Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Tight estimates for non-stationary consensus with fixed underlying spanning tree

David Angeli [*] **Pierre-Alexandre Bliman** [**]

[*] *Dipartimento di Sistemi e Informatica, University of Florence, Via di S. Marta 3, 50139 Firenze, Italy and INRIA, Rocquencourt BP105, 78153 Le Chesnay cedex, France. Email:* angeli@dsi.unifi.it
[**] *INRIA, Rocquencourt BP105, 78153 Le Chesnay cedex, France. Email:* pierre-alexandre.bliman@inria.fr

Abstract: This article is devoted to estimating the speed of convergence towards consensus for a general class of discrete-time multi-agent systems. In the systems considered here, both the topology of the interconnection graph and the weight of the arcs are allowed to vary as a function of time. Under the hypothesis that some spanning tree structure is preserved along time, and that some nonzero minimal weight of the information transfer along this tree is guaranteed, an estimate of the contraction rate is given. The latter is expressed explicitly as the spectral radius of some matrix depending upon the tree depth and the lower bounds on the weights.

Keywords: Cooperative systems; multi-agent systems; rate of convergence; time-varying systems; robustness stability.

1. INTRODUCTION

Appeared in the areas of communication networks, control theory and parallel computation, the analytical study of ways for reaching consensus in a population of agents is a problem of broad interest in many fields of science and technology; see Angeli et al. (2006b) for references. Of particular interest is the question of estimating how quickly consensus is reached on the basis of few qualitative (mainly topological) information as well as basic quantitative information on the network (mainly the strength of reciprocal influences).

Originally, this problem was considered in the context of stationary networks. For Markov chains that are homogeneous (that is stationary in the vocabulary of dynamical systems), it amounts to quantify the speed at which steady-state probability distribution is achieved, and is therefore directly related to finding an a priori estimate to the second largest eigenvalue of a stochastic matrix. Classical works on this subject are due to Cheeger (1969); Diaconis et al. (1991); see also Rosenthal (1995) for a survey on improved bounds. The latter concern reversible Markov chains, for example when the transition matrix is symmetric, see e.g. Fill (1991) for the non-reversible case.

Among the classical contributions which instead deal with time-varying interactions we refer to the work of Cohn (1989), where asymptotic convergence is proved, but neglecting the issue of relating topology and guaranteed convergence rates. Tsitsiklis *et al.* also provided important qualitative contributions to this subject, see Tsitsiklis (1984, 1986); Bertsekas et al. (1989), as well as Moreau (2005). See also Angeli et al. (2006a) for further nonlinear results. In particular, the role of connectivity of the communication graph in the convergence of consensus and spanning trees has been recognised and finely analysed Moreau (2005); Cao et al. (2005); Olshevsky et al. (2006).

More recently, important contributions in characterizing convergence to consensus in a time-varying set-up were proven by several authors, see for instance Bertsekas et al. (1989); Moreau (2005).

In a previous paper Angeli et al. (2006b), several criteria were provided to estimate quantitatively the contraction rate of a set of agents towards consensus, in a discrete time framework. The attempt there consisted in following the spread of the information over the agent population, along one or more spanning-trees. Ensuring a lower bound to the matrix entries of the agents already attained by the information flow along the spanning-tree, rather than the nonzero contributions as classically, permitted to obtain tighter estimates with weaker assumptions. Distinguishing between different sub-populations, of agents already touched by spanning-tree and agents not yet attained, and using lower bounds on the influence of the former ones, one is able to establish rather precise convergence estimates.

As a matter of fact, rapid consensus can be obtained in two quite different ways — either by dense and isotropic communications (based, say, on a complete graph), or by very unsymmetric and sparse relations (with a star-shaped graph with a leading root). In the first case many spanning trees cover the graph, while in the second configuration a unique one does the job.

The present article is a continuation of Angeli et al. (2006b). Emphasis is put on propagation of a unique spanning tree and on the resulting consequences in terms of convergence speed. It is demonstrated that in the particular case where such a spanning tree structure is guaranteed to exist at any time, ensuring minimal weight to the transmission of information along the tree (from the root to the leafs) indeed enforces some minimal convergence rate, whose expression is particularly simple. A worst-case estimate is provided, expressed as the spectral radius of certain matrix whose size equals the depth of the tree and whose coefficients depend in a simple way of the assumed minimal weights. This results in a sensible improvement over existing evaluations.

The paper is organized as follows. Section 2 contains the problem formulation and a presentation of the main result, together with the minimal amount of technical tools to allow for its comprehension. A comparison system is introduced afterwards in Section 3, whose study is central to establish the convergence estimate. The original method for analysis of this system is used in Section 4 to get convergence rate estimate (therein is stated the main result of the paper, Theorem 2), and some properties of the latter are studied. This result is commented in Section 5, before some concluding remarks. For space reasons, most proofs are omitted, they may be consulted in the complete version of this text, see Angeli et al. (2007).

Notations

The i-th vector of the canonical basis in the space $\mathbb{R}^n$ ($1 \leq i \leq n$) is denoted e_i^n; the vector with all components equal to 1 in $\mathbb{R}^n$ is written $\mathbf{1}^n$. When the context is clear, we omit the exponent and just write e_i, resp. $\mathbf{1}$ to facilitate reading. We also use brackets to select components of vectors. All these notation are standard, and for a vector $x \in \mathbb{R}^n$, the i-th component is written alternatively x_i, $[x]_i$, $(e_i^n)^\mathsf{T} x$ or $e_i^\mathsf{T} x$.

The systems considered here will be composed of n agents: accordingly, we let $\mathcal{N} \doteq \{1, \ldots, n\}$.

As usual, identity and zero square matrices of dimension $q \times q$ are denoted I_q and 0_q respectively. We denote J_q the $q \times q$ matrix with ones on the sub-diagonal and zeros otherwise: $(J_q)_{i,j} = \delta_{i=j+1}$. Here, and later in the text, δ denotes the Kronecker symbol, equal to 1 (resp. 0) when the condition written in the subscript is fulfilled (resp. is not). For self-containedness, recall that a real square matrix M is said stochastic (row-stochastic) if it is nonnegative with each row sum equal to 1.

The spectral radius of a square matrix M is denoted $\lambda_{\max}(M)$. Last, we use the notion of nonnegative matrices, meaning real matrices which are componentwise nonnegative. Accordingly, the order relations $\leq$ and $\geq$ envisioned for matrices are meant componentwise.

2. PROBLEM FORMULATION AND PRESENTATION OF THE MAIN RESULT

Our aim is to estimate the speed of convergence towards consensus for the following class of time-varying linear systems:

$$x(t+1) = A(t)x(t) \tag{1}$$

where $A(t) \doteq (a_{i,j}(t))_{(i,j)} \in \mathbb{R}^{n \times n}$ is a sequence of stochastic matrices (in particular, $A(t)\mathbf{1} = \mathbf{1}$; this is exactly the dual of what happens in the case of non-homogeneous Markov chains, where the probability distribution, written as a row vector $\pi(t)$, verifies rather a relation like $\pi(t+1) = \pi(t)A(t)$).

Let us first introduce some technical vocabulary to present in simplest terms the main result of the paper, afterwards enunciated in Section 4. The definition of the quantity we intend to estimate is as follows.

Definition 1. (Contraction rate). We call *contraction rate* of system (1) the number $\rho \in [0, 1]$ defined as:

$$\rho \doteq \sup_{x(0)} \limsup_{t \to +\infty} \left(\frac{\max_{i \in \mathcal{N}} x_i(t) - \min_{i \in \mathcal{N}} x_i(t)}{\max_{i \in \mathcal{N}} x_i(0) - \min_{i \in \mathcal{N}} x_i(0)} \right)^{\frac{1}{t}},$$

where the supremum is taken on those $x(0)$ for which the denominator is nonzero. □

The contraction rate is thus related to the speed of convergence to zero of the agent set diameter. In what follows, the latter plays the role of a Lyapunov function to study convergence to agreement. For stationary systems, as is well known, the number ρ is indeed the second largest eigenvalue of the matrix A. More in general, it corresponds to the second largest Lyapunov exponent of the considered sequence of matrices $A(t)$.

Definition 2. (Communication graph). We call *communication graph* of system (1) *at time* t the directed graph defined by the ordered pairs $(j, i) \in \mathcal{N} \times \mathcal{N}$ such that $a_{i,j}(t) > 0$. □

In the present context, we use indifferently the terms “node” or “agent”.

We now introduce assumptions on the existence of a constant hierarchical structure embedded in the communication graph, and on minimal weights attached to the corresponding links.

Assumption 1. For a given positive integer $T_d > 0$, called the *depth of the communication graph*, assume the existence of nested sets $\mathcal{N}_0, \ldots, \mathcal{N}_{T_d}$ such that

- $\mathcal{N}_0$ is a singleton (whose element is called the *root*);
- $\mathcal{N}_k \subset \mathcal{N}_{k+1}$;
- $\mathcal{N}_{T_d} = \mathcal{N} = \{1, \ldots, n\}$.

Assume in addition, for given nonnegative real numbers α, β, γ, that, for all $t \geq 0$ and all $k \in \{1, 2, \ldots, T_d\}$

$$a_{i,i}(t) \geq \alpha \quad \text{if } i \in \mathcal{N}_0, \tag{2a}$$

$$\sum_{j \in \mathcal{N}_k \setminus \mathcal{N}_{k-1}} a_{i,j}(t) \geq \beta \quad \text{if } i \in \mathcal{N}_k \setminus \mathcal{N}_{k-1}, \tag{2b}$$

$$\sum_{j \in \mathcal{N}_{k-1}} a_{i,j}(t) \geq \gamma \quad \text{if } i \in \mathcal{N}_k \setminus \mathcal{N}_{k-1}. \tag{2c}$$

□

As an example, the sets $\mathcal{N}_k$ may be induced by some fixed spanning tree embedded in the communication graph: the existence of a distinguished agent, the root, is presupposed and, although the matrices $A(t)$ and the underlying communication graphs are allowed some variations, information progress from this root along a (time-varying) tree to attain all the agents. The number T_d bounds from above the minimal time for the information to attain the most distant agents from the root. Likely, we call

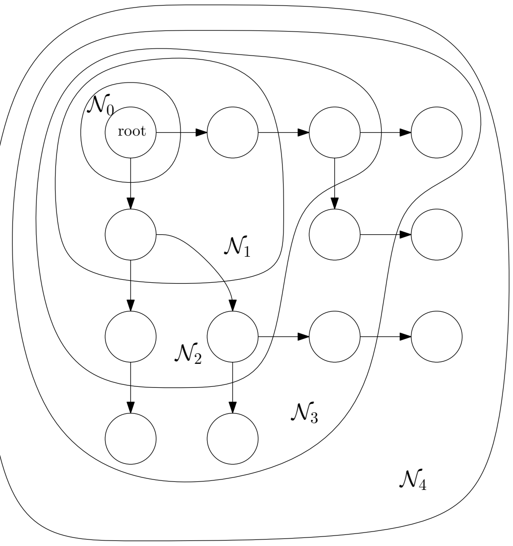


Fig. 1. The nested sets and the spanning tree

$d_i \doteq \min\{k \ : \ i \in \mathcal{N}_k\}$ the *depth of agent i*. The set $\mathcal{N}_k$ indeed consists of all the agents i whose depth d_i is guaranteed by Assumption 1 to be at most equal to k. An example of (fixed) communication graph and the associated nested sets is shown in Figure 1.

In addition to the spanning tree structure, Assumption 1 imposes some minimal weights to the information transmitted downstream along this structure (this is the role played by γ), and also to the information used between agents located at same depth. Concerning the latter, expressed by condition (2b), remark that it is fulfilled by self-loops, that is when

$$a_{i,i}(t) \geq \beta \quad \text{if } d_i > 0$$

(because by definition, $i \in \mathcal{N}_{d_i} \setminus \mathcal{N}_{d_i-1}$ for $d_i > 0$); but it is indeed weaker: it allows just as well communications between agents whose depths are equal. The constraint on the self-loops of the root agent, measured by α, is different than for the other agents (β); this is done on purpose, and permits to treat simultaneously the case of leaderless coordination and 'pure' coordination with a leader (case corresponding to $\alpha = 1$).

Last, notice that, the matrices $A(t)$ being stochastic, one should have:

$$\alpha, \ \beta + \gamma \leq 1$$

for Assumption 1 to be fulfilled.

We are now in position to present the contents of Theorem 2. The latter states that, under the conditions exposed above, *the rate of convergence of system* (1) *is* at most *equal to the spectral radius of the* $T_d \times T_d$ *matrix* $\zeta_{T_d}(\alpha, \beta, \gamma)$ *defined by*

$$\zeta_{T_d}(\alpha,\beta,\gamma) = \begin{pmatrix} \beta^\star & 0 & \dots & 0 & 1-\alpha^\star-\beta^\star \\ \alpha^\star & \beta^\star & \ddots & \vdots & 1-\alpha^\star-\beta^\star \\ 0 & \alpha^\star & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & \beta^\star & 1-\alpha^\star-\beta^\star \\ 0 & \dots & 0 & \alpha^\star & 1-\alpha^\star \end{pmatrix},$$

with $\alpha^\star = \min\{\alpha, \gamma\}$, $\beta^\star = \min\{\beta + \gamma, \alpha\} - \alpha^\star$. A major characteristic of this estimate is that it is *independent* of the number n of agents: it only depends upon α, β, γ and the depth T_d.

3. A COMPARISON SYSTEM FOR THE DIAMETERS EVOLUTION

We now build an auxiliary time-varying system, with a simpler structure than (1), and with the property that the asymptotic contraction rate of the original system can be bounded from above by carrying out suitable computations on this newly introduced system. Our main result for the present section (demonstrated in Angeli et al. (2007)) is a statement relating convergence of (1) towards consensus of a comparison system introduced below.

Theorem 1. Assume system (1) fulfills Assumption 1, for given nonnegative numbers α, β, γ (such that $\alpha, \beta+\gamma \leq 1$). Let $\Delta(t)$ be defined by

$$\Delta(t) \doteq \begin{pmatrix} \max_{i\in\mathcal{N}_0} x_i(t) - \min_{i\in\mathcal{N}_0} x_i(t) \\ \max_{i\in\mathcal{N}_1} x_i(t) - \min_{i\in\mathcal{N}_1} x_i(t) \\ \vdots \\ \max_{i\in\mathcal{N}_{T_d}} x_i(t) - \min_{i\in\mathcal{N}_{T_d}} x_i(t) \end{pmatrix}.$$

Then, $\Delta(t)$ satisfies the following inequality:

$$\Delta(t+1) \leq \left(\begin{array}{c|c} 1 & 0_{T_d\times 1} \\ \hline \alpha^\star & \\ 0_{(T_d-1)\times 1} & \zeta_{T_d}(\alpha^\star,\beta^\star) \end{array}\right) \Delta(t) \ , \qquad (3)$$

where $\zeta_{T_d}(\alpha^\star, \beta^\star) \in \mathbb{R}^{T_d\times T_d}$ is defined by:

$$\zeta_{T_d}(\alpha^\star,\beta^\star) \doteq (1-\alpha^\star-\beta^\star)\ \mathbf{1}e_{T_d}^{\mathsf{T}} + \beta^\star I_{T_d} + \alpha^\star J_{T_d}, \quad (4)$$

$$\alpha^\star \doteq \min\{\alpha,\gamma\}, \quad \beta^\star \doteq \min\{\beta+\gamma,\alpha\} - \alpha^\star \ . \quad (5)$$

□

Recall that inequality (3) is meant componentwise.

Remark 1. Two special cases of interest as far as application of Theorem 1 are obtained for the following values of parameters:

(1) $\alpha = 1$: viz. communication graph admits a *leader*; under such premises, expressions for $\alpha^\star$ and $\beta^\star$ simplify as follows:
$$\alpha^\star = \gamma \qquad \beta^\star = \beta$$
(2) $\alpha = \beta$, viz. root agent is not different from any other member of the group in terms of self-confidence on his own position in the formation of consensus:
$$\alpha^\star = \min\{\beta,\gamma\} \qquad \beta^\star = \max\{0, \beta-\gamma\}$$

□

4. CONVERGENCE RATE ESTIMATE AND PROPERTIES

Based on Theorem 1, we now provide Theorem 2, which states properly the property announced in the beginning of the paper.

Theorem 2. Consider the linear time-varying dynamical system (1), with $A(t)$ stochastic. Assume Assumption 1 is fulfilled. Then, the contraction rate towards consensus can be bounded according to the following formula:

$$\rho \leq \rho_{T_d}(\alpha^\star, \beta^\star) \doteq \lambda_{\max}(\zeta_{T_d}(\alpha^\star, \beta^\star)) \ , \tag{6}$$

with $\zeta_{T_d}, \alpha^\star, \beta^\star$ given in (4) and (5). □

Theorem 2 is demonstrated in Appendix. Recall that stochasticity of $A(t)$ implies that the nonnegative scalar α, β, γ verify: $\alpha \leq 1$, $\beta + \gamma \leq 1$.

Theorem 2 provides a tight estimate for the contraction rate of (1) on the basis of the parameters α, β and γ, and of the depth T_d of the sequence of tree matrices. We emphasize the fact that the result holds for time-varying systems. Indeed, Theorem 2 is an inherently robust result, as Assumption 1 allows for much uncertainty in the definition of system (1). This robustness is meant with respect to variations of the communication graph (provided these variations don't violate the set conditions of Assumption 1), and with respect to variations of the coefficients of the matrix $A(t)$ (provided they respect the quantitative constraints in Assumption 1).

A central fact is that the value in (6) does not depend upon the number of agents involved in the network: rather the depth of the graph is involved, which is quite natural.

Some properties of the estimate are now given in Theorems 3 and 4 (see Angeli et al. (2007) for a proof). They are indeed useful to have a grasp on the asymptotic behaviour of the contraction estimate, as well as on their monotonicity properties; the latter are in agreement with the increase of decrease of information available by varying the parameters α, β and γ.

Theorem 3. Let $\alpha^\star, \beta^\star \in (0, 1]$. Then for any $T \in \mathbb{N}$, $\rho_T(\alpha^\star, \beta^\star) = \lambda_{\max}(\zeta_T(\alpha^\star, \beta^\star))$ has the following properties.

- $\rho_T(\alpha^\star, \beta^\star)$ is the largest *real* root of the polynomial equation
$$\left(\frac{s-\beta^\star}{\alpha^\star}\right)^T + \left(\frac{s-\beta^\star}{\alpha^\star}\right)^{T-1} + \cdots + \frac{s-\beta^\star}{\alpha^\star} + 1$$
$$= \left(\frac{1-\beta^\star}{\alpha^\star}\right)\left(\left(\frac{s-\beta^\star}{\alpha^\star}\right)^{T-1} + \cdots + \frac{s-\beta^\star}{\alpha^\star} + 1\right) .$$
- For any $T \in \mathbb{N}$, $\rho_T(\alpha^\star, \beta^\star) \leq \rho_{T+1}(\alpha^\star, \beta^\star)$.
- For any $T \in \mathbb{N}$, $1 - \alpha^\star, \beta^\star < \rho_T(\alpha^\star, \beta^\star) < 1$.
- $\rho_T(\alpha^\star, \beta^\star) \leq \alpha^\star + \beta^\star$ if and only if $T \leq \frac{\alpha^\star}{1-\alpha^\star-\beta^\star}$.
- $\rho_T(\alpha^\star, \beta^\star) \to 1$ when $T \to +\infty$, and more precisely
$$1 - \rho_T(\alpha^\star, \beta^\star)$$
$$= (1-\alpha^\star-\beta^\star)\left(\frac{\alpha^\star}{1-\beta^\star}\right)^T + o\left(\left(\frac{\alpha^\star}{1-\beta^\star}\right)^T\right) .$$

□

The following result studies the variation of ρ_T as a function of α, β, γ. When considering ρ_T as a function of these quantities, we write $\rho_T(\alpha, \beta, \gamma)$, meaning $\rho_T(\alpha^\star, \beta^\star)$ for $\alpha^\star(\alpha, \beta, \gamma), \beta^\star(\alpha, \beta, \gamma)$ defined as in (5).

Theorem 4. For any $T \in \mathbb{N}$,

- the function $(\alpha^\star, \beta^\star) \mapsto \rho_T(\alpha^\star, \beta^\star)$ is nonincreasing on the set $\{(\alpha^\star, \beta^\star) \in [0, 1]^2 \ : \ \alpha^\star + \beta^\star \leq 1\}$;

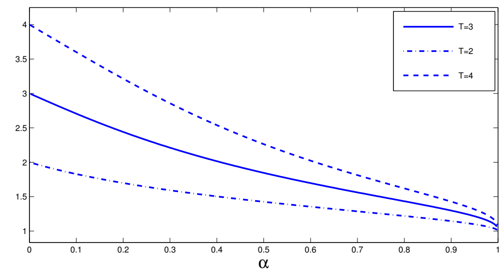


Fig. 2. Ratios between spectral gaps

- the function $(\alpha, \beta, \gamma) \mapsto \rho_T(\alpha, \beta, \gamma)$ is nonincreasing on the set $\{(\alpha, \beta, \gamma) \in [0, 1]^3 \ : \ \beta + \gamma \leq 1\}$;
- if $\beta + \gamma = \beta' + \gamma'$, then $\rho_T(\alpha, \beta, \gamma) \leq \rho_T(\alpha, \beta', \gamma')$ when $\beta \geq \beta'$.

Moreover, for any $T \in \mathbb{N}$,

- $\rho_T(\alpha, \beta, \gamma) = 1$ if and only if $\alpha = 0$ or $\gamma = 0$.
- $\rho_T(\alpha, \beta, \gamma) = \beta = 1 - \gamma$ if and only if $\alpha = \beta + \gamma = 1$.

□

Notice that the estimates given in the last two points of Theorem 4 are tight: they are reached for the following *stationary* systems:

$$\begin{aligned} \text{Case } \alpha = 0\text{:} \quad & A = J_n + e_1 e_n^\mathsf{T} \\ \text{Case } \gamma = 0\text{:} \quad & A = I_n \\ \text{Case } \alpha = \beta + \gamma = 1\text{:} \quad & A = \beta I_n + (1-\beta)(J_n + e_1 e_1^\mathsf{T}) \end{aligned}$$

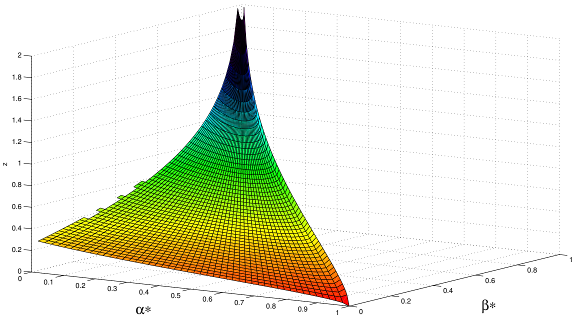
5. DISCUSSION AND INTERPRETATION OF THE RESULTS

It is interesting to compare our results with the classical estimate $\rho \leq \sqrt[T_d]{1-\alpha^{T_d}}$ which is obtained by assuming a lower-bound α on the diagonal entries as well as on the non-zero entries of $A(t)$. In our set-up this is obtained by letting $\alpha = \gamma = \beta = \alpha^\star$ and $\beta^\star = 0$. In order to have an idea on the quality of the two estimates, we plot the ratio of the *spectral gaps*,

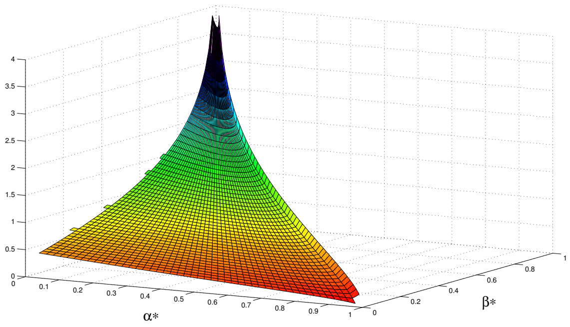
$$\frac{1 - \rho_{T_d}(\alpha^\star, 0)}{1 - \sqrt[T_d]{1 - {\alpha^\star}^{T_d}}}$$

for $T_d = 2, 3, 4$ in Fig. 2. As it is possible to see, the new estimates are consistently tighter than the classic ones; in the best case, viz. for $\alpha^\star \approx 0$, the ratio of spectral gaps approaches T_d. So, the quality of the estimates actually improves with respect to the classic bound, as the horizon T_d increases.

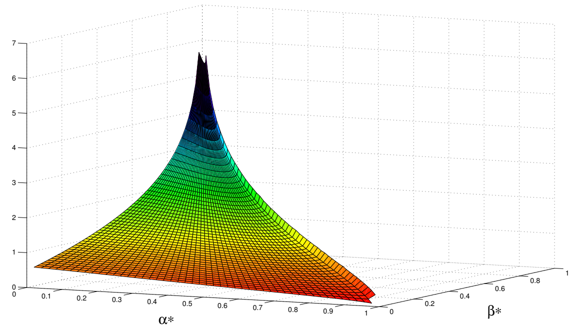
When additional information is available, for instance when the coefficient α, β, γ as given in (2) are known, then contraction rate estimates become much tighter with respect to their classical counterparts which are not able to discriminate between inner loops of the root node and inner-loops of individual agents, as well as strength of inter-agent communication links. In order to carry out a comparison, notice that under the assumption of a prescribed α, β, γ tree matrix bounding from below $A(t)$, we may assume for the classical estimate the following value of $\alpha := \min\{\alpha, \beta, \gamma\}$ which indeed is always smaller



(a)



(b)



(c)

Fig. 3. Ratios of spectral gap: (a) $T = 2$, (b) $T = 3$, (c) $T = 4$. The vertical axis is graduated in a $\log_{10}$ scale.

than $\alpha^\star = \min\{\alpha, \gamma\}$. Hence, the corresponding spectral gaps satisfy:

$$1 - \sqrt[T_d]{1 - \min\{\alpha, \beta, \gamma\}^{T_d}} \leq 1 - \sqrt[T_d]{1 - \min\{\alpha, \gamma\}^{T_d}}$$

so that, we may compare the classical estimate with the new one by considering the following ratios:

$$\frac{1 - \rho_{T_d}(\alpha^\star, \beta^\star)}{1 - \sqrt[T_d]{1 - \min\{\alpha, \beta, \gamma\}^{T_d}}} \geq \frac{1 - \rho_{T_d}(\alpha^\star, \beta^\star)}{1 - \sqrt[T_d]{1 - \alpha^{\star T_d}}}$$

We plotted the function at the right-hand side of the previous inequality in a $\log_{10}$ scale as a function of $\alpha^\star$ and $\beta^\star$. In general the ratio depends critically on the tree depth T_d, hence we only plot it for relatively small tree depths. In particular the results shown in Fig. 3 were obtained. Notice that the relative quality of the estimates again increases with T_d, and already for $T_d = 4$ a significant portion of parameters space lies in the area in which estimates differ by a 10^4 factor. The dependence of ρ_{T_d} upon $\alpha^\star$ and $\beta^\star$ is shown in Fig. 4 for $T = 2, 3, 4$. This also clearly shows the different monotonicity properties highlighted in the previous Section.

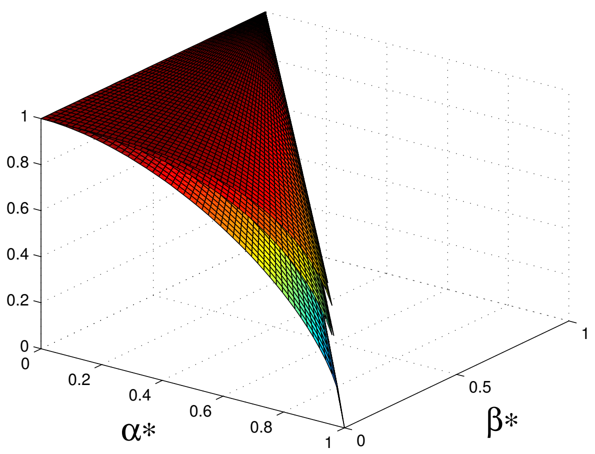


Fig. 4. The function $\rho_{T_d}(\alpha^\star, \beta^\star)$ for $T_d = 2, 3, 4$ (from bottom to top)

.

6. CONCLUSION

We provide a novel and tight estimate of the contraction rate of infinite products of stochastic matrices, under the assumption of prescribed lower bounds on the influence between different sets of agents which naturally arise by following the information spread along the interaction graph. This improves previously known bounds and, when additional information is assumed, exploits the additional structure for tightening of several orders of magnitude the previously available estimates. The other crucial factor in determining the overall convergence rate is the time T_d needed to the information to propagate from some root node (which may or may not play the role of a leader) to the other nodes. The bound can be computed as the Perron-Frobenius eigenvalue (the spectral radius) of a positive T_d-dimensional matrix, whose entries depend in a relatively simple way on the parameters characterizing the hypothetic lower bounds available. Some monotonicity and asymptotic properties of the bound are also proved.

REFERENCES

D. Angeli, P.-A. Bliman (2006a). Stability of leaderless discrete-time multi-agent systems, *Mathematics of Control, Signals and Systems* **18** no 4 (2006) 293–322

D. Angeli, P.-A. Bliman (2006b). Convergence speed of unsteady distributed consensus: decay estimates along the settling spanning trees, available at `arxiv.org/abs/math.OC/0610854`

D. Angeli, P.-A. Bliman (2007). Tight estimates for convergence of some non-stationary consensus algorithms, available at `arxiv.org/abs/0706.0630`

D.P. Bertsekas, J.N. Tsitsiklis (1989). *Parallel and distributed computation*, Prentice-Hall International; also downloadable at `https://dspace.mit.edu/handle/1721.1/3719`

M. Cao, D.A. Spielman, A.S. Morse (2005). A Lower Bound on Convergence of a Distributed Network Consensus Algorithm. In *Proceedings of the Joint European*

Control Conference/IEEE Conference on Decision and Control, Sevilla, Spain

J. Cheeger (1969). A lower bound for the smallest eigenvalue of the Laplacian, *Problems in Analysis, Papers dedicated to Salomon Bochner*, Princeton University Press, Princeton, 195–199

H. Cohn (1989). Products of stochastic matrices and applications, *Int. Journal of Mathematics and Mathematical Science* **12** no 2, 209–233

P. Diaconis, D. Stroock (1991). Geometric bounds for eigenvalues of Markov chains, *Annals of Applied Probability* **1**, 36–61

J.A. Fill (1991). Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process, *Ann. Appl. Probab.* Vol. 1 no 1, 62–87.

L. Moreau (2005). Stability of multi-agent systems with time-dependent communication links, *IEEE Trans. Automat. Control* **50** (2), 169–182

A. Olshevsky, J.N. Tsitsiklis (2006). Convergence speed in distributed consensus and averaging, In *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, California

J.S. Rosenthal (1995). Convergence rates of Markov chains, *SIAM Review* **37** (3), 387–405

J.N. Tsitsiklis (1984). *Problems in Decentralized Decision Making and Computation*, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology; available at `http://hdl.handle.net/1721.1/15254`

J.N. Tsitsiklis, D.P. Bertsekas, M. Athans (1986). Distributed Asynchronous Deterministic and Stochastic Gradient Optimization Algorithms, *IEEE Trans. Automat. Control* **31** no 9, 803–812

Appendix A. PROOF OF THEOREM 2

All the factors in (3) being nonnegative, the order relation is compatible with multiplication. One then obtains, for all $t \in \mathbb{N}$,

$$\Delta(t) \leq \left(\begin{array}{c|c} 1 & 0_{T_d\times 1} \\ \hline \alpha^\star & \\ 0_{(T_d-1)\times 1} & \zeta_{T_d}(\alpha^\star,\beta^\star) \end{array}\right)^t \Delta(0)$$

$$\leq \left(\begin{array}{c|c} 1 & 0_{T_d\times 1} \\ \hline \alpha^\star & \\ 0_{(T_d-1)\times 1} & \zeta_{T_d}(\alpha^\star,\beta^\star) \end{array}\right)^t \begin{pmatrix} 0 \\ \mathbf{1}^{T_d} \end{pmatrix} \Delta_{T_d+1}(0) \ ,$$

where the fact that $\Delta_1(t) \equiv 0 \leq \Delta_k(t) \leq \Delta_{k+1}(t) \leq \Delta_{T_d+1}(t)$, $1 \leq k \leq T_d+1$, have been taken into account.

One deduces that

$$\Delta(t) \leq \begin{pmatrix} 0_{T_d\times 1} \\ \zeta_{T_d}(\alpha^\star,\beta^\star)^t \end{pmatrix} \mathbf{1}^{T_d} \Delta_{T_d+1}(0)$$

and

$$\Delta_{T_d+1}(t) = e_{T_d+1}^{(T_d+1)\mathsf{T}} \Delta(t) \leq e_{T_d}^{T_d\mathsf{T}} \zeta_{T_d}(\alpha^\star,\beta^\star)^t \mathbf{1}^{T_d} \Delta_{T_d+1}(0) \ ,$$

from which it ensues

$$\limsup_{t\to+\infty} \left(\frac{\Delta_{T_d+1}(t)}{\Delta_{T_d+1}(0)}\right)^{1/t} \leq \limsup_{t\to+\infty} \left(e_{T_d}^{T_d\mathsf{T}} \zeta_{T_d}(\alpha^\star,\beta^\star)^t \mathbf{1}^{T_d}\right)^{1/t}$$

$$\leq \limsup_{t\to+\infty} \left\|\zeta_{T_d}(\alpha^\star,\beta^\star)^t\right\|^{1/t} = \lambda_{\max}(\zeta_{T_d}(\alpha^\star(\lambda),\beta^\star(\lambda))) \ .$$

This yields (6) and achieves the proof of Theorem 2.