CONVERGENCE OF LINEAR AND NON-LINEAR VERSIONS OF VICSEK'S MODEL

JULIEN M. HENDRICKX AND VINCENT D. BLONDEL

Abstract. The Vicsek model describes the evolution of a system composed by different agents moving in the plane. Each agent has a constant speed and updates its heading using a local rule depending on the headings of its "neighbors". Although the original model by Vicsek is non-linear, most of the convergence results obtained so far deal with linearized versions. In this paper, we introduce a new linear model in which the relative importance of each neighbor can vary with the distance and we prove the convergence of all the agents headings. For this purpose, we derive a theorem on the convergence of long products of stochastic products that applies to infinite set of matrices. Using this result we also prove convergence properties for the original non-linear Vicsek model.

Moreau [7] obtains similar results to ours but using a proof technique based on convexity and system theory. We present here proofs that are based on elementary linear algebra tools. The results we obtain are somewhat weaker than those of Moreau but have the advantage of being indistinctly applicable to continuous and discontinuous systems.

1. Introduction

In 1995, Vicsek et al. [11] proposed a discrete-time model of n autonomous agents all moving in the plane with the same speed but with different headings. Each agent is updating its heading using a local rule based on the average of its own heading and that of its "neighbors". Two agents are neighbors if they are within distance r of each other. In the original model by Vicsek, the average of the headings is performed by computing the normalized vectorial sum of the speeds. However, most of the convergence results obtained until now apply only to the linearized version of this model, in which the average is done directly on the headings θ_i:

$$\theta_i(t+1) = \frac{1}{1+|N_i(t)|}\left(\theta_i(t) + \sum_{j\in N_i(t)} \theta_j(t)\right), \tag{1}$$

where $N_i(t)$ is the set of neighbors of i at time t. This update can be written in matrix form as $\theta(t+1) = A_t\theta(t)$, with $[A_t]_{ij} = (1+|N_i(t)|)^{-1}$ whenever $i=j$ or i is connected to j, and 0 otherwise.

The convergence of systems similar to (1) has been analyzed in a number of recent contributions, see [8] for a survey. In 2003, Jadbabaie, Lin and Morse gave a sufficient condition for all the headings to converge to the same limit [4]. Their condition is based on Wolfowitz's theorem on convergence of stochastic matrices products and requires the existence of an infinite sequence of contiguous nonempty bounded time intervals across which all the agents are linked together (i.e., the graph obtained by linking two agents if they are at least once neighbors during this time interval is connected). These requirements come from the fact that Wolfowitz's theorem can only deal with a finite number of matrices (see Section 2).

Department of Mathematical Engineering, Université catholique de Louvain, Avenue Georges Lemaitre 4, B-1348 Louvain-la-Neuve, Belgium; hendrickx@inma.ucl.ac.be, blondel@inma.ucl.ac.be.

Convergence can also be proved by using the earlier results obtained by Tsitsiklis and his co-authors in the area of distributed algorithms [1, 9, 10]. Those results, which do not rely on Wolfowitz theorem, allow time-varying coefficients and communication delays (under certain conditions): if at time t an agent i uses the value of an agent j to update his value, it can use an outdated value $\theta_j(\tau_{ij}(t))$ (with $\tau_{ij}(t) \leq t$) instead of the most recent one $\theta_j(t)$. Note that in our paper, all communications and influences are assumed to be instantaneous.

More recently, Li and Wang gave a weaker sufficient condition for the convergence of the system (1). They require the graph obtained by connecting two vertices if the corresponding agents are neighbors for an infinite number of time steps to be connected [5].This result implies that each agent's heading always converge to some limiting value. If the graph is not connected, one can indeed see that after a certain number of time steps, the different connected components become totally independent and can be considered separately.

Although the results of Li and Wang allow one to consider any form of neighborhood relation, it can be objected that the neighbors are always assumed equivalent, and that each agent always considers itself as a neighbor among the others. This last assumption is particularly questionable if one considers the model as an approximation of an analogous continuous time process. We therefore introduce in Section 3 a generalization of the linearized Vicsek model that allows a decay of the relative influence of the agents, and a different treatment for one agent's own heading. As a consequence, if one uses the matrix representation $\theta(t+1) = A_t\theta(t)$, the number of possible A_t is not necessarily finite, and Wolfowitz's theorem cannot be used. However, we prove in Section 3 that the results of [5] can be generalized to this model. Actually, we prove it for a generic version that includes this model, the usual linear one, and several others. It indeed allows an explicit dependance of the neighborhood relation and the relative importance on the neighbors on time and on all the variables of the system and their history.

For this purpose, we use a result about convergence of infinite stochastic matrices products that we derive in Section 2. As compared to Wolfowitz's theorem, the main advantage of our result is that it can be applied to products for which the matrices belong to an infinite set. However, it requires the symmetry of the zero/non-zero structure and the positiveness of all the diagonal elements. The presence of a uniform lower bound on the positive elements of the matrices is also needed, but this condition is also implicitly required by Wolfowitz's theorem since it is satisfied by finite set of matrices. We show that the possibility of considering an infinite number of matrices allows us to apply our results to a wide class of non-linear systems. We can indeed replace an initial condition and a non-linear system by a linear system (defined by an infinite sequence of matrices) that produces exactly the same sequence of states for this initial condition. Proving convergence of the linear system associated to each initial condition is then sufficient to prove the convergence of the non-linear system.

Finally, we consider in Section 4 the initial non-linear version of the Vicsek model. Using the approach described above, we prove convergence under the same assumptions as those for the linear case, provided that all the initial headings belong to $]-\pi/2, \pi/2[$, and we show on an example that convergence is not guaranteed when this last condition is not satisfied.

The results presented here can also be found in a more detailed form in [3]. Equivalent results can be found in the seminal paper by Moreau [7]. Moreau uses set valued Lyapunov functions and convexity arguments to prove the convergence of a generic non-linear multi-agent systems, which he particularizes then to the linear case. We begin here by proving this last result using elementary linear algebra tools (similarly to what has been independently done by Lorenz [6] in the context of opinion dynamics) and show that this linear result can in fact be applied to a wide class of non-linear systems. As compared to Moreau's results on non-linear systems, our approach implies some restrictions, the main one being the existence of a uniform lower bound on the influence that the agents can have on each other. But, our way to prove the convergence of non-linear systems using results on linear systems allows us to consider indistinctly continuous and discontinuous systems, which cannot be done by a direct application of Moreau's results.

2. Convergence of infinite matrices products

In [4], Jadbabaie et al. use Wolfowitz's theorem [12] to prove the convergence of the linear version of the Vicsek model under certain assumptions. This theorem states that if square matrices of the sequence $A_0, A_1, A_2, \dots, A_t, \dots$ are all taken from some finite set of ergodic stochastic matrices of $\Re^{n\times n}$ having the property that each finite length product of its matrices is also ergodic, then

$$\lim_{t\to\infty} A_t A_{t-1} \dots A_0 = \mathbf{1}c^T,$$

for a certain positive $c \in \Re^n$, $||c||_1 = 1$. A stochastic matrix $A \in \Re^{n\times n}$ is ergodic if the digraph on n vertices obtained by connecting i to j if $A_{ij} > 0$ is strongly connected and aperiodic (that is, the greatest common divisor of the lengths of its cycles is 1). The hypothesis of a finite number of different matrices forces Jadbabaie et al. to require all the agents to be periodically connected (see [4]). The results that we present here do not require all the matrices to belong to a finite set, nor to be ergodic. On the other hand, they can only be applied if all the matrices have a positive diagonal and a symmetric zero/non-zero structure. They also require a uniform lower bound on the non-zero elements, which is the case for Wolfowitz's theorem as well since there the number of different matrices has to be finite.

Let us first introduce the following notations. Two matrices A and B are said to be of the *same type* ($A \sim B$) if $a_{ij} \neq 0 \Leftrightarrow b_{ij} \neq 0$. The type can be represented by a directed graph that has an edge between i and j if $a_{ij} \neq 0$. By definition, if $(i,j) \in A$ and $A \sim B$, then $(i,j) \in B$. If for all i,j, $(i,j) \in A \Rightarrow (i,j) \in B$, we say $A \subseteq B$ or $A \subset B$ if the inclusion is strict. A matrix A is *type-symmetric* if $A \sim A^T$, and the graph representing its type is undirected.

Consider now an infinite sequence $\mathbb{A} = (A_0, A_1, A_2, \dots)$ of type-symmetric matrices (where A_s can be equal to A_q for $s \neq q$). To each row/column, we associate a vertex, and we say that two vertices are *infinitely connected* if for all $t > 0$, there exists a $t' > t$ such that $(i,j) \in A_{t'}$. We define the undirected graph $G(\mathbb{A})$ associated to this sequence by having an edge between each pair of infinitely connected vertices. Note that a vertex can be infinitely connected to itself and this graph can thus contain loops. Before proving our main results, we need the two following lemmas.

Lemma 1. *Let $\mathbb{A}$ be an infinite sequence of type-symmetric stochastic matrices with positive diagonal elements such that the associated graph $G(\mathbb{A})$ is connected. Then, there exists a $t^* > 0$ such that the product $A_{t^*}A_{t^*-1}\dots A_1A_0$ does not have any zero elements.*

Proof. We denote $A_tA_{t-1}\dots A_1A_0$ by P_t. Observe that since the diagonal elements are always positive, for all t, $P_t \subseteq P_{t+1}$. We now consider two vertices i,j, and show the existence of a t_{ij} such that $(i,j) \in P_{t_{ij}}$, and thus that for all $t \geq t_{ij}$, $(i,j) \in P_t$.

Since $G(\mathbb{A})$ is connected, there exists a path $i = i_0, i_i, \dots, i_{l-1}, i_l = j$. By hypothesis, $(i,i) \in P_t$ for all $t > 0$. We now show the result recursively along this path. Suppose indeed that there is a t_k such that $(i, i_k) \in P_{t_k}$. Since i_k and i_{k+1} are infinitely connected, there exists a $t_{k+1} > t_k$ for which $(i_k, i_{k+1}) \in A_{t_{k+1}}$. Because all the matrices are non-negative and because $P_{t_{k+1}} = A_{t_{k+1}}P_{t_{k+1}-1}$, this implies that $(i, i_{k+1}) \in P_t$ for all $t > t_{k+1}$. □

Lemma 2. *Let A and P be two stochastic matrices with positive diagonal elements. If A is type-symmetric and $Q = AP \sim P$, then the smallest positive element of Q is larger than or equal to the smallest positive element of P.*

Proof. Let us first assume that $(i,j) \in P$ and $(i,k) \in A$, which implies $(k,i) \in A$ due to the type-symmetry of A. We have

$$q_{kj} = \sum_l a_{kl} p_{lj} \geq a_{ki} p_{ij} > 0,$$

and therefore $(k,j) \in P \sim Q$. We now take $(i,j) \in P$, compute q_{ij} to show that it is larger than or equal to p, smallest positive element of P,

$$q_{ij} = \sum_k a_{ik} p_{kj} = \sum_{k,(i,k)\in A} a_{ik} p_{kj}.$$

Since $(i,j) \in P$, for all $(i,k) \in A$ we have $(k,j) \in P > 0$ and thus $p_{kj} \geq p$. Because A is stochastic, this yields

$$q_{ij} \geq \sum_{k,(i,k)\in A} a_{ik} p = p \sum_k a_{ik} = p.$$

$\square$

Note that this result cannot be applied if A is not type-symmetric, as seen with the following counterexample:

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ 0 & 1 \end{pmatrix}.$$

We can now prove a sufficient condition for the convergence of an infinite product of stochastic matrices.

Proposition 1. *Let $\mathbb{A} = (A_0, A_1, A_2, \dots)$ be an infinite sequence of type-symmetric stochastic matrices with positive diagonal elements such that the associated graph $G(\mathbb{A})$ is connected. If there exists a uniform lower bound $\mu > 0$ on the positive elements of all the matrices of $\mathbb{A}$, then there exists a positive $c \in \Re^n, ||c||_1 = 1$ such that*

$$\lim_{t\to\infty} A_t A_{t-1} \dots A_1 A_0 = \mathbf{1} c^T \tag{2}$$

Proof. As in the proof of Wolfowitz's Theorem [12], we rewrite the infinite product by regrouping some matrices together in such a way to have a uniform upper bound on the coefficient of ergodicity λ (see under) of the matrices obtained. The main difference is that in [12] this upper bound resulted from the finite number of available matrices, while here, it is a consequence of Lemma 2 and of the uniform lower bound on the positive elements.

Let us introduce the following measure of the difference between the rows of a matrix [12],

$$\delta(A) = \max_j \left(\max_{i_1,i_2} |a_{i_1 j} - a_{i_2 j}| \right).$$

One can see that $\delta(A) = 0$ if and only if all the rows of A are identic.

Denote $A_t A_{t-1} \dots A_1 A_0$ by P_t. We begin by showing that the convergence of $\delta(P_t)$ is sufficient to prove this proposition. Suppose that for any $\epsilon' > 0$, there exists an integer $\tau(\epsilon')$ such that for all $t > \tau(\epsilon')$, $\delta(P_t) < \epsilon'$. Consider then an $\epsilon > 0$, take a $t > \tau(\epsilon/2n)$, and a $t' > t$. By definition of δ, P_t could be written as

$$P_t = \mathbf{1} b^T + E,$$

for a certain $b \in \Re^n$, and with $||E||_\infty \leq n\epsilon/2n = \epsilon/2$. So,

$$P_{t'} = (A_{t'} \dots A_{t+1}) \left(\mathbf{1} b^T + E\right) = \mathbf{1} b^T + (A_{t'} \dots A_{t+1}) E,$$

and we have

$$\begin{aligned} ||P_{t'} - P_t||_\infty &= ||A_{t'} \dots A_{t+1} E - E||_\infty \\ &\leq ||A_{t'} \dots A_{t+1}||_\infty ||E||_\infty + ||E||_\infty \ . \\ &\leq \epsilon/2 + \epsilon/2 = \epsilon. \end{aligned}$$

The P_t form thus a Cauchy sequence and therefore converge to a certain limit P. Since δ is a continuous function, we have $\delta(P) = \lim_{t\to\infty} \delta(P_t) = 0$. All the rows of P are thus identical and P can be written as $P = \mathbf{1}c^T$ for a certain $c \in \Re^n$. Because all the matrices involved in the infinite product are stochastic and have a positive diagonal, c is trivially be positive and satisfies $||c||_1 = 1$.

To prove the convergence of $\delta(P_t)$ to 0, we use the *coefficient of ergodicity* λ introduced by Hajnal [2],

$$\lambda(A) = 1 - \min_{i_1,i_2} \sum_j \min\left(a_{i_1 j}, a_{i_2 j}\right). \tag{3}$$

Note that $0 \leq \lambda(A) \leq 1$ if A is stochastic, and $\lambda(A) = 0$ if and only if $\delta(A) = 0$. Moreover, it can be shown [2] that for any k and any sequence of matrices $A_0, \dots, A_k$,

$$\delta\left(A_k A_{k-1} A_0\right) \leq \lambda\left(A_k\right) \lambda\left(A_{k-1}\right) \dots \lambda\left(A_0\right).$$

So it would be sufficient to be able to rewrite P_t as

$$P_t = \left(A_t A_{t-1} \dots A_{t_{k(t)}+1}\right) \left(Q_{k(t)} Q_{k(t)-1} \dots Q_1\right),$$

where $Q_s = A_{t_s} \dots A_{t_{s-1}+1}$ for $0 = t_0 < t_1 < \dots < t_{k(t)}$ ($\lim_{t\to\infty} k(t) = \infty$), and such that for all s, $\lambda(Q_s) \leq d < 1$ for a certain uniform upper bound d. We would indeed then have

$$\lim_{t\to\infty} \delta(P_t) \leq \lim_{t\to\infty} \lambda(Q_{k(t)}) \dots \lambda(Q_1) \leq \lim_{t\to\infty} d^{k(t)} = 0. \tag{4}$$

Note that $k(t)$ has to tend to infinity with t, but we require no upper bound on $t_{k+1} - t_k$. We now show that such a function k always exists.

By Lemma 1, there is a t_1 such that $Q_1 := A_{t_1} \dots A_0$ does not have any zero elements. Applying this result to the product $A_t \dots A_{t_1+1}$ gives us then t_2, and doing this recursively, we obtain $0 = t_0 < t_1 < \dots < t_{k(t)}$. Since this can be done indefinitely, $\lim_{t\to\infty} k(t) = \infty$. By construction, all the elements of any Q_s are positive. We now show that they are also larger than $\mu^{n(n-1)+1}$ and thus that for all s, $\lambda(Q_s)$ defined in (3) is no greater than $d := 1 - n\mu^{n(n-1)+1}$, where μ is the uniform lower bound on the elements of any matrix of $\mathbb{A}$. By (4), this is sufficient to show the convergence of P_t.

Let us consider a particular k and denote $A_t A_{t-1} \dots A_{t_{k-1}+1}$ by R_t. Since the diagonal elements of theses matrices are positive, we know that $R_t \subseteq R_{t+1}$. There are thus at most $n(n-1)$ values of t such that $R_{t+1} \not\sim R_t$. By Lemma 2, these values of t are the only ones for which the smallest positive element r_{t+1} of R_{t+1} can be smaller than r_t, smallest positive element of R_t. Consider such a t. By the non-negativity of the considered matrices, we have

$$r_{t+1} \geq r_t \left(\min_{(i,j)\in A_{t+1}} a_{ij}\right) \geq r_t \mu.$$

So, for all $t \in \{t_{i-1}+1, \dots, t_i\}$, $r_t \geq \mu^{n(n-1)} r_{t_{i-1}+1} \geq \mu^{n(n-1)+1}$ because $R_{t_{k-1}+1} = A_{t_{k-1}+1}$. Since $Q_k = R_{t_k}$, we have

$$\lambda(Q_k) \leq d = 1 - n\mu^{n(n-1)+1}.$$

By (4), this implies the convergence if $\delta\left(P_T\right)$ to 0, which as explained above is sufficient to prove this proposition. □

In the last proposition, the hypothesis of a lower bound on the smallest positive element cannot be removed. Consider for example

$$P_t = \begin{pmatrix} \epsilon^t & 1-\epsilon^t \\ 1-\epsilon^t & \epsilon^t \end{pmatrix} \begin{pmatrix} \epsilon^{t-1} & 1-\epsilon^{t-1} \\ 1-\epsilon^{t-1} & \epsilon^{t-1} \end{pmatrix} \cdots \begin{pmatrix} \epsilon & 1-\epsilon \\ 1-\epsilon & \epsilon \end{pmatrix}. \tag{5}$$

One can show that the diagonal elements are smaller than $\frac{\epsilon}{1-\epsilon}$ is k is odd and larger than 1-$\frac{\epsilon}{1-\epsilon}$ if k is even. So if $\epsilon < 1/3$, P_t cannot converge. But, the hypothesis of uniform lower bound is actually too strong. It is in fact used to have a uniform bound $d < 1$ on the value of the coefficients of ergodicity λ (see (3) and (4)), which is a sufficient but not necessary condition to guarantee the convergence to zero of their product. Consider indeed

$$P_t = \begin{pmatrix} 1-\frac{1}{t+k} & \frac{1}{t+k} \\ \frac{1}{t+k} & 1-\frac{1}{t+k} \end{pmatrix} \begin{pmatrix} 1-\frac{1}{t-1+k} & \frac{1}{t-1+k} \\ \frac{1}{t-1+k} & 1-\frac{1}{t-1+k} \end{pmatrix} \cdots \begin{pmatrix} 1-\frac{1}{k} & \frac{1}{1+k} \\ \frac{1}{k} & 1-\frac{1}{k} \end{pmatrix},$$

for a certain $k > 2$. There is no uniform positive lower bound on the element of the matrices. Moreover, the coefficient of ergodicity $\lambda(A_s)$ of the s^{th} matrix of this product is $1 - 2/(k+s)$, which can be arbitrarily close to 1 for large values of s. However, it is possible to prove that

$$\lim_{t\to\infty}(1-\frac{2}{k+t})(1-\frac{2}{k+t-1})\dots(1-\frac{2}{k}) = 0,$$

which implies the convergence of P_t to a matrix with identical rows when t tends to infinity (by symmetry, this limiting matrix is $\frac{1}{2}\mathbf{1}\mathbf{1}^T$). There might thus exist weaker conditions that could replace the uniform lower bound one.

A result similar to Proposition 1 can be obtained by considering the connected components of the graph $G(\mathbb{A})$.

Theorem 1. *Let $\mathbb{A} = (A_0, A_1, A_2, \dots)$ be an infinite sequence of type-symmetric stochastic matrices with positive diagonal elements and $G(\mathbb{A})$ the associated graph. If there exists a uniform lower bound $\mu > 0$ on the positive elements of all the matrices of $\mathbb{A}$, then there exists A such that*

$$\lim_{t\to\infty} A_t A_{t-1} \dots A_1 A_0 = A$$

Moreover, A can be written as $A = DQ$ where D is block diagonal (after a row and column permutation PDP^T) and Q is a $n\times n$ type-symmetric stochastic matrix with a positive diagonal. The blocks of D correspond to the connected components G_l of $G(\mathbb{A})$, and each of them can be written as $\mathbf{1}c_l^T$, for a certain positive c_l such that $||c_l||_1 = 1$.

Proof. By definition of $G(\mathbb{A})$, there exists $t^* > 0$ such that for all $t > t^*$, $[A_t]_{ij}$ cannot be positive if i and j are not connected in $G(\mathbb{A})$. Let $Q = A_{t^*}A_{t^*-1}\dots A_0$, we can rewrite the original product as

$$\lim_{t\to\infty}(A_t A_{t-1}\dots A_{t^*+1})\, Q.$$

By definition of t^*, all the A_t $(t > t^*)$ are block diagonal (after a row and column permutation), the block partition being the same in each matrix and corresponding to the different connected components of $G(\mathbb{A})$. The graph associated to each of the block sequence is therefore trivially connected, and the desired result follows from the application of Proposition 1 to each infinite block product. ∎

From this theorem follows a natural result on the convergence of certain discrete-time dynamic systems.

Corollary 1. *Let $\theta(0)$ be a vector in $\Re^n$, and consider the recurrence $\theta(t+1) = A_t\theta(t)$, where the A_t are type-symmetric stochastic matrices with a positive diagonal and have a lower bound μ on their positive elements. There exists $\theta \in \Re^n$ such that*

$$\theta = \lim_{t\to\infty} \theta(t).$$

Moreover, if i and j are in the same connected component of the graph associated to the sequence $\mathbb{A} = (A_0, A_1, A_2, \dots)$, then $\theta_i = \theta_j$.

Proof. By Theorem 1, we have

$$\lim_{t\to\infty} \theta(t) = \left(\lim_{t\to\infty} A_{t-1}A_{t-2}\dots A_1A_0\right)\theta(0) = DQ\theta(0),$$

where D and Q have the same properties as in Theorem 1. $\theta(t)$ converges thus to $\theta = DQ\theta(0)$, and the rest of the result follows directly from the structure of D. □

This corollary gives us a sufficient condition for the convergence of two elements of the $\theta(t)$ to a same value. It would be of course more satisfactory to have a necessary and sufficient, but one can see that the convergence sometimes depends on the initial condition. Consider indeed $A_t = I$ for all t. The system converges to $\theta = \theta(0)$. Although the elements of $\theta(t)$ never influence each other, two of them could "converge" to a same value if the corresponding entries of the initial condition $\theta(0)$ were equal. Another possibility would be a necessary and sufficient condition on the matrix sequence for two elements of $\theta(t)$ to converge to the same value for all initial condition $\theta(0)$. But, no such condition could rely only on the graphs associated to the matrices and matrix sequences. Imagine that there is a $q \in \mathbb{N}_0$ for which $A_q = \frac{1}{N}\mathbf{1}\mathbf{1}^T$. Independently of all the other matrices of the sequence (if they are stochastic), we would have $\theta_i(t) = \theta_j(t)$ for all $t > q$ and for any i and j. On the other hand, one can easily build a matrix which contains also a matrix A_q such that $G(A_q) = K_n$ but that does not imply the convergence of all the θ_i to a common value.

As already mentioned, the matrices of Theorem 1 and Corollary 1 do not need to belong to a finite set unlike in the applications of Wolfowitz's theorem [12]. This allows us to apply those results to non-linear and even non-continuous systems. Consider a generic discrete-time dynamic system defined by $\theta(t+1) = f(t, \theta(t), \Theta(t))$ for $\theta(t) \in \Re^n$ and where $\Theta(t) \in \Re^{n\times(t+1)}$ contains all the previous values of $\theta(t)$. Note that no assumption is made about the linearity or the continuity of f. Consider also the sequence of vectors $\hat\theta(0), \hat\theta(1), \hat\theta(2), \dots$ produced by this system for a particular initial condition $\hat\theta(0)$. If one can prove the existence of an infinite sequence of matrices $\mathbb{A} = (A_0, A_1, A_2, \dots)$ satisfying the hypotheses of Corollary 1 and such that for all t, $\hat\theta(t+1) = A_t\hat\theta(t+1)$, then this corollary guarantees the convergence of $\hat\theta(t)$ when t tends to infinity, i.e. that the considered dynamic system converges if the initial condition is $\hat\theta(0)$. So, if for each initial condition $\theta(0)$ in a certain set Ω there exists such a matrix sequence, then the dynamic system is guaranteed to converge for all initial conditions of Ω. Moreover, for each of these matrix sequences, we can define the associated graph $G(\mathbb{A})$ (as in the previous results of this section). If on Ω this graph does not depend on the particular value of the initial conditions, then the results of Corollary 1 about a unique limiting value for each connected component of the graph can also be applied. An example of such application of Corollary 1 to non-linear systems is provided in Section 4 where the convergence of the initial non-linear Vicsek model is proved.

3. Linear Vicsek model

In this section, we propose a new version of the linearized Vicsek model in which the relative importance of each neighbor depends on its distance to the considered agent, and compare it to the usual one. We prove convergence of all the heading for a generic version of this kind of

Vicsek model that contains the two mentioned versions.

Let us first recall the usual linearized version of the model. n autonomous agents are moving in the plane at a constant speed v with a heading $\theta_i(t)$ that depends on the agent and on time. At each time step their position (x_i, y_i) is updated by

$$x_i(t+1) = x_i(t) + v\Delta t \cos\theta_i(t) \qquad y_i(t+1) = y_i(t) + v\Delta t \sin\theta_i(t), \tag{6}$$

and the headings by

$$\theta_i(t+1) = \frac{1}{1+|N_i(t)|}\left(\theta_i(t) + \sum_{j\in N_i(t)} \theta_j(t)\right), \tag{7}$$

where $N_i(t)$ is the set of neighbors of i at the time step t, two agents being neighbors if they are separated by a distance smaller than a certain radius r.

A natural objection to this headings update model is that all the neighbors of an agent are considered as equivalent, independently of their relative positions. Moreover, each agent considers itself as one neighbor among the others, while it would be reasonable to assume that its own heading has more influence than those of the others. This is especially true if one considers this discrete-time system as an approximation of a corresponding continuous system where the derivative of the heading depends on the neighbors headings. In that case, one would indeed expect the importance of the neighbors headings to be inversely proportional to the time discretization. We therefore propose the following generalization of the usual update model, for the same neighborhood relation:

$$\theta_i(t+1) = \frac{g\left(|N_i(t)|\right)\theta_i(t) + \sum_{j\in N_i(t)} f(d_{ij}(t))\theta_j(t)}{g\left(|N_i(t)|\right) + \sum_{j\in N_i(t)} f\left(d_{ij}(t)\right)}, \tag{8}$$

where $d_{ij}(t)$ is the distance between agent i and j at time t and f_i is a decreasing function that admits a uniform positive lower bound in the circle of radius r centered the origin. g is a positive function that allows the agents to give more or less importance to their own heading, depending of the number of their neighbors. One possibility would be to take a constant g and $f(d) = b + ce^{-(d/d_0)^2}$ for some positive b, c and d_0. If r is large, this would mean that the close neighbors would have a relative importance $f \simeq b + c$, while the one on the limit of the agent sightseen would have $f \simeq b$. Another possibility for f would be to take a positive multiple-step function, which would not immediately be allowed by the results of [7] since it is discontinuous.

Let $\theta(t)$ be the column vector containing all the headings. It is possible to express the update models (7) and (8) using the matrix form $\theta(t+1) = A_t\theta(t)$, where A_t is a stochastic matrix. One can see that in the usual version (7), since the number of agents and thus the number of neighborhood graphs are finite, there is a finite number of possible A_t. On the other hand, even if there is still a finite number of neighborhood graphs, the matrices corresponding to the update rule (8) depends also on the distances, and their number can therefore be infinite; Wolfowitz's theorem thus cannot be applied, but we can use Theorem 1. Instead of proving our convergence result for the model (8), we consider the following generic version of linearized Vicsek model, that contains the usual version, ours (8), and several others. The positions are still updated by (6), but the headings update is

$$\theta_i(t+1) = a_{ii}\left(t, S(t)\right)\theta_i(t) + \sum_{j:(i,j)\in E(t,S(t))} a_{ij}\left(t, S(t)\right)\theta_j(t), \tag{9}$$

where $S(t)$ represents the system state (i.e. headings and positions of all the agents) at time t and all its history, and $E(t, S)$ a symmetric neighborhood relation. Note that the definition of the neighborhood relation and the coefficients a_{ij} can depend on the time and on the system state and history, but we impose the existence of a uniform lower bound $\mu > 0$ on the a_{ij} and that (9) represents a convex combination, i.e.

$$a_{ii}\left(t,S\right) + \sum_{j:(i,j)\in E(t,S)} a_{ij}\left(t,S\right) = 1, \qquad \forall t, S, i.$$

One can see that the headings update models (7) and (8) are particular cases of the generic version (9). Now, as in these particular cases, one can re-express (9) by $\theta(t+1) = A_t\theta(t)$, where $[A_t]_{ij} := a_{ij}\left(t, S(t)\right)$ if $i = j$ or $(i,j) \in E(t, S(t))$, and 0 else. The matrices A_t are stochastic and satisfy the conditions of Corollary 1. So, if we define a graph associated to this system by connecting two agents if for all time step t, there exists a time step $t' > t$ at which they are neighbors, or equivalently if they are neighbors at an infinite number of time steps, we have the following result:

Theorem 2. *Considering the headings update system described in (9), the heading of each agent converges to a limiting value. Moreover, if two agents belong to the same connected component of the associated graph obtained by connecting the agents that are are neighbors at an infinite number of time steps, their limiting heading are identical.*

Theorem 2 provides a sufficient condition for all the headings to converge to a same limit. It suffices indeed that the graph associated to the system is connected. As explained in Section 2, there is probably no condition relying only on the matrices type that would be necessary and sufficient for all the headings to converge to a same limit for all initial condition. However, if we take a particular system as described in (7), some better results might be found, using for example the properties of corresponding neighborhood relation.

In Section 2, we showed with (5) that although the hypothesis of a uniform lower bound on the elements of the matrices was not necessary, it could not simply be removed. We now show a practical example of Vicsek model where the absence of this uniform lower bound leads to two headings converging to different values while the agents remain always neighbors. Consider the headings update model (8) with $g \equiv 1$, $f(d) = e^{-(d/d_0)^2}$ and an infinite r. For each agent, we have

$$\theta_i(t+1) = \frac{\theta_i(t) + c\sum_{j\neq i} e^{-(d_{ij}(t)/d_0)^2}\theta_j(t)}{1 + c\sum_{j\neq i} e^{-(d_{ij}(t)/d_0)^2}}.$$

Imagine now a system composed of two agents initially at the origin $(0,0)$ with initial headings $\theta(0) = (\pi/3, -\pi/3)^T$, and suppose that $c = 1/16$ and $d_0 = V\Delta t/\sqrt{2}$. One can show by induction that the following inequalities hold

$$\begin{aligned} d_{12}(t) &\geq td_0, \\ \theta_1(t) = -\theta_2(t) &\geq \tfrac{\pi}{4}\left(1 - \tfrac{1}{8}\tfrac{e - e^{-(t-1)^2}}{e-1}\right) \geq \pi/4. \end{aligned}$$

Since their evolution is monotonous and bounded, both headings converge. But, although the agents always remain connected, their limiting headings are not identical.

4. Non-linear Vicsek model

In this section, we prove similar convergence properties for the non linear model as those of the linear model. The same result is obtained by Moreau [7], but the originality of our approach is that although the system is non-linear, we prove its convergence using results about the convergence of linear systems (Corollary 1), following the method described at the end of Section 2.

Let $v_i(t) = v\,(\cos(\theta_i(t)), \sin(\theta_i(t)))$ be the speed of agent i at time t. The positions are updated by (6) as in the linear version, but the update of the speed vectors is given by

$$w_i(t+1) = \frac{1}{1+|N_i(t)|}\left(v_i(t) + \sum_{j\in N_i(t)} v_j(t)\right), \quad v_i(t) = v\frac{w_i(t+1)}{||w_i(t+1)||}, \tag{10}$$

where $N_i(t)$ is the set of neighbors of i at time t. Since we do not use any other information about the neighborhood relation, we just assume that is symmetric. Note that a special treatment is needed if $||w_i(t+1)|| = 0$. We define the graph associated to the system as in the linear case by connecting the agents that are neighbors at an infinite number of time steps.

Theorem 3. *Considering the speed update model defined by (10), if all the initial headings belong to* $]-\frac{\pi}{2},\frac{\pi}{2}[$*, all the speeds converge to a limiting vector (of norm* v*). Moreover, if two agents belong to the same connected component of the graph associated to the system, their limiting speed are identical.*

Proof. Let us consider an initial condition of this system. The idea of the proof is to build a linear system that gives, for this particular initial condition, exactly the same sequence of headings (as explained at the end of Section 2). Such a system is not unique but we show that it is always possible to obtain one that satisfies the hypotheses of Corollary 1 and for which the associated graph is the same as for the initial non-linear system. Since this can be done for any initial condition as long as all the initial headings belong to $]-\frac{\pi}{2},\frac{\pi}{2}[$, it will imply the desired result.

On $]-\frac{\pi}{2},\frac{\pi}{2}[$ there is a unique continuous correspondence between a heading $\theta_i(t)$ and its tangent. Moreover, once the tangent is known, the sine and cosine of the angle are uniquely determined, and so is the corresponding speed since all the speeds have by hypothesis the same norm v. The headings update (10) model can thus be reexpressed as

$$\tan(\theta_i(t+1)) = \frac{\sin(\theta_i(t)) + \sum_{j\in N_i(t)} \sin(\theta_i(t))}{\cos(\theta_i(t)) + \sum_{j\in N_i(t)} \cos(\theta_i(t))}.$$

Note that the problem of a zero denominator does not appear anymore here because all the cosine are positive. Using trigonometry arguments, it is possible to prove the following result by induction on m:

Consider $\phi_1, \phi_2, \ldots, \phi_m \in]-\frac{\pi}{2},\frac{\pi}{2}[$ *such that* $\cos\phi_j > \epsilon$ *for all* j*, and define* $\phi(m) \in]-\frac{\pi}{2},\frac{\pi}{2}[$ *by*

$$\tan(\phi(m)) = \frac{\sum_{j=1}^m \sin(\phi_j)}{\sum_{j=1}^m \cos(\phi_j)}. \tag{11}$$

There exists $c_1, c_2, \ldots, c_m \geq \mu(m,\epsilon) > 0$ *(where* μ *is a positive function of* m *and* ϵ*) such that* $\sum_{j=1}^m c_j = 1$*, and*

$$\sum_{j=1}^m c_j \tan(\phi_j) = \tan(\phi(m)).$$

Because all the initial headings belong to $]-\frac{\pi}{2},\frac{\pi}{2}[$, one can also show by induction that at no time step there is a heading having a cosine smaller than $\epsilon_\theta = \min_i \theta_i(0)$. The above result can thus be applied to (10); it gives an expression of $\tan(\theta_i(t+1))$ as a convex combination of all

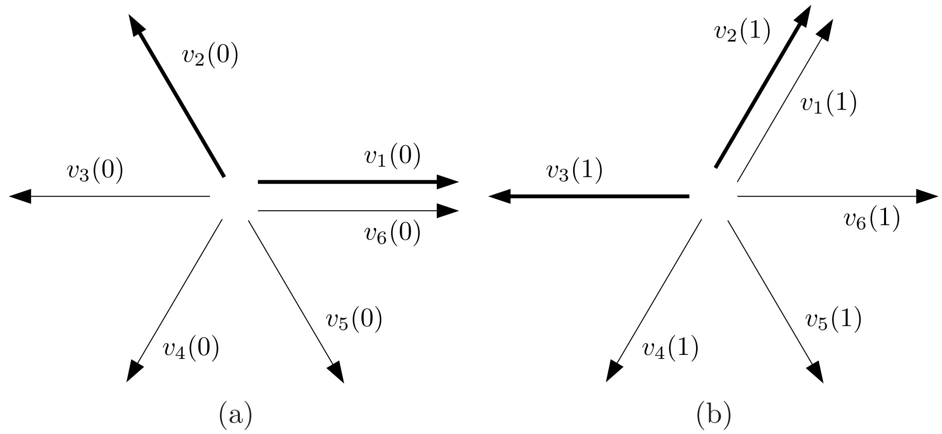

FIGURE 1. Agents speed at $t = 0$ (a) and $t = 1$ (b). The bold arrows correspond to agents that are neighbors. One can see the possible apparition of cycles, and thus the absence of convergence.

the $\tan(\theta_j(t))$ (for all the $j \in N_i(t)$ and for $j = i$) where the coefficients are always larger than a certain $\mu > 0$ that does not depend on time. Rewriting this using matrices, we have

$$\tan(\theta(t+1)) = A_t A_{t-1} \dots A_0 \tan(\theta(0)),$$

where all positive elements are larger than a uniform lower bound $\mu > 0$. By (4), the diagonal elements are all positive. Moreover, the non-diagonal elements $[A_t]_{ij}$ are positive if and only if i and j are neighbors. This implies that graph associated to the matrix sequence is the same as the one associated to the non-linear system and also that the matrices $A_0, A_1, \dots$ are type-symmetric since the neighborhood relation is symmetric. We can then invoke Corollary 1 to guarantee the convergence of the $\tan(\theta(t))$ and therefore of $\theta(t)$, as well as the desired result about a unique limiting heading for each connected component of the graph associated to the system.

□

Our hypothesis about the initial headings could seem artificial, but there exists cases where no convergence takes place if this condition is not satisfied. Consider indeed

$$\theta(0) = \begin{pmatrix} 0 & \frac{2}{3}\pi & \pi & \frac{4}{3}\pi & \frac{5}{3}\pi & 0 \end{pmatrix}^T,$$

and suppose that at time step t, only the agents $1 + (t \mod 6)$ and $1 + (t + 1 \mod 6)$ are neighbors. As shown in Figure 1, one can see that the evolution of the speeds is cyclic with $\theta(t + 6) = \theta(t)$. The hypothesis about the initial condition in Theorem 3 can thus not be removed. However, we were not able to build such a cyclic situation if the usual neighborhood relation is used, i.e., two agents are neighbors if they are separated by a distance no greater than a certain pre-specified radius r.

5. CONCLUDING REMARKS

In this paper, we gave an alternative proof of the convergence of any infinite product of type-symmetric matrices with positive diagonal to a stochastic matrix of order one, provided that there exists a lower bound on the positive elements and that the graph associated to the matrix sequence is connected. We saw that the uniform lower bound condition is too strong and can thus be relaxed. The weakness of our connectivity condition and of the graph definition allowed us then to generalize our results to non-connected graph; two vertices of different connected components becoming indeed totally independent after a finite number of step in the sequence, one can apply our first result to each connected component. As a corollary, we proved the convergence of a wide class of linear discrete-time dynamic systems. Moreover, since we were not restricted to finite set of matrices as in [12], we were also able to use this result in non-linear (possibly discontinuous) cases.

We applied our result to a generalization of linear (symmetric) Viscek models and proved the convergence of all the headings to a limiting value. We also gave a sufficient condition about the connectivity of the graph associated to the system for all the headings to converge to the same limiting value. We then turned our attention to the original non-linear Vicsek model and proved the same convergence properties as in the linearized case, under the assumption that all the initial headings belonged to an interval smaller than π.

Acknowledgments

The work of Julien M. Hendrickx and Vincent D. Blondel is supported by the Concerted Research Action (ARC) "Large Graphs and Networks" of the French Community of Belgium and by the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. This work was also supported in part by DoD AFOSR URI for "Architectures for Secure and Robust Distributed Infrastructures", F49620-01-1-0365 (led by Stanford University). The scientific responsibility rests with its author(s). Julien Hendrickx holds a FNRS fellowship (Belgian Fund for Scientific Research).

References

[1] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proc. of the 44nd IEEE Conf. on Decision and Control*, Seville, Spain, December 2005.
[2] J. Hajnal. Weak ergodicity in non-homogeneous Markov chains. *Proc. Cambridge Philos. Soc.*, 54:233–246, 1958.
[3] J.M. Hendrickx and V.D. Blondel. Convergence of different linear and non-linear vicsek models. CESAME research report RR 2005.57, Université catholique de Louvain.
[4] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Automat. Control*, 48(6):988–1001, 2003.
[5] S. Li and H. Wang. Multi-agent coordination using nearest neighbor rules: revisiting the vicsek model. *CoRR*, cs.MA/0407021, 2004.
[6] J. Lorenz. A stabilization theorem for continuous opinion dynamics. *Physica A*, 355(1):217–223, 2005.
[7] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Trans. Automat. Control*, 50(2):169–182, 2005.
[8] R. Olfati-Saber, J.A. Fax, and R.M. Murray. Consensus and cooperation in networked multi-agent systems. *To appear in IEEE Proceedings*, 2006.
[9] J.N. Tsitsiklis. *Problems in decentralized decision making and computation*. PhD thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, http://web.mit.edu/ jnt/www/Papers/PhD-84-jnt.pdf, 1984.
[10] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9), 1986.
[11] T. Vicsek, A. Czirok, I. Ben Jacob, I. Cohen, and O. Schochet. Novel type of phase transitions in a system of self-driven particles. *Phys. Rev. Lett.*, 75:1226–1229, 1995.
[12] J. Wolfowitz. Products of indecomposable, aperiodic, stochastic matrices. *Proc. Amer. Math. Soc.*, 14:733–737, 1963.