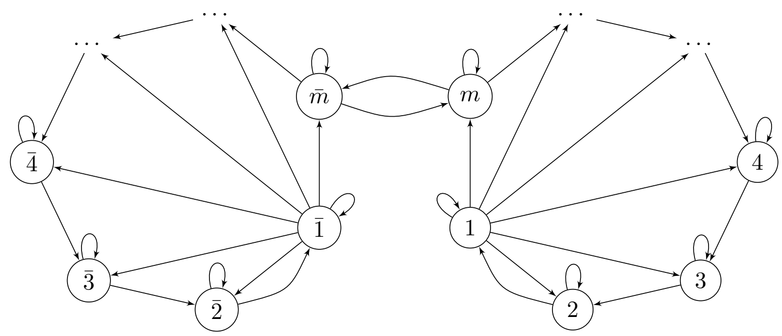

Figure 2: The butterfly graph

By symmetry, this also defines the Perron vector for the remaining indices between $m+1$ and $2m$ since $\pi_i = \pi_{n-i+1}$. Then we easily arrive at

$$\alpha(B) = \pi_m \, B_{m\,1} = \frac{1}{5\,.\,2^{m-1}},$$

which directly gives the following analytic bound in Proposition 11

$$\beta_a = 1 - \frac{1}{5(2m-1)\,2^{m-1}}.$$

For $\kappa(B^\dagger B)$ and $\kappa(B^\dagger B)$, we compute the estimates $\beta_b(B^\dagger B)$ and $\beta_{DS}(B^\dagger B)$ given by

$$\kappa(B^\dagger B) \leqslant \frac{\alpha(B^\dagger B)}{\delta_*(H)} \quad \text{and} \quad \kappa(B^\dagger B) \leqslant \frac{\alpha(B^\dagger B)}{\delta(H)(\pi_{\max})^2 b(H)},$$

where $H = G_{B^\dagger B}$. The directed graph H consists in two cliques with the sets of nodes $1, 2, \ldots, m$ and $\overline{1}, \overline{2}, \ldots, \overline{m}$, connected by the edges $(m-1, \overline{m})$, $(m, \overline{m-1})$, $(m, \overline{m})$ and the three edges in the reverse direction. Thus H has $2(m^2+3)$ edges, $d_{\max}(H) = m+2$, $\delta(H) = 3$, and $\delta_*(H) = 1$. The bottleneck measure is $b(H) = m/3$. A rather tedious computation gives

$$\alpha(B^\dagger B) = \pi_{m-1} \, B^\dagger B_{(m-1)\,m} = \frac{1}{3\,.\,5\,.\,2^{m-1}}.$$

Since $\pi_{\max} = 1/5$, we arrive at the two following geometric bounds

$$\beta_b = 1 - \frac{1}{3\,.\,5\,.\,2^{m-1}} \quad \text{and} \quad \beta_{DS} = 1 - \frac{5}{3\,m\,2^{m-1}}.$$

The bound β_b is better than both β_a and β_{DS}. Thus we arrive at

$$\varrho \leqslant 1 - \frac{1}{3\,.\,5\,.\,2^{m-1}} \ .$$

The subset $S = \{1, 2, \ldots, m\}$ satisfies $\pi(S) = 1/2$ and

$$\sum_{i \in S, j \notin S} \pi_i B^\dagger B_{i\,j} = \frac{1}{5.2^{m-2}}.$$

	EqualNeighbor	FixedWeight	Metropolis
ring	$1-\frac{16}{3n^2}$	$1-\frac{2}{Qn}$	$1-\frac{16}{3n^2}$
hypercube	$1-\frac{1}{(\log_2 n)^2}$ $[DS]$	$1-\frac{1}{Q}$	$1-\frac{1}{(\log_2 n)^2}$ $[DS]$
star	$1-\frac{1}{6n}$ $[b]$	$1-\frac{1}{2q_{\max}}$	$1-\frac{1}{3n}$ $[DS]$
two-star	$1-\frac{1}{9n}$ $[b]$	$1-\frac{1}{3Q}$	$1-\frac{8}{3n^2}$
binary tree	$1-\frac{1}{4n\log_2 n}$	$1-\frac{2}{Qn}$	$1-\frac{1}{2n\log_2 n}$
grid	$1-\frac{2}{5n\sqrt{n}}$	$1-\frac{1}{Q\sqrt{n}}$	$1-\frac{2}{5n\sqrt{n}}$
barbell	$1-\frac{8}{(n+1)^3}$	$1-\frac{2}{Qn}$	$1-\frac{1}{(n+1)^2}$

Figure 3: Bounds for networked systems with n agents and bidirectional links.

The lower bound in Cheeger's inequalities gives

$$\lambda_2(B^\dagger B) \geqslant 1 - \frac{1}{5.2^{m-4}}.$$

This lower bound is of the same order as β_b, which shows that the convergence rate of the EqualNeighbor algorithm is $1-\theta\left(2^{-m}\right)$.

Acknowledgements. I thank Eric Fuzy and Patrick Lambein-Monette for useful discussions, and Jean-Benoît Bost and Raphaël Bost for their help during the completion of this paper.

References

[1] Michael O. Albertson. The irregularity of a graph. *Ars Combinatoria*, 46:219–225, 1997.

[2] Ming Cao, A. Stephen Morse, and Brian D. O. Anderson. Reaching a consensus in a dynamically changing environment: a graphical approach. *SIAM Journal on Control and Optimization*, 47(2):575–600, 2008.

[3] Bernadette Charron-Bost and Patrick Lambein-Monette. Consensus: A little learning goes a long way. In preparation, 2020.

[4] Bernard Chazelle. The total s-energy of a multiagent system. *SIAM Journal on Control and Optimization*, 49(4):1680–1706, 2011.

[5] Felipe Cucker and Steve Smale. Emergent behavior in flocks. *IEEE Transactions on Automatic Control*, 52:852–862, 2007.

[6] Persi Diaconis. *Group representations in probability and statistics*, volume 11 of *Lecture Notes–Monograph Series*. Springer, 1988.

[7] Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of Markov chains. *The Annals of Applied Probability*, 1(1):36–61, 1991.

[8] D. Frank Hsu and Tomasz Luczak. On the k-diameter of k-regular k-connected graphs. *Discrete Mathematics*, 133:291–296, 1994.

[9] Satoshi Ikeda, Izumi Kumo, Norihiro Okumoto, and Masafumi Yamashita. Impact of local topological information on random walks on finite graphs. In *Proceedings of the 30th International Colloquium on Automata, Languages, and Programming, ICALP03*, volume 2719 of *Lecture Notes in Computer Science*, pages 1054–1067. Springer, 2003.

[10] Alexander Kolmogoroff. Zur Theorie der Markoffschen Ketten. *Mathematische Annalen*, 112(1):155–160, 1936.

[11] Henry J. Landau and Alexander M. Odlyzko. Bounds for eigenvalues of certain stochastic matrices. *Linear Algebra and Its Applications*, 38:5–15, 1981.

[12] Alexander Lubotzky. *Discrete Groups, Expanding Graphs and Invariant Measures*, volume 11 of *Lecture Notes Monograph Series*. Springer, 1989.

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

[14] Angelia Nedić and Ji Liu. On convergence rate of weighted-averaging dynamics for consensus problems. *IEEE Transactions on Automatic Control*, 62(2):766–781, 2017.

[15] Angelia Nedic, Alex Olshevsky, and Michael G. Rabbat. Network topology and communication-computation tradeoffs in decentralized optimization. *Proceedings of the IEEE*, 106(5):953–976, 2018.

[16] Angelia Nedic, Alexander Olshevsky, Asuman E. Ozdaglar, and John N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *IEEE Transactions on Automatic Control*, 54(11):2506–2517, 2009.

[17] Yoshiaki Nonaka, Hirotaka Ono, Kunihiko Sadakane, and Masafumi Yamashita. The hitting and cover times of metropolis walks. *Theoretical Computer Science*, 411(16–18):1889–1894, 2010.

[18] Alex Olshevsky. Linear time average consensus and distributed optimization on fixed graphs. *SIAM Journal on Control and Optimization*, 55(6):3990–4014, 2017.

[19] Alex Olshevsky and John N. Tsitsiklis. Convergence speed in distributed consensus and averaging. *SIAM Review*, 53(4):747–772, 2011.

[20] Alex Olshevsky and John N. Tsitsiklis. Degree fluctuations and the convergence time of consensus algorithms. *IEEE Transactions on Automatic Control*, 58(10):2626–2631, 2013.

[21] Alistair J. Sinclair. Improved bounds for mixing rates of. markov chains and multicommodity flow. *Combinatorics, Probability, & Computing*, 1:351–370, 1992.

[22] Alistair J. Sinclair and Mark R. Jerrum. Approximate counting, uniform generation, and rapidly mixing Markov chains. *Information and Computation*, 82:93–133, 1989.

[23] Lin Xiao and Stephen Boyd. Fast linear iterations for distributed averaging. *Systems & Control Letters*, 53(1):65–78, 2004.