Asymptotic Consensus Without Self-Confidence

Thomas Nowak*

Abstract

This paper studies asymptotic consensus in systems in which agents do not necessarily have self-confidence, i.e., may disregard their own value during execution of the update rule. We show that the prevalent hypothesis of self-confidence in many convergence results can be replaced by the existence of aperiodic cores. These are stable aperiodic subgraphs, which allow to virtually store information about an agent's value distributedly in the network. Our results are applicable to systems with message delays and memory loss.

1 Introduction

Asymptotic consensus is a phenomenon observed in certain biological, physical, and sociological systems. It is also utilized in some engineered man-made computer systems. The phenomenon consists in agents communicating in a very simple fashion to asymptotically reach agreement on a common real value. In nature, it can be observed (e.g.,[14, 9, 16, 7]) in bird flocking, firefly synchronization, synchronization of coupled oscillators, or opinion spreading. In engineering, it is used for sensor fusion, dynamic load balancing protocols, robot formation protocols, replication techniques, or rendezvous in space.

There is a very simple algorithm for asymptotic consensus that works in a large class of environments: In every computation step of a process, it updates its value to some average of all values it has received, and then sends out its new value. This simple algorithm has two remarkable properties: Firstly, it is very simple and yet manages to solve asymptotic consensus in a surprisingly large number of different environments. Secondly, it is an algorithm that can be *observed* in nature. More specifically, it serves as a widely accepted model in biology, physics, and sociology to explain various phenomena such as bird flocking, synchronization of coupled oscillators, and opinion spreading. It thus stands to reason to expect the algorithm to have a certain robustness against adverse environments. Of course, one can think of using it to attain approximate agreement in man-made, engineered, systems. And indeed, it is actually used, for example, in sensor fusion. For engineered systems, the viewpoint is not one of observing and explaining a given system, but of *analyzing* it for prediction of its future behavior or for assessing the need to improve the system. The speed of convergence in the context of asymptotic consensus is a measure for the stabilization time, or the transient phase, of the system. Obviously, the sharper the analysis of the system and its performance, the tighter it can be integrated into the timing constraints of a larger system, and hence the larger the potential performance of the larger system.

The analysis becomes significantly harder if the communication graphs, or the weights, change over time, if communication delays are introduced and if nodes are susceptible to certain faults. If one admits the dynamicity of the communication graph, then one has already accounted for a large class of faults, namely link faults. The addition of communication delays covers timing faults on links. A class of faults that has received considerably less attention in the literature is that of memory faults, either by memory loss or memory delays, i.e., the value read from local memory is not that of the most recent write operation. Memory delays become more probable with the advent of modern pipelined architectures and memories with weakened consistency properties. The present paper has as its goal the study of systems in which processes cannot, or do not, access their most recent value, but may read an older one or disregard it altogether. In the context of natural asymptotic consensus systems like in sociology, this phenomenon is more naturally called a *lack of self-confidence* and has its specific interest in the analysis of such systems. The paper extends a variety of convergence results known for cases with self-confidence to cases

*Thomas Nowak is with the Département d'Informatique, École normale supérieure, Paris, France. Email: thomas.nowak@ens.fr

without and identifies the importance of having a certain replacement for self-confidence, which we call aperiodic cores. Self-confidence is a specific instance of this notion. Moreover, we discuss an explicit example showing the boundary between convergence and non-convergence in the context of aperiodic cores, shedding a more precise light on the frontier.

In linear algebraic terms, the study of asymptotic consensus is the study of infinite backwards products of stochastic matrices. The first convergence result for products of stochastic matrices is the Perron-Frobenius theorem, which states that the powers of an ergodic stochastic matrix converge to a rank 1 stochastic matrix. It was first generalized to a non-constant product of matrices by Wolfowitz [20] who showed that if every *finite* product of matrices of a set $\mathcal{M}$ of matrices is ergodic, then every backwards of matrices in $\mathcal{M}$ converges to a rank 1 stochastic matrix. The strict finiteness and ergodicity conditions in Wolfowitz' theorem were found to be inappropriate for many applications. Subsequently, Wolfowitz' theorem was extended in several directions (see, for example, [1], [13, Section II.G], or [2]). However, no direct generalization of Wolfowitz' theorem or the Perron-Frobenius theorem was obtained. This is due to the fact that these results all assume a strictly positive diagonal in all occurring matrices. In this sense, the results on asymptotic consensus in dynamic settings are no strict generalizations of the Perron-Frobenius theorem or Wolfowitz' theorem, precisely because of the fact that they require a strictly positive diagonal. One goal of this paper is to remedy this deficiency; by providing convergence results for asymptotic consensus in dynamic settings without this hypothesis. Thus, our results are both strict generalizations of the Perron-Frobenius theorem and existing convergence theorems in asymptotic consensus.

The rest of the paper is organized as follows: Section 2 introduces the model, discusses related work, and gives necessary preliminary results. The notion of aperiodic cores is defined in Section 3 and the first new convergence result based on this notion follows in Section 4. We generalize the definition of aperiodic cores in Section 5 by introducing the notion of clusterings. This is useful to talk about hierarchic systems with local leader agents, as they naturally appear in the reduction from non-synchronous to synchronous settings. We apply this notion in Sections 6, 7, and 8 to show quite general convergence theorems in various environments, together with upper bounds on the convergence rate where applicable. Each of our theorems is followed by a corollary in form of an already known result in the literature. We do this to facilitate finding the context in terms of classical results in which the present paper generalizes the state of the art. Section 9 concludes the paper with some final remarks.

2 Asymptotic Consensus

In particular in computer science for multi-agent systems whose agents start with a private value and repeatedly form averages of perceived values of others. These types of multi-agent systems are not only used in computer networks, but have also been found to model various physical and biological phenomena like the behavior of bird flocks [1, 13]. Mathematically, they translate into long and infinite backwards products of stochastic matrices.

2.1 Computational Model

The distributed computing model in which we study asymptotic consensus is the following: There are n distinguishable agents, each agent $i \in [n] = \{1, 2, \dots, n\}$ possessing a real state variable x_i and communicating by exchanging messages. There is a global discrete time base, referred to by nonnegative integers in $\mathbb{N} = \{0, 1, 2, \dots\}$. At every time $t \in \mathbb{N}$, we denote the content of the agents' state variables by $x_i(t)$. The initial value of state variable x_i is $x_i(0)$. At every time $t \in \mathbb{N}$, every agent sends the content of its state variable to all other agents. Messages may be delayed and/or lost. All agents simultaneously update their state variable at all positive times $t = 1, 2, 3, \dots$ to some weighted average value of the received values, at most one of each other agent, and its current content of its own state variable.

Since the new content of the state variable is a mean value, there exists a $\Delta_{i,j}(t) > 0$ for every $j \in [n]$ such that

$$x_i(t) = \sum_{j=1}^{n} A_{i,j}(t) \cdot x_j\big(t - \Delta_{i,j}(t)\big) \tag{1}$$

with

$$\sum_{j=1}^{n} A_{i,j}(t) = 1 \ . \tag{2}$$

A *configuration* of asymptotic consensus is a collection of real values, one for each agent's state variable, i.e., a vector in $\mathbb{R}^n$. An *execution* of asymptotic consensus is an infinite sequence of configurations $x(t) \in \mathbb{R}^n$ following the evolution (1) for some choice of the $A_{i,j}(t)$ and the $\Delta_{i,j}(t)$. An execution *reaches asymptotic consensus* if $x(t)$ converges and all component-wise limits $\lim_{t\to\infty} x_i(t)$ are equal.

We call an *averaging matrix* a matrix whose entries are all nonnegative and whose row sums are all 1. In other words, it is a row stochastic matrix. Equation (2) assures that the collection of the $A_{i,j}(t)$ is an averaging matrix for all t. A *delay matrix* for time t is a matrix of integers between 1 and t. For every t, the collection of the $\Delta_{i,j}(t)$ is a delay matrix for t. Hence an execution is determined by the initial configuration $x(0)$, the sequence of the averaging matrices $A(t)$, and the sequence of the delay matrices $\Delta(t)$. A pair consisting of a sequence of averaging matrices $A(t)$ and a sequence of vectors $\Delta(t)$ such that every $\Delta(t)$ is a delay matrix for t is referred to as a *setting.* An *environment* is a nonempty set of settings. We say that a setting or an environment reaches asymptotic consensus if all of its executions do.

An important parameter of a setting is the maximum entry of the delay matrices, if it exists. We call a setting B*-bounded* if all entries of its delay matrices are at most B. A 1-bounded setting is called *synchronous* and is determined uniquely by the sequence of averaging matrices. If the nonzero entries of the averaging matrices are lower bounded by some positive α, then we say that the setting has *minimal confidence* α. It has *self-confidence* if all diagonal entries are positive. The *communication digraph* of a stochastic matrix A in $\mathbb{R}^{n\times n}$ has node set $[n]$ and contains an edge (i, j) if and only if $A_{i,j} > 0$.

We note that not every non-synchronous setting reaches asymptotic consensus; *not even* with self-confidence and strongly connected bidirectional communication graphs. The following example shows this. The problem arises if the delay $\Delta_{i,i}(t)$ is strictly greater than 1, i.e., node i does not use its most recent value for the update rule. It is one of the goals of the present paper to study sufficient conditions that enable convergence even if $\Delta_{i,i}(t) > 1$ for some, or even all, i and t.

Example 1. *With $n = 2$ agents, we choose the averaging matrices*

$$A(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad A(t) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \text{ for } t \geq 2$$

and the initial vector $x(0) = {}^t(0, 1)$. Thus there is self-confidence and a minimal confidence of $1/2$. For the delay matrices, we choose

$$\Delta(1) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \Delta(t) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \text{ for } t \geq 2 \ ,$$

i.e., for times $t \geq 2$, there is a delay to itself at every agent of 2 (even though the delay to the other agent is 1). The communication graph for $t \geq 2$ is shown in Fig. 1(a). One can show that $x_1(2t) \to 1/3$ as $t \to \infty$ whereas $x_1(2t + 1) \to 2/3$. Similarly, $x_2(2t) \to 2/3$ and $x_2(2t + 1) \to 1/3$. That is, the system is asymptotically periodic with period 2. The issue becomes clearer when looking at the equivalent synchronous system as studied by Cao, Morse, and Anderson [3]. Its communication graph for $t \geq 2$ is depicted in Fig 1(b). This equivalent synchronous communication graph has a period of 2. We introduce their reduction in more detail at the end of Section 2.2.

In a synchronous setting, the evolution of configurations $x(t)$ is governed by the linear recursive law

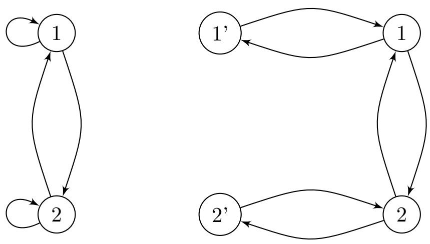
$$x(t) = A(t) \cdot x(t - 1)$$

where $A(t)$ is a row stochastic matrix. Defining the product matrices

$$P(t) = A(t) \cdot A(t - 1) \cdots A(1) \ ,$$

we have

$$x(t) = P(t) \cdot x(0) \ .$$



(a) Non-synchronous

(b) Equiv. synchronous

Figure 1: Communication graphs for $t \geq 2$ in the original non-synchronous and the equivalent synchronous setting in Example 1

In particular, the sequence of state vectors is determined by the initial vector $x(0)$ and the sequence of row stochastic matrices $A(t)$.

In the following sections, we will also use the notation

$$P(t,s) = A(t) \cdot A(t-1) \cdots A(s+1)$$

for partial products. It is $P(t) = P(t,0)$ for all t and $P(t,s) = I$, the identity matrix, if $t \leq s$. If all $A(t)$ are equal to a constant matrix A, then $P(t) = A^t$.

2.2 Related Work

In this subsection, we list several convergence theorems in the literature that our results generalize. All of them suppose self-confidence.

Tsitsiklis introduced the *bounded intercommunication* assumption. It states that if an edge (i,j) appears in infinitely many communication digraphs, then is appears in one of the digraphs

$$G\left(A(t)\right), G\left(A(t+1)\right), \dots, G\left(A(t+B-1)\right)$$

for a fixed B and all t.

Theorem 2 (Tsitsiklis [17])**.** *A synchronous setting with averaging matrices $A(1), A(2), \dots$ with self-confidence and minimal confidence α reaches asymptotic consensus if the digraph G_∞ formed by the edges appearing in infinitely many communication digraphs is strongly connected and the bounded intercommunication assumption holds.*

Moreau and Hendrickx and Blondel independently showed that the bounded intercommunication assumption can be replaced by the assumption that every communication digraph is bi-directional:

Theorem 3 (Moreau [12], Hendrickx and Blondel [8])**.** *A synchronous setting with averaging matrices $A(1), A(2), \dots$ with self-confidence and minimal confidence α reaches asymptotic consensus if the digraph G_∞ is strongly connected and every communication digraph is bi-directional.*

Blondel et al. generalized this result to B-bounded settings:

Theorem 4 (Blondel et al. [1])**.** *A B-bounded setting with averaging matrices $A(1), A(2), \dots$ with self-confidence and minimal confidence α reaches asymptotic consensus if the digraph G_∞ is strongly connected and every communication digraph is bi-directional.*

Touri and Nedić generalized the assumption of bi-directional digraphs to digraphs that are completely reducible. Charron-Bost recently showed its extension to B-bounded settings.

Theorem 5 (Touri and Nedić [18], Charron-Bost [4])**.** *A B-bounded setting with averaging matrices $A(t)$ with self-confidence and minimal confidence α reaches asymptotic consensus if the digraph G_∞ is strongly connected and every communication digraph is completely reducible.*

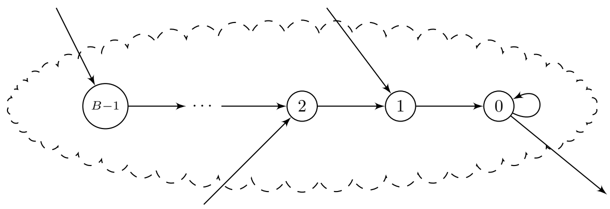


Figure 2: The B copies of an agent in Cao, Morse, and Anderson's reduction

If an execution $x(t)$ reaches asymptotic consensus, one can ask the question of the speed at which this convergence occurs. Olshevsky and Tsitsiklis noted that this speed tends to be exponential and have hence defined the *rate of convergence* as

$$\lim_{t\to\infty} \|x(t) - x^*\|_2^{1/t} \ .$$

Cao, Morse, and Anderson studied *coordinated* communication digraphs, i.e., digraphs that have a node j such that every other node has a path to j. They obtained the following result:

Theorem 6 (Cao, Morse, and Anderson [2, 3])**.** *A B-bounded setting with sequence of averaging matrices $A(1), A(2), \ldots$ with self-confidence and minimal confidence α reaches asymptotic if every communication digraph is coordinated. Moreover, the rate of convergence is less than* 1.

To prove their result, they described a reduction of B-bounded settings to synchronous settings, albeit with B times as many agents as the original setting [3, Section 4.1]. The idea is to replicate every agent B times, but to shift the copies in time, i.e., at time t there is one copy holding the value $x_i(t)$, one $x_i(t-1)$, and so on until $x_i(t-B+1)$. This results in synchronous setting for asymptotic consensus. The replication of agents is illustrated in Fig. 2. Only the copy for the current value $x_i(t)$ has links to other agents' copies. Nonetheless, no such restriction exists for incoming edges. In the new resulting communication digraphs, even if all agents have self-loops in the original communication digraphs, not all nodes have them.

2.3 Dobrushin Semi-Norm for Stochastic Matrices

All stochastic matrices have 1 as an eigenvalue of maximum modulus. If the matrix is irreducible, the corresponding right-eigenspace is one-dimensional and generated by the column vector $\mathbf{1} = {}^t(1, 1, \ldots, 1)$. When studying such matrices, we are hence led to consider the distance of vectors to this eigenspace. Indeed, we will see that considering this distance is an appropriate tool for products of stochastic matrices.

The Dobrushin vector semi-norm on $\mathbb{R}^n$ is defined by setting

$$\delta(x) = \inf_{y \in \mathbb{R}\cdot\mathbf{1}} \|x - y\|_\infty \ .$$

This vector semi-norm induces the Dobrushin matrix semi-norm on $\mathbb{R}^{n\times n}$ by defining it in the operator norm fashion:

$$\delta(A) = \sup_{\substack{x\in\mathbb{R}^n \\ \delta(x)\neq 0}} \frac{\delta(Ax)}{\delta(x)}$$

Clearly, $\delta(A) = 0$ if the image of A is contained in the subspace $\mathbb{R}\cdot\mathbf{1}$.

We now give an example of a matrix whose semi-norm is strictly less than 1, but that has neither a strictly positive column nor a strictly positive diagonal. The matrix is equal to

$$A = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

and its digraph is depicted in Fig. 3. In fact, $\delta(A)$ is equal to 1/2.

The following lemma characterizes the matrices with a Dobrushin semi-norm strictly smaller than 1. It uses the notion of a *scrambling* matrix. A stochastic matrix A is scrambling if for all indices i_1, i_2

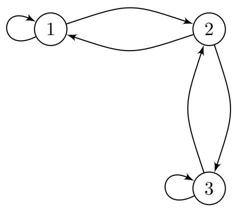


Figure 3: Digraph $G\,(A)$ of matrix A

there exists an index j such that both $A_{i_1,j} > 0$ and $A_{i_2,j} > 0$. Note that, a fortiori, A is scrambling if it has a strictly positive column. Its proof follows from the formula

$$\delta(A) = \max_{i_1,i_2 \in [n]} \sum_{j=1}^{n} (A_{i_1,j} - A_{i_2,j})_+$$

for the Dobrushin matrix semi-norm where $(x)_+ = \max\{x, 0\}$ denotes the positive part of x.

Lemma 7 ([2, 4])**.** *Let A be a stochastic matrix. We always have $\delta(A) \leq 1$ and $\delta(A) < 1$ if and only if A is scrambling. In this case, $\delta(A) \leq 1 - \alpha$ where α is the smallest nonzero entry of A.*

The next lemma shows the utility of δ to show convergence and asymptotic agreement.

Lemma 8. *The sequence of partial backwards products $P(t)$ converges to a rank 1 stochastic matrix if and only if $\delta\big(P(t)\big) \to 0$ as $t \to \infty$.*

Proof. If $P(t)$ converges to a rank 1 stochastic matrix P, then $\delta(P) = 0$. By continuity of δ and monotonicity of $\delta\big(P(t)\big)$, necessarily $\delta\big(P(t)\big) \to 0$.

To prove the converse implication, we now assume that $\delta\big(P(t)\big) \to 0$. We show that, for every $x \in \mathbb{R}^n$, the sequence of vectors $P(t) \cdot x$ converges by showing that it is Cauchy. This then concludes the proof because stochasticity is preserved when taking limits and δ is continuous.

Let $\varepsilon > 0$. Because also $\delta\big(P(t) \cdot x\big) \to 0$, there exists some T such that $\delta\big(P(T) \cdot x\big) \leqslant \varepsilon/2$. Letting $y \in \mathbb{R} \cdot \mathbf{1}$ such that $\delta\big(P(T) \cdot x\big) = \|P(T) \cdot x - y\|_\infty$, we calculate for every $t \geqslant T$:

$$\begin{aligned}
\|P(t) \cdot x - P(T) \cdot x\|_\infty &\leqslant \|P(t,T) \cdot P(T) \cdot x - y\|_\infty \\
&\quad + \|P(T) \cdot x - y\|_\infty \\
&= \|P(t,T) \cdot (P(T) \cdot x - y)\|_\infty \\
&\quad + \|P(T) \cdot x - y\|_\infty \\
&\leqslant 2 \cdot \|P(T) \cdot x - y\|_\infty \\
&= 2 \cdot \delta\big(P(T) \cdot x\big) \leqslant \varepsilon
\end{aligned}$$

because $y = P(t,T) \cdot y$ since $P(t,T)$ is stochastic and y is a multiple of $\mathbf{1}$. This shows that $P(t) \cdot x$ is indeed a Cauchy sequence. $\square$

We now provide a tool to prove convergence of the matrix semi-norm of a product to zero by stating a sufficient condition for the semi-norm of a factor to be constantly bounded away from 1. It shows in particular that the semi-norm of a stochastic matrix is at most 1.

2.4 Graph Interpretation of Matrix Products

Let i and j be nodes of a digraph G. A *walk* in G from i to j is a finite sequence of adjacent nodes in G that starts at i and ends at j. Its length is the number of nodes in the sequence minus one.

The following lemma characterizes positivity of entries in products of stochastic matrices solely in terms of the matrices' associated digraphs. It should be noted that, because we study backward products, the walks grow at the start node and not at the end node.

Lemma 9. *Let $0 \le s \le t$ and $i, j \in [n]$. Then $P_{i,j}(t,s)$ is positive if and only if there exist $i_t, i_{t-1}, \ldots, i_s \in [n]$ with $i_t = i$ and $i_s = j$ such that $(i_\tau, i_{\tau-1})$ is an edge of $G\big(A(\tau)\big)$ for all $s+1 \le \tau \le t$.*

If a strongly connected digraph is aperiodic, there exist walks of arbitrary length between all pairs of nodes as long as the length is greater or equal to a number called the *exponent* (sometimes also *index*) of the digraph. Formally, we denote the smallest T such that there is a walk from i to j of length t for all nodes i and j such that j is reachable from i in G and all $t \ge T$ by $T(G)$. Wielandt provided an upper bound on the exponent, although many more followed [5, 15, 10, 6, 11]. Wielandt's bound is the best possible upper bound in terms of only the number of nodes. If other parameters of the graph are known, however, tighter bounds exist. Since the exponent $T(G)$ appears in some of our bounds, it may be worthwhile to find a more precise bound for the specific graph appearing in a given application framework.

Theorem 10 (Wielandt [19])**.** *Let G be a strongly connected aperiodic digraph with n nodes Then the exponent of G is bounded by*

$$T(G) \le W(n) = \begin{cases} n^2 - 2n + 2 & \text{if } n \ge 2 \\ 0 & \text{if } n = 1 \end{cases} .$$

3 Aperiodic Cores

Classically, in asymptotic consensus, self-confidence of the agents is assumed. That is, every communication digraph contains self-loops at all nodes. This can model the fact that an agent does not ignore or forget its own previous value. We generalize the existence of self-loops, however: A missing self-loop in a specific communication digraph can model memory loss of an agent. We replace the assumption of self-loops to *aperiodic cores*, which are sub-digraphs of all of the settings' communication digraphs. They can be seen as a "distributed safety net against memory loss". In this sense, existence of self-loops is the assumption of a non-distributed safety measure against memory loss or temporary self-distrust. Their function in the proofs is similar to that of self-loops, but they are more general. A parameter that we use over and over in our results is that of the exponent of the aperiodic core. If one assumes self-loops, then H only consists of self-loops at all nodes and this parameter is equal to 0. So, in our theorem statements, if one assumes self-confidence, then $T(H) = 0$.

We call a node j in a digraph G a *leader* of another node i if G contains a path from i to j. A digraph is *j-coordinated* if j is a leader of every node. In this case, node j is called a *leader* of G. A digraph is *coordinated* if it is j-coordinated for some j. If j is a node of a digraph G, we say that G is *j-aperiodic* if j's strongly connected component in G is primitive. A digraph H is a *core* of a sequence $G_1, G_2, \ldots$ of digraphs if H is a sub-digraph of every G_t.

4 Coordinated Aperiodic Cores

We start with assuming that there is a core that is coordinated and leader-aperiodic. The assumption of a core in particular applies if the communication digraph is constant. We hence get a direct generalization of the constant ergodic case:

Theorem 11. *A synchronous setting with averaging matrices $A(t)$ with spanning core H and minimal confidence α reaches asymptotic consensus if there exists some agent j_0 such that H is j_0-coordinated and j_0-aperiodic. Moreover, the rate of convergence is at most $1 - \alpha^{T(H)}/T(H)$.*

We remark that Theorem 11 in particular shows that the setting of Example 1 reaches asymptotic consensus if we change the delay $\Delta_{2,1}(t) = 2$, i.e., *increase* the message delay from agent 1 to agent 2, for $t \ge 2$. Indeed, the resulting equivalent synchronous setting has an aperiodic core from time $t = 2$ on, as is shown in Fig. 4. Note that, as the resulting stochastic matrix for the synchronous system is ergodic and constant, that also the Perron-Frobenius theorem shows convergence to asymptotic consensus in this case. However, embedding this structure into a slightly larger but simple system of 3 agents, as in Fig. 5 (the aperiodic core is almost the whole graph and is shown in bold; only a single edge changes continuously over time) shows the need the generalization that Theorem 11 provides.

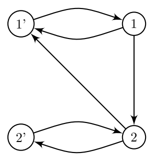
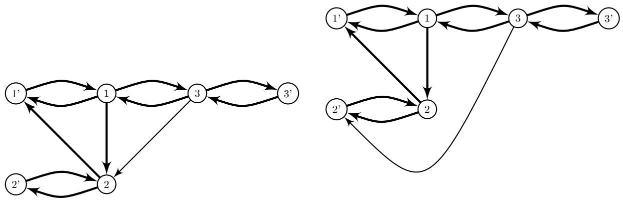


Figure 4: Variant of Example 1 that converges



(a) $G(2t)$ (b) $G(2t+1)$

Figure 5: Equivalent synchronous communication graphs that alternate in time

We prove this theorem in the rest of the subsection.

In general, given a sequence of stochastic matrices $A(1), A(2), \ldots$ in $\mathbb{R}^{n\times n}$ and a node $j \in [n]$, we define $S_j(t,s)$ to be the set of indices $i \in [n]$ such that $P_{i,j}(t,s)$ is positive. Denote by $\mu_j(t,s)$ the smallest (positive) $P_{i,j}(t,s)$ with $i \in S_j(t,s)$. We also define $S_j(t) = S_j(t,0)$ and $\mu_j(t) = \mu_j(t,0)$.

It is easy to see that $\mu_j(t,s) \geq \alpha^{t-s}$ if α is the minimal confidence. This will be our main tool to bound the convergence rate: If $S_j(t,s) = [n]$, then $\delta\left(P(t,s)\right) \leq 1 - \alpha^{t-s}$ by Lemma 7. And if we can show $S_j(t,s) = [n]$ whenever $t - s \geq T$ where T is some constant, then

$$\begin{aligned}\lim_{t\to\infty} \delta\left(P(t)\right)^{1/t} &= \lim_{k\to\infty} \delta\left(P(kT)\right)^{1/kT} \\ &\leq \left(1-\alpha^T\right)^{1/T} \leq 1 - \alpha^T/T \ .\end{aligned}$$

Because all hypotheses we consider are time-invariant, it is sufficient to show $S_j(T) = [n]$.

For Theorem 11, we choose $T = T(H)$: We show that $S_j(T(H)) = [n]$. This is done by reducing the problem to one with a constant matrix. So let A be any stochastic matrix whose digraph $G(A)$ is equal to H. If A^t has a positive column, then so does $P(t)$ because H is a sub-digraph of every communication digraph. This shows the claim since $T(G(A)) = T(H)$.

5 Clusterings

We pair the idea of the distributed safety net in form of an aperiodic core with the notion of *clusters*, which have a leader that is the sole agent of the cluster to regard values of agents other than the cluster's. We will prove that it is not necessary for every agent to be contained in an aperiodic component, but only for the cluster leaders.

A natural example of these clusterings occurs in the reduction of B-bounded settings with self-confidence to synchronous ones (see Fig. 2), for which $T(H) = B - 1$. If we do not assume self-confidence in B-bounded settings, then asymptotic consensus is not necessarily reached, even if the averaging matrices are constant and ergodic. By proving results on cluster-aperiodic cores in synchronous settings, we are hence also proving results on B-bounded settings with self-confidence.

A digraph is a *cluster* with leader l if it is l-coordinated. A *clustering* $\mathcal{C}$ is a collection of node-disjoint clusters $C_1, C_2, \ldots, C_m$ together with respective leaders $l_1, l_2, \ldots, l_m$. A digraph is $\mathcal{C}$*-aperiodic* if every cluster C_j is a sub-digraph, every node is contained in some cluster, and it is l-aperiodic for every leader l_j of $\mathcal{C}$. Fig. 6 shows an example of a $\mathcal{C}$-aperiodic digraph.

A digraph *respects* a clustering $\mathcal{C}$ if the only edges leaving a cluster are the leader's. Given a digraph that respects clustering $\mathcal{C}$, the corresponding *cluster digraph* is the digraph when collapsing all clusters of $\mathcal{C}$ to single node.

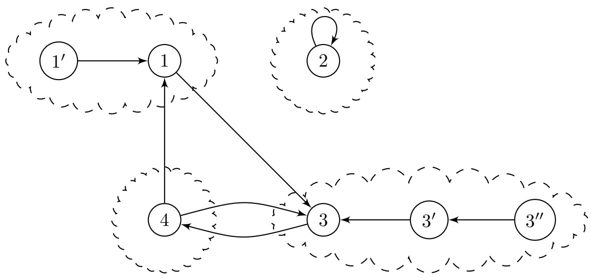


Figure 6: $\mathcal{C}$-aperiodic digraph with leaders $1, 2, 3, 4$

6 Dynamic Coordinated Communication Digraphs

We now prove that asymptotic consensus is also reached if there is no coordinated core, but that coordination at every time step suffices.

Theorem 12. *A synchronous setting with averaging matrices $A(1), A(2), \ldots$ with a $\mathcal{C}$-aperiodic spanning core H and minimal confidence α reaches asymptotic consensus if every communication digraph respects $\mathcal{C}$ and is coordinated. Moreover, the rate of convergence is at most*

$$1 - \alpha^{(n-1)^2(T(H)+1)}/(n-1)^2(T(H)+1)$$

where n is the number of clusters in $\mathcal{C}$.

Corollary 13. *A B-bounded setting with averaging matrices $A(1), A(2), \ldots$ with self-confidence and minimal confidence α reaches asymptotic consensus if every communication digraph is coordinated. Moreover, the rate of convergence is at most $1 - \alpha^{(n-1)^2 B}/(n-1)^2 B$.*

Corollary 13, without the explicit bound on the rate of convergence is included in Theorem 6.

We prove the theorem in the rest of the subsection. Recall that n is the number of nodes, not the number of clusters. We also note that $T(C_j) \leq T(H)$ for every cluster C_j in $\mathcal{C}$.

The sets $S_j(t)$ satisfy a weak form of monotonicity if the sequence of communication graphs have an aperiodic core. If there are self-loops in all communication digraphs, then clearly $S_j(t) \subseteq S_j(t+1)$, which is a special case of the following lemma.

Lemma 14. *If H is a spanning $\mathcal{C}$-aperiodic core and all communication digraphs respect $\mathcal{C}$, then $S_j(t_1) \subseteq S_j(t_2)$ whenever $t_2 - t_1 \geq T(H)$ and j is a leader of $\mathcal{C}$.*

Proof. Let $i \in S_j(t_1)$. Since all communication digraphs respect the clustering, i's leader l_i appears in some earlier set: $l_i \in S_j(t_1')$ with $t_1' \leq t_1$.

Because H is l_i-aperiodic and $t_2 - t_1' \geq T(H)$, there exists a walk of length $t_2 - t_1'$ from i to l_i in H by the definition of $T(H)$. The fact that H is a sub-digraph of all $G\big(A(\tau)\big)$ shows that $P_{i,l_i}(t_2, t_1')$ is positive by Lemma 9.

Hence

$$P_{i,j}(t_2) = \sum_k P_{i,k}(t_2, t_1') \cdot P_{k,j}(t_1') \geq P_{i,l_i}(t_2, t_1') \cdot P_{l_i,j}(t_1')$$

is positive, which shows $i \in S_j(t_2)$. ☐

The following lemmas are used to lower bound the steps need until $S_j(t) = [n]$.

Lemma 15. *If H is a spanning $\mathcal{C}$-aperiodic core, all communication graphs respect $\mathcal{C}$, j is a leader of $\mathcal{C}$, $t \geq T(H)$, and $G\left(A(t+1)\right)$ is j-coordinated, then either $S_j(t) = [n]$ or $S_j(t+1) \setminus S_j(t) \neq \emptyset$.*

Proof. The hypothesis that $t \geq T(H)$ guarantees that $j \in S_j(t)$ by Lemma 14. Every node has a path to j, and hence to $S_j(t)$, in $G\left(A(t+1)\right)$. Now, if $S_j(t) \neq [n]$, there is some $i \in [n] \setminus S_j(t)$ that has an outgoing neighbor k_0 in $S_j(t)$, i.e., $A_{i,k_0}(t+1) > 0$. The condition $k_0 \in S_j(t)$ means $P_{k_0,j}(t) > 0$ and hence

$$\begin{aligned} P_{i,j}(t+1) &= \sum_k A_{i,k}(t+1) \cdot P_{k,j}(t-1) \\ &\geq A_{i,k_0}(t+1) \cdot P_{k_0,j}(t) > 0 \ , \end{aligned}$$

which shows $i \in S_j(t+1)$. □

Lemma 16. *Let H be a spanning $\mathcal{C}$-aperiodic core, all communication graphs respect $\mathcal{C}$ and j be a leader of $\mathcal{C}$. If l is any leader of some cluster C of $\mathcal{C}$ and $l \in S_j(t)$, then $C \subseteq S_j(t + T(H))$.*

Proof. Because C is l-aperiodic and l-coordinated, we have $C \subseteq S_l(\tau)$ for all $\tau \geq T(C)$. Because $T(H) \geq T(C)$, the lemma follows with an application of Lemma 9. □

Set $t_m = m \cdot (T(H)+1)$. For $m \geq 1$, let j_m be a leader of the digraph $G(A(t_m))$ and also of $\mathcal{C}$. Lemma 14 specialized to $s = t_{m-1}$ and $t = t_m - 1 = t_{m-1} + T(H)$ gives $S_j(t_m - 1) \supseteq S_j(t_{m-1})$ for all leaders j and all $m \geq 1$. Lemma 15 applied to $t = t_m$ and $j = j_m$ gives: $S_{j_m}(t_m) \supsetneq S_{j_m}(t_{m-1})$ if $S_{j_m}(t_m - 1) \neq [n]$.

If $m = (n-1)^2 = (n-2)n + 1$, then some $j_0 \in [n]$ appears at least $n-1$ times in the sequence of leaders $j_1, j_2, \ldots, j_m$. By the above and Lemma 16, it is hence $S_{j_0}(t_m) = [n]$, which shows the theorem.

7 Dynamic Communication Digraphs with Fixed Leader

In this subsection, we assume a *fixed* leader in every communication digraph and are able to show a tighter bound on the rate of convergence. The case of strongly connected communication digraphs is a special case.

Theorem 17. *A synchronous setting with averaging matrices $A(1), A(2), \ldots$ with a $\mathcal{C}$-aperiodic spanning core H and minimal confidence α reaches asymptotic consensus if every communication digraph respects $\mathcal{C}$ and there is an agent j_0 such that every communication digraph is j_0-coordinated. Moreover, the rate of convergence is at most*

$$1 - \alpha^{(n-1)(T(H)+1)}/(n-1)(T(H)+1) \tag{3}$$

where n is the number of clusters in $\mathcal{C}$.

Corollary 18. *A B-bounded setting with averaging matrices $A(1), A(2), \ldots$ with self-confidence and minimal confidence α reaches asymptotic consensus if there is an agent j_0 such that every communication digraph is j_0-coordinated. Moreover, the rate of convergence is at most $1 - \alpha^{(n-1)B}/(n-1)B$.*

Corollary 18, without the explicit bound on the rate of convergence is included in Theorem 6.

We use the notation of the previous subsection. The theorem follows similarly by noticing that, in this case, $j_m = j_0$ for all $m \geq 1$ and hence j_0 appears $n-1$ times in the sequence of leaders $j_1, j_2, \ldots, j_{n-1}$.

8 Completely Reducible Communication Digraphs

We now show that one can replace the assumption of coordination by the assumption of completely reducibility at every time step and eventual weak connectivity.

Theorem 19. *A synchronous setting with averaging matrices $A(1), A(2), \ldots$ with a $\mathcal{C}$-aperiodic spanning core H and minimal confidence α reaches asymptotic consensus if every communication digraph respects $\mathcal{C}$, all cluster communication digraphs are completely reducible, and the digraph G_∞ formed by all edges that appear in infinitely many cluster communication digraphs is weakly connected.*

Corollary 20. *A B-bounded setting with averaging matrices $A(1), A(2), \ldots$ with self-confidence and minimal confidence α reaches asymptotic consensus if every communication digraph is completely reducible and the digraph G_∞ of edges that appear in infinitely many communication digraphs is weakly connected.*

Corollary 20 for synchronous settings is Theorem 5.

We prove this theorem in the rest of this subsection. We do not use the exact same proof strategy as in the previous subsection: We show the existence of a T such that

$$\delta\left(P(T)\right) \le 1 - \alpha^{n(T(H)+1)} \ .$$

This suffices to show the theorem because the conditions in the theorem are time-invariant and repeated application thus shows that $\delta\left(P(t)\right) \to 0$. Even though we cannot bound T with the hypotheses of the theorem, we *can* bound the semi-norm uniformly, which is critical for the proof to work. Lemma 8 then concludes the proof.

We first show that G_∞ is completely reducible. For that, we show the following basic lemma.

Lemma 21. *Every union of completely reducible digraphs is completely reducible.*

Proof. Let $\mathcal{G}$ be a set of completely reducible digraphs and let $H = \bigcup \mathcal{G}$ be their union. Let i and j be two nodes in H and suppose that there exists a path P from i to j in the union digraph H. We will show that there then exists a path from j to i in H. This is trivial if $i = j$ so suppose the contrary, i.e., that P is nonempty.

Let $i_0, i_1, \ldots, i_n$ be P's sequence of nodes. For every $1 \le k \le n$, the edge e_k is in some digraph $G \in \mathcal{G}$. Now, because G is completely reducible, there exists a path P_k in G from e_k to e_{k-1}. But then the composite walk $P_n \cdot P_{n-1} \cdots P_1$ is a walk in H from j to i. ☐

Hence G_∞ is completely reducible because Lemma 21 shows that

$$G_\infty = \lim_{T\to\infty} \bigcup_{t \ge T} G(A(t))$$

is a decreasing limit of a sequence of completely reducible digraphs. Because all digraphs are finite, this sequence is eventually constant. Hence its limit G_∞ is equal to one of the sequence's elements and hence completely reducible.

The next lemma captures the essence of the complete reducibility assumption: If $S_j(t)$ does not change, then $\mu_j(t)$ does not decrease. Together with the weak monotonicity of Lemma 14 and eventual connectivity, we are able to show the theorem.

Lemma 22. *Under the hypotheses of Theorem 19, if j is a leader of $\mathcal{C}$ and $S_j(t) = S_j(t+1)$, then $\mu_j(t+1) \ge \mu_j(t)$.*

Proof. Let $P_{i,j}(t+1)$ be positive, i.e., $i \in S_j(t+1) = S_j(t)$. By definition of $S_j(t)$, we have

$$P_{i,j}(t+1) = \sum_{k \in S_j(t)} A_{i,k}(t+1) \cdot P_{k,j}(t) \ . \tag{4}$$

Because $S_j(t) = S_j(t+1)$, we derive that $A_{i,k}(t+1)$ is zero whenever $i \notin S_j(t)$ and $k \in S_j(t)$. Because every node of a cluster is leader-coordinated, every the nodes of a cluster are either all in $S_j(t)$ or all outside of $S_j(t)$. Hence, because the cluster digraph $A(t+1)$ is completely irreducible, we also have that $A_{i,k}(t+1)$ is zero whenever $i \in S_j(t)$ and $k \notin S_j(t)$.

By assumption, we have $i \in S_j(t)$, and hence by the above and by stochasticity of $A(t+1)$:

$$1 = \sum_k A_{i,k}(t+1) = \sum_{k \in S_j(t)} A_{i,k}(t+1) \tag{5}$$

Because $P_{k,j}(t) \ge \mu_j(t)$ for all $k \in S_j(t)$, combination of Equations (4) and (5) yields $P_{i,j}(t+1) \ge \mu_j(t)$. ☐

Choose any leader j_0 of $\mathcal{C}$. For every $i \in [n]$, let t_i be the least nonnegative integer such that $C_i \subseteq S_{j_0}(t_i)$. All t_i are well-defined as G_∞ is strongly connected. By permuting indices, we can assume without loss of generality that $t_1 \le t_2 \le \cdots \le t_n$. Because $P(0)$ is the identity matrix, we have $S_{j_0}(0) = \{j_0\}$ and hence $t_1 = 0$.

We inductively show

$$\mu_{j_0}(t_m) \ge \alpha^{(m-1)(T(H)+1)} \tag{6}$$

for all $1 \leq m \leq n$. This is true for $m = 1$. To prove the inductive step, we distinguish two cases: (A) $t_m - t_{m-1} < T(H)$ and (B) $t_m - t_{m-1} \geq T(H)$.

In case (A), we have

$$\mu_{j_0}(t_m) \geq \alpha^{t_m - t_{m-1}} \cdot \mu_{j_0}(t_{m-1}) \geq \alpha^{(m-1)(T(H)+1)}$$

by the induction hypothesis.

In case (B), we have $S_{j_0}(t) = S_{j_0}(t_{m-1})$ for all t with $t_{m-1} + T(H) \leq t \leq t_m - 1$ by Lemma 14 and the definition of t_m. Repeated application of Lemma 22 hence yields $\mu_{j_0}(t_m - 1) \geq \mu_{j_0}\big(t_{m-1} + T(H)\big)$. We thus have

$$\begin{aligned}\mu_{j_0}(t_m) &\geq \alpha \cdot \mu_{j_0}(t_m - 1) \geq \alpha \cdot \mu_{j_0}(t_{m-1} + T(H)) \\ &\geq \alpha^{T(H)+1} \cdot \mu_{j_0}(t_{m-1}) \geq \alpha^{(m-1)(T(H)+1)}\end{aligned}$$

by the induction hypothesis.

In particular, we have shown Equation (6) for $m = n$. Now set $T = t_n + T(H)$. By Lemmas 14 and 16, $S_{j_0}(T) = [n]$ for all and $\mu_{j_0}(T) \geq \alpha^{n(T(H)+1)}$. This concludes the proof of the theorem.

9 Conclusion

The paper introduced the novel notion of aperiodic cores and showed that the prevalent hypothesis of self-confidence can be replaced by the hypothesis of the existence of an aperiodic core in a large variety of convergence results for asymptotic consensus in dynamic settings. In particular, we discussed and explored the case of non-synchronous environments, for which we gave an explicit example of a 2-bounded system with 2 agents that could not be handled by existing convergence theorems. We also highlighted the need to be careful in these matters by showing that a small variant of the example does not reach asymptotic consensus (and does not even converge). In a linear algebraic view, our results are strict generalizations of the Perron-Frobenius theorem, which was not the case for most results on asymptotic consensus in the literature, as they require self-confidence.

References

[1] Vincent D. Blondel, Julien M. Hendrickx, Alex Olshevsky, and John N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference (CDC-ECC)*, pp. 2996–3000. IEEE, New York, NY, 2005.

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

[3] Ming Cao, A. Stephen Morse, Brian D. O. Anderson. Reaching consensus in a dynamically changing environment: convergence rates, measurement delays, and asynchronous events. *SIAM Journal on Control and Optimization* 47(2):601–623, 2008.

[4] Bernadette Charron-Bost. Orientation and connectivity based criteria for asymptotic consensus, 2013. `arXiv:1303.2043 [cs.DC]`

[5] A. L. Dulmage and N. S. Mendelsohn. Gaps in the exponent set of primitive matrices. *Illinois Journal of Mathematics* 8(4):642–656, 1964.

[6] D. A. Gregory, S. J. Kirkland, and N. J. Pullman. A bound on the exponent of a primitive matrix using Boolean rank. *Linear Algebra and its Applications* 217:101–116, 1995.

[7] Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation* 9(3), 20022

[8] Julien M. Hendrickx and Vincent D. Blondel. Convergence of linear and non-linear versions of Vicsek's model. CESAME Research Report 2005.57, Université catholique de Louvain, Louvain-la-Neuve, 2005.

[9] Ali Jadbabaie, Jie Lin, and A. Stephen Morse. Coordination of groups of mobile autonomous stability agents using nearest neighborhood rules. *IEEE Transactions on Automatic Control* 48(6):988–1001, 2003.

[10] K. H. Kim. An extension of the Dulmage-Mendelsohn theorem. *Linear Algebra and its Applications* 27:187–197, 1979.

[11] Glenn Merlet, Thomas Nowak, Hans Schneider, and Sergeĭ Sergeev. Generalizations of bounds on the index of convergence to weighted digraphs. *Discrete Applied Mathematics* 178:121–134, 2014. An extended abstract was presented at CDC 2014.

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

[13] Reza Olfati-Saber, J. Alex Fax, and Richard M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE* 95(1):215–233, 2007.

[14] Craig W. Reynolds. Flocks, herds, and schools: a distributed behavioral model. *ACM SIGGRAPH Computer Graphics* 21(4):25–34, 1987.

[15] Štefan Schwarz. On a sharp estimation in the theory of binary relations on a finite set. *Czechoslovac Mathematical Journal* 20(4):703–714, 1970.

[16] H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Flocking in fixed and switching networks. *IEEE Transactions on Automatic Control* 52(5):863–868, 2007.

[17] John N. Tsitsiklis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, Massachusetts Institute of Technology, 1984.

[18] Behrouz Touri and Angelia Nedić. Product of random stochastic matrices. *IEEE Transactions on Automatic Control* 59(2):437–448, 2013.

[19] Helmut Wielandt. Unzerlegbare, nicht negative Matrizen. *Mathematische Zeitschrift* 52(1):642–648, 1950.

[20] J. Wolfowitz. Products of indecomposable, aperiodic, stochastic matrices. *Proceedings of the AMS* 14(5):733–737, 1963.