Convergence Rate and Convergence Time

Bernadette Charron-Bost

In this note, we generalize the polynomial bound on convergence time of averaging algorithms established by Nedic et al. [2] in the case of doubly stochastic matrices. Our contribution consists in observing that the arguments developed in [2] provide an upper bound on the singular values of general (non doubly) stochastic matrices.

Interestingly, our polynomial bound also applies to the *fixed-weight* algorithm with a time-varying bidirectional topology. It thus unifies the time complexity results in [2], [1] and in [3].

1 Singular values of a stochastic matrix

1.1 Preliminaries

Let n be a positive integer and let $[n] = \{1, \ldots, n\}$.

Let $\pi \in \mathbb{R}^n$ be a positive probability vector in $\mathbb{R}^n$. We define

$$< x, y >_\pi = \sum_{i=1}^{n} \pi_i \, x_i \, y_i$$

that is an inner product on $\mathbb{R}^n$. For any $n \times n$ square matrix A, $A^{\dagger_\pi}$ denotes A's adjoint for the inner product $< .,. >_\pi$. We easily check that

$$A^{\dagger_\pi}_{ij} = \frac{\pi_j}{\pi_i} A_{ji} \, .$$

In the case A is an ergodic matrix, its Perron vector denoted $\pi(A)$ is a positive vector. For simplicity, we write $A^\dagger$ instead of $A^{\dagger_{\pi(A)}}$. We easily check that if A is a stochastic matrix, then $A^\dagger$ is also a stochastic matrix with the same Perron vector, i.e., $\pi\big(A^\dagger\big) = \pi(A)$. Therefore $A^\dagger A$ is self-adjoint for the inner product $< .,. >_{\pi(A)}$.

1.2 A formula *à la Green*

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

Proposition 1. *If π is a positive probability vector in $\mathbb{R}^n$ and L is a square matrix of size n such that $\mathbb{1} \in \ker(L)$ and $L^{\dagger_\pi} = L$, then for all vector $x \in \mathbb{R}^n$, we have*

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

Proof. First we observe that

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

Because of the properties of L, the first two terms are both equal to $-\sum_{i=1}^n \pi_i L_{ii} x_i^2$ and so

$$\sum_{(i,j)\in[n]^2} \pi_i L_{ij} (x_i - x_j)^2 = -2\left(\sum_{i=1}^n \pi_i L_{ii} x_i^2 + \sum_{i\neq j} \pi_i L_{ij} x_i x_j\right).$$

Besides, we have

$$\langle x, Lx\rangle_\pi = \sum_{i=1}^n \sum_{j=1}^n \pi_i L_{ij} x_i x_j = \sum_{i=1}^n \pi_i L_{ii} x_i^2 + \sum_{i\neq j} \pi_i L_{ij} x_i x_j$$

and the lemma then follows. □

We immediately derive the following corollary for ergodic stochastic matrices.

Corollary 2. *Let A be an ergodic stochastic matrix of size n and let π denote its Perron vector. For all vector $x \in \mathbb{R}^n$,*

$$< x, x >_\pi - < x, A^\dagger A.x >_\pi = \frac{1}{2} \sum_{(i,j)\in[n]^2} \pi_i \left(A^\dagger A\right)_{i,j} (x_i - x_j)^2 .$$

As an immediate consequence of the above corollary, we obtain that the restriction of the quadratic form

$$Q_A(x) = \langle x, x - A^\dagger A.x\rangle_\pi$$

to the orthogonal complement of $\mathbb{R}\mathbf{1}$ in $\mathbb{R}^n$ is positive definite.

1.3 An upper bound on the singular values of a stochastic matrix

Let A be an ergodic stochastic matrix of size n with positive diagonal entries. Since $A^\dagger A$ is a stochastic matrix, the n singular values of A (which are the square roots of $A^\dagger A$) in the increasing order satisfy $0 \leqslant \sigma_n \leqslant \ldots \leqslant \sigma_2 \leqslant \sigma_1 = 1$. By the Perron-Frobenius theorem, we have

$$\sigma_2 < 1 .$$

The aim of the section is to refine the latter inequality.

First we fix some notation: the Perron vector of A is denoted by π and we let

$$\mu_A = \min\{\pi_i A_{ij} \mid A_{ij} > 0\} .$$

Let Δ be the real vector space generated by $\mathbf{1} = (1, \ldots, 1)^T$, and $\Delta^\perp$ be the orthogonal complement of Δ in $\mathbb{R}^n$. We denote by δ the semi-norm on $\mathbb{R}^n$ defined by

$$\delta(x) = \max_{i=1,\ldots,n}(x_i) - \min_{i=1,\ldots,n}(x_i)$$

which is a norm on $\Delta^{\perp}$.

Moreover, let $\{\varepsilon_1, \ldots, \varepsilon_n\}$ be an orthonormal basis in which $A^{\dagger}A$ is diagonalizable and such that $A^{\dagger}A\varepsilon_i = \sigma_i^2 \varepsilon_i$ and $\varepsilon_1 = \mathbf{1}$. We start by two lemmas which are both slight variations of two results established in [2].

Lemma 3 (Lemma 5 in [2])**.** *Let $N_1 \cup N_2$ a partition of $[n]$ into two disjoint sets. If there exist two indices $i \in N_1$ and $j \in N_2$ such that $A_{ij} > 0$, then*

$$\sum_{i \in N_1,\, j \in N_2} \pi_i \big(A^{\dagger}A\big)_{ij} \geqslant \mu_A/2\,.$$

Proof. First we observe that

$$\pi_i \big(A^{\dagger}A\big)_{ij} = \sum_{k \in N} \pi_k A_{ki} A_{kj}\,.$$

Moreover, since A is a stochastic matrix, for every index i, we have either $\sum_{j \in N_2} A_{ij} \geqslant 1/2$ or $\sum_{j \in N_1} A_{ij} \geqslant 1/2$. That leads us to consider the two following cases.

1. There exists an index $i^* \in N_1$ such that $\sum_{j \in N_2} A_{i^* j} \geqslant 1/2$. Then we have

$$\sum_{i \in N_1,\, j \in N_2} \pi_i \big(A^{\dagger}A\big)_{ij} \geqslant \sum_{j \in N_2} \pi_{i^*} A_{i^* i^*} A_{i^* j} \geqslant \frac{\pi_{i^*} A_{i^* i^*}}{2}\,.$$

2. Otherwise for each index $i \in N_1$, we have $\sum_{j \in N_1} A_{ij} \geqslant 1/2$. It follows that

$$\sum_{i \in N_1,\, j \in N_2} \pi_i \big(A^{\dagger}A\big)_{ij} \geqslant \sum_{i \in N_1,\, j \in N_2} \sum_{k \in N_1} \pi_k A_{ki} A_{kj} = \sum_{k \in N_1,\, j \in N_2} \pi_k A_{kj} \sum_{i \in N_1} A_{ki}\,.$$

   Hence we have

$$\sum_{i \in N_1,\, j \in N_2} \pi_i \big(A^{\dagger}A\big)_{i,j} \geqslant \frac{1}{2} \sum_{i \in N_1,\, j \in N_2} \pi_i A_{ij}\,.$$

   By assumption, there exist $k_1 \in N_1$ and $k_2 \in N_2$ such that $A_{k_1 k_2} > 0$ and

$$\sum_{i \in N_1,\, j \in N_2} \pi_i \big(A^{\dagger}A\big)_{ij} \geqslant \frac{\pi_{k_1} A_{k_1 k_2}}{2}\,.$$

In both cases, we then derive that $\sum_{i \in N_1,\, j \in N_2} \pi_i \big(A^{\dagger}A\big)_{ij} \geqslant \frac{\mu_A}{2}$. □

Lemma 4 (Lemma 8 in [2])**.** *For every vector $x \in \mathbb{R}^n$, we have*

$$Q_A(x) \geqslant \frac{\mu_A}{2n} \big(\delta(x)\big)^2\,.$$

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

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

By reordering the terms in the last sum, we obtain

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

Then we use Lemma 3 to show that for each $d \in \{1, \ldots, n-1\}$, we have

$$\sum_{i=1}^{d} \sum_{j=d+1}^{n} \pi_i \big(A^\dagger A\big)_{ij} \geqslant \mu_A/2 .$$

Therefore

$$Q_A(x) \geqslant \frac{\mu_A}{2} \sum_{d=1}^{n-1} (x_{d+1} - x_d)^2 .$$

By Cauchy-Schwarz, we obtain

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

which completes the proof.

□

Lemma 5. *Each vector in the orthonormal basis $\{\varepsilon_2, \ldots, \varepsilon_n\}$ of $\Delta^\perp$ has a norm δ greater than 1, i.e.,*

$$\forall i \in \{2, \ldots, n\},\ \delta(\varepsilon_i) > 1 .$$

Proof. We denote by $\{e_1, \ldots, e_n\}$ the standard basis and we let

$$\varepsilon_i = u_{i,1} e_1 + \cdots + u_{i,n} e_n .$$

By definition of the inner product $\langle .,. \rangle_\pi$, we have $\|e_k\|_\pi^2 = \pi_k$ and $\langle e_k, e_\ell \rangle_\pi = 0$ when $k \neq \ell$. It follows that for every $i \in \{2, \ldots, n\}$,

$$\|\varepsilon_i\|_\pi^2 = u_{i,1}^2 \pi_1 + \cdots + u_{i,n}^2 \pi_n = 1$$

and

$$< \epsilon_i, \epsilon_1 >_\pi = < \epsilon_i, \mathbf{1} >_\pi = u_{i,1}\pi_1 + \cdots + u_{i,n}\pi_n = 0 .$$

From these equalities, we derive that there exist two different indices k_i and ℓ_i such that

$$|u_{k_i}| > 1 \quad \text{and} \quad u_{k_i} u_{\ell_i} < 0 ,$$

which implies

$$\delta(\varepsilon_i) > 1 .$$

□

Proposition 6. *If A is an ergodic stochastic matrix of size n with positive diagonal entries, then each singular value of A different from 1 is at most equal to $\sqrt{1 - \frac{\mu_A}{2n}}$.*

Proof. We easily check that

$$Q_A(\varepsilon_2) = 1 - (\sigma_2)^2 \ .$$

Besides combining Lemmas 4 and 5, we obtain

$$Q_A(\varepsilon_2) > \frac{\mu_A}{2n} \,,$$

which shows the upper bound on σ_2. □

1.4 Simpler proof and improvement in the case of a self-adjoint matrix

In the case A is a self-adjoint ergodic matrix, a simplified version of the above proof, in which Lemmas 4 and 5 are omitted, proves that any A's eigenvalue λ different from 1 satisfies

$$\lambda < 1 - \frac{\mu_A}{2D_A} \,, \tag{1}$$

where D_A denotes the diameter of the graph associated to A; see Chazelle [1].

Moreover λ lies within at least one Gershgorin disc $D(A_{ii}, 1 - A_{ii})$, i.e.,

$$-1 + 2A_{ii} \leqslant \lambda \leqslant 1 \ .$$

It then follows that

$$|\lambda| < \max\{1 - \frac{\mu_A}{2D_A}, 1 - 2\alpha_A\} \,,$$

where $\alpha_A = \min_{i=1,\ldots,n} (A_{i\,i})$. We easily check that $\mu_A \leqslant \alpha_A$, and thus

$$|\lambda| < 1 - \frac{\mu_A}{2D_A} \,.$$

If $\lambda_1(A) \ldots \lambda_n(A)$ denote A's eigenvalue ranged in the order of magnitude, then $\lambda_1(A) = 1$, and the second Weyl inequality gives

$$\sigma_2(A) \leqslant \lambda_2(A) \,.$$

The upper bound in Proposition 6 then can be easily improved into

$$\sigma_2(A) \leqslant 1 - \frac{\mu_A}{2n} \,.$$

2 A convergence theorem and its applications

We consider a sequence $\big(A_t\big)_{t\in\mathbb{N}}$ of $n \times n$ matrices that satisfies the following assumptions:

A1 : Every matrix A_t is an ergodic stochastic matrix with a positive diagonal.

A2: All the matrices A_t have the same Perron vector denoted by π:

$$\forall t \in \mathbb{N}, \quad \pi(A_t) = \pi \,.$$

A3: There exists some positive lower bound α on the positive entries of the matrices A_t:

$$\forall t \in \mathbb{N}, \forall (i,j) \in \{1,\ldots,n\}^2, \ (A_t)_{i\,j} \in \{0\} \cup [\alpha, 1] \,.$$

Let $x(0) \in \mathbb{R}^n$, and let $\big(x(t)\big)_{t\in\mathbb{N}}$ denote the sequence of vectors in $\mathbb{R}^n$ defined by:

$$x(t+1) = A_t\, x(t)\,. \tag{2}$$

Theorem 7. *Under assumptions A1-3, the sequence $\big(x(t)\big)_{t\in\mathbb{N}}$ converges to a vector x^* that is colinear to $\mathbf{1}$ and*

$$\lim_{t\to+\infty} \|x(t)-x^*\|^{1/t} \leqslant 1-\frac{\mu}{4n}\,,$$

where $\mu = \inf\big\{\pi_i\big(A_t\big)_{ij} : \big(A_t\big)_{ij} > 0\big\}$.

Proof. By (2), we have:

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

Since $A^\dagger_{t-1}$ is a stochastic matrix, $A^\dagger_{t-1}\,\mathbf{1} = \mathbf{1}$ and so

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

Therefore, the orthogonal projection of $x(t)$ on Δ is constant.

We let $a = \langle x(0), \mathbf{1}\rangle_\pi$ and

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

Then

$$V(t) = \|x(t)\|^2_\pi - 2a\langle x(t), \mathbf{1}\rangle_\pi + a^2\|\mathbf{1}\|^2_\pi = \|x(t)\|^2_\pi - a^2\,,$$

and

$$V(t) - V(t+1) = \langle x(t), x(t)\rangle_\pi - \langle A_t\, x(t), A_t\, x(t)\rangle_\pi = \langle x(t), x(t) - A^\dagger_t A_t\, x(t)\rangle_\pi\,.$$

By Corollary 2, it follows that V is non-increasing; we shall prove that $V(t)$ tends to 0.

By Proposition 6, A_t has n real singular values $\sigma_1, \ldots, \sigma_n$ that satisfy

$$0 \leqslant \sigma_n \leqslant \ldots \sigma_2 \leqslant \sqrt{1-\frac{\mu}{2n}} < \sigma_1 = 1$$

with $\mu = \min\big\{\pi_i\big(A_t\big)_{ij} : (i,j) \in E\big(G_t\big) \wedge t \in \mathbb{N}\big\}$. Let $\{\varepsilon_1, \ldots, \varepsilon_n\}$ be an orthonormal basis for the inner product $\langle ., .\rangle_\pi$ such that for each index $i \in [n]$,

$$A^\dagger_t A_t \varepsilon_i = \sigma_i^2 \varepsilon_i\,.$$

In particular, $\varepsilon_1 = \mathbf{1}$.

Let $z_1, \ldots, z_n$ the components of $x(t)$ in this basis[1], namely,

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

Hence $z_1 = \langle x(0), \mathbf{1}\rangle_\pi$ and so

$$V(t) = z_2^2 + \cdots + z_n^2\,.$$

Moreover, we have

$$A^\dagger_t A_t x(t) = z_1\varepsilon_1 + z_2\sigma_2^2\varepsilon_2 + \cdots + z_n\sigma_n^2\varepsilon_n$$

[1]The real numbers z_i, λ_i, and the vectors ε_i all depend on t, but t is not explicit in our notation as no confusion can arise.

and thus

$$V(t) - V(t+1) = z_2^2(1-\sigma_2^2) + \cdots + z_n^2(1-\sigma_n^2)\,.$$

Hence

$$V(t) - V(t+1) \geqslant (1-\sigma_2^2)V(t)\,.$$

From the upper bound on the second singular value in Proposition 6, it follows that

$$V(t) \leqslant \left(1 - \frac{\mu}{2n}\right)^t V(0)\,.$$

Hence $\lim_{t\to\infty} V(t) = 0$, and so

$$\lim_{t\to\infty} x(t) = a\,\mathbf{1}\,.$$

To complete the proof, we use the inequality $(1-u)^{1/2} \leqslant 1 - u/2$ which holds for any $u \in [0,1]$. □

We now consider a system with n agents $\{1,\ldots,n\}$, a local variable x_i for each agent i, and an averaging algorithm $\mathcal{A}$ with with non-vanishing and bounded weights, i.e., if G_t is the communication graph at round t and $x_i(t)$ is the value of x_i at the end of round t:

$$x_i(t) = \sum_{k\in In_i(G_t)} w_{ik}(t)x_k(t-1)\,, \tag{3}$$

with a positive lower bound on all the weights $w_{ik}(t)$.

We consider an execution of $\mathcal{A}$ with an initial state $x(0) \in \mathbb{R}^n$ and a communication pattern (sequence of communication graphs) $\left(G_t\right)_{t\in\mathbb{N}}$. We say that $\mathcal{A}$ *achieves asymptotic consensus* in this execution if the sequence $x(t)$ converges to a vector x^* that is colinear to $\mathbf{1}$. The *convergence rate* is then defined as

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

where $\|.\|$ is any norm on $\mathbb{R}^n$, and the *convergence time* is

$$T(\varepsilon) = \inf\{\tau \in \mathbb{N} \;:\; \forall t \geqslant \tau,\; V(t) \leqslant \varepsilon V(0)\,.$$

Let A_t denote the stochastic matrix associated to the update rule (3) at round t in this execution. The central assumption is that all the stochastic matrices A_t share the same Perron vector π:

$$\forall t \in \mathbb{N},\quad \pi(A_t) = \pi\,.$$

We denote

$$\mu = \inf\left\{\pi_i\left(A_t\right)_{ij} : \left(A_t\right)_{ij} > 0\right\}.$$

From Theorem 7, we immediately deduce the following corollary.

Corollary 8. *The algorithm $\mathcal{A}$ achieves asymptotic consensus with convergence rate $\varrho \leqslant 1 - \frac{\mu}{4n}$ and convergence time $T(\varepsilon) \leqslant \frac{2n}{\mu}\log\left(1/\varepsilon\right)$.*

In the *reversible case* where every matrix A_t is self-adjoint, the remark in Section 1.4 provides a better upper bound on the second singular value of A_t, and leads to the following improvements on the convergence rate and convergence time.

Corollary 9. *In the reversible case, the algorithm $\mathcal{A}$ achieves asymptotic consensus with convergence rate $\varrho \leqslant 1 - \frac{\mu}{2n}$ and convergence time $T(\varepsilon) \leqslant \frac{n}{\mu}\log\left(1/\varepsilon\right)$.*

2.1 Applications and previous results

Corollary 8 has several applications to (1) the EqualNeighbor algorithm with a fixed bidirectional topology, (2) the FixedWeight algorithm with a dynamic bidirectional topology, and finally the Metropolis algorithm with a dynamic bidirectional topology.

2.1.1 EqualNeighbor algorithm with a fixed bidirectional topology

With the EqualNeighbor algorithm, we have

$$w_{ik}(t) = \frac{1}{d_i(t)} ,$$

where $d_i(t)$ is the number of i's in-neighbors in G_t. If G_t is bidirectional, then the i-th entry of the Perron vector of A_t is equal to:

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

where $d(t) = \sum_{i=1}^{n} d_i(t)$. Therefore $\mu \geqslant 1/n^2$. Moreover, the stochastic matrix A associated in each round is self-adjoint.

Corollary 9 then applies to any execution of the EqualNeighbor algorithm with a fixed bidirectional topology.

Theorem 10 ([3])**.** *In any execution with a fixed bidirectional connected topology, the EqualNeighbor algorithm achieves asymptotic consensus with convergence rate* $\varrho \leqslant 1 - \frac{1}{2n^3}$ *and convergence time* $T(\varepsilon) \leqslant n^3 \log\big(1/\varepsilon\big)$.

2.1.2 FixedWeight algorithm with a dynamic bidirectional topology

For each agent i, let q_i denote an upper bound on the number of i's in-neighbors in a given communication pattern (G_t). Weights in the FixedWeight algorithm are given by:

$$w_{ik}(t) = \begin{cases} 1/q_i & \text{if } \ j \in N_i^+(t) \setminus \{i\} \\ 1 - \left(d_i(t) - 1\right)/q_i & \text{if } \ j = i \\ 0 & \text{otherwise } . \end{cases}$$

We easily check that if G_t is bidirectional, then the i-th entry of the Perron vector of A_t is equal to:

$$\pi_i(A_t) = \frac{q_i}{q} ,$$

where $q = \sum_{i=1}^{n} q_i$.

It then follows that in any bidirectional communication pattern, the Perron vector is constant and $\mu = 1/q \geqslant 1/n^2$. Moreover, for every round t, the stochastic matrix A_t is self-adjoint.

Corollary 9 then applies to any execution of the FixedWeight algorithm with a bidirectional communication pattern.

Theorem 11 ([1])**.** *In any execution with a communication pattern composed of bidirectional connected communication graphs, the FixedWeight algorithm achieves asymptotic consensus with convergence rate* $\varrho \leqslant 1 - \frac{1}{2n^3}$ *and convergence time* $T(\varepsilon) \leqslant n^3 \log\big(1/\varepsilon\big)$.

2.1.3 Metropolis algorithm with a dynamic bidirectional topology

With the Metropolis algorithm, we have

$$w_{ik}(t) = \begin{cases} \frac{1}{\max(d_i(t), d_j(t))} & \text{if } \ j \in N_i^+(t) \setminus \{i\} \\ 1 - \sum_{j \in N_i^+(t)\setminus\{i\}} \frac{1}{\max(d_i(t), d_j(t))} & \text{if } \ j = i \\ 0 & \text{otherwise } \ . \end{cases}$$

We easily check that if G_t is bidirectional, then each matrix A_t is symmetric, and so doubly stochastic. It follows that the Perron vector of A_t is collinear to $\mathbf{1}$, its i-th entry is equal to:

$$\pi_i(A_t) = 1/n\,,$$

and $\mu \geqslant 1/n^2$. Moreover A_t is self-adjoint for the inner product $\langle .,. \rangle_\pi$.

Corollary 9 then applies to any execution of the Metropolis algorithm with a bidirectional communication pattern.

Theorem 12 ([2])**.** *In any execution with a communication pattern composed of bidirectional connected communication graphs, the Metropolis algorithm achieves asymptotic consensus with convergence rate* $\varrho \leqslant 1 - \frac{1}{2n^3}$ *and convergence time* $T(\varepsilon) \leqslant n^3 \log\big(1/\varepsilon\big)$*.*

References

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

[2] 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.

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