Natural Algorithms *

Bernard Chazelle †

Abstract

We provide further evidence that the study of complex self-organizing systems can benefit from an algorithmic perspective. The subject has been traditionally viewed through the lens of physics and control theory. Using tools typically associated with theoretical computer science, we settle an old question in theoretical ecology: bounding the convergence of bird flocks. We bound the time to reach steady state by a tower-of-twos of height linear in the number of birds. We prove that, surprisingly, the tower-of-twos growth is intrinsic to the model. This unexpected result demonstrates the merits of approaching biological dynamical systems as "natural algorithms" and applying algorithmic techniques to them.

1 Introduction

What do migrating geese, flocking cranes, bait balls of fish, prey-predator systems, and synchronously flashing fireflies have in common? All of them are instances of *natural algorithms*, ie, algorithms designed by evolution over millions of years. By and large, their study has been the purview of dynamical systems theory within the fields of zoology, ecology, evolutionary biology, etc. The main purpose of this work is to show that tools from theoretical computer science might be of benefit to the study of natural algorithms. We consider two standard bird flocking models that have been extensively studied in the literature and for which convergence bounds have been elusive. We establish both upper and lower bounds on the time to reach steady state. In doing so, we exhibit a remarkably exotic behavior.

Bird flocking has received considerable attention in the scientific and engineering literature, including the now-classical *Boids* model of Reynolds [15, 19–21]. Close scrutiny has been given to leaderless models where birds update their velocities by averaging them out over their nearest neighbors. Computer simulations support the intuitive belief that, by repeated averaging, each bird should eventually converge to a fixed speed and heading. This has been proven theoretically, but how long it takes for the system to converge has remained

an open question. This paper addresses this problem by giving an upper bound on the convergence time. It is surprisingly high: a tower-of-twos in the number of birds. A bigger surprise still is that this rate of growth is inevitable.

The existential question (does the system converge?) has been settled in many different ways, and it is useful to review its history briefly. A "recurrent connectivity" assumption stipulates that, over any time interval of a fixed length, every pair of birds should be able to communicate with each other, directly or indirectly via other birds. Jadbabaie, Lin, and Morse [5] proved the first of several convergence results under that assumption (eg, [11,12,17,20]). Several authors extended these results to variable-length intervals [4,8,10]. They established that the bird group always ends up as a collection of separate flocks (perhaps only one), each one converging toward its own speed and heading. Some authors have shown how to do away with the recurrent connectivity assumption by changing the model suitably. Tahbaz-Salehi and Jadbabaie [18], for example, assume that the birds fly on the surface of a torus. Cucker and Smale [3] use a broadcast model that extends a bird's influence to the entire group while scaling it down as a function of distance. In a similar vein, Ji and Egerstedt [6] introduce a hysteresis rule to ensure that connectivity only increases over time. Recent work suggests the use of topological as well as metric criteria [1], but the bulk of work on leaderless flocking has been based on nearest-neighbor rules. We are not aware of any bounds on the convergence time.

Rather than trying to build an abstract framework to capture as many models as possible, we consider two specific examples that are highly representative of the many variants considered in the literature. Model **K** (for kinematic) is a variant of the classical Vicsek model [22]. The control does not allow for variations in speed—only headings can change—so the model is nonholonomic. Model **D** (for dynamic) averages over velocities: it includes inertia and is fully actuated, as in [3]. In both models, the flocking network is a geometric graph: given n birds $\mathcal{B}_1, \ldots, \mathcal{B}_n$, represented at time t by points $x_1(t), \ldots, x_n(t)$ in some Euclidean space, the *flocking network* G_t has a vertex for each bird and an edge between any two of them within distance 1

of each other. By convention, G_t has no self-loops. The connected components of G_t are called *flocks*. If $d_i(t)$ denotes the number of birds adjacent to $\mathcal{B}_i$ at time t, the total number of birds within the closed unit disk centered at $\mathcal{B}_i$ is precisely $d_i(t) + 1$.

Model K. The input consists of a speed $\sigma > 0$ and two vectors: the initial position of the birds, $x_1(t), \ldots, x_n(t)$, for $t = 0$, and their flight headings, $\theta_1(t), \ldots, \theta_n(t)$, for $t = 1$. In this model, birds live in two dimensions, so we represent each $x_i(t)$ as a complex number and assume that $0 \leq \theta_i(t) < 2\pi$. For any integer $t \geq 1$ and $1 \leq i \leq n$,

$$x_i(t) = x_i(t-1) + \sigma e^{i\theta_i(t)},$$

where

$$\theta_i(t+1) = \frac{1}{d_i(t)+1}\Big(\theta_i(t) + \sum_{(i,j)\in G_t} \theta_j(t)\Big),$$

We can rewrite this recurrence more simply in matrix form.[1] Let $P(t)$ be the n-by-n matrix defined by $p_{ij}(t) = (d_i(t)+1)^{-1}$ if $i = j$ or $(i,j) \in G_t$. For $t \geq 1$,

$$(1.1) \qquad \begin{cases} x(t) = x(t-1) + \sigma e^{i\theta}(t); \\ \theta(t+1) = \kappa + P(t)\theta(t). \end{cases}$$

We inserted the vector $\kappa \in \mathbb{C}^n$ for added generality. It may depend on t as well as on the positions and velocities at any time. One can think of it as a control parameter or an adversarial mechanism; or, more simply, as in the Vicsek model, as noise.

Model D. The input consists of the initial position $x(0)$ and velocity $v(1)$. Both vectors belong to E^d (for constant $d \geq 1$). For $t \geq 1$ and $1 \leq i \leq n$,

$$x_i(t) = x_i(t-1) + v_i(t),$$

where

$$v_i(t+1) - v_i(t) = c_i(t) \sum_{(i,j)\in G_t} (v_j(t) - v_i(t)),$$

where $c_i(t)$ satisfies $0 < c_i(t)d_i(t) < 1$. For simplicity, we assume that $c_i(t)$ may vary only when G_t does. The model has a simple mechanical interpretation. The difference $v_i(t+1) - v_i(t)$ is the discrete analogue of

the acceleration. By Newton's Law, $F = ma$, a bird is subject to a force that grows in proportion to the difference between its velocity and its neighbors'. Again, it is convenient to express the dynamics in matrix form. Let $P(t) = I_n - C_t L_t$, where $C_t = \operatorname{diag} c(t)$ and L_t is the Laplacian of G_t:

$$(L_t)_{ij} = \begin{cases} d_i(t) & \text{if } i = j; \\ -1 & \text{if } (i,j) \in G_t; \\ 0 & \text{else.} \end{cases}$$

One final piece of notation is the Kronecker product $\otimes$, which we need to extend the averaging to all the coordinates.[2] We form the vector $x(t)$ by stacking $x_1(t), \ldots, x_n(t)$ together into one big column vector of dimension dn. Given a matrix A, the product $(A \otimes I_d)x(t)$ stacks together the result of multiplying A by the vector formed by k-th coordinate of each $x_i(t)$. In this way, for $t \geq 1$,

$$(1.2) \qquad \begin{cases} x(t) = x(t-1) + v(t); \\ v(t+1) = \mathcal{K}(P(t) \otimes I_d)v(t). \end{cases}$$

As with model **K**, we've inserted a dn-by-dn matrix $\mathcal{K}$ for control or noise.

We add to both models a tie-breaking *hysteresis rule*: an existing edge (i,j) of G_t remains in place if the distance between $\mathcal{B}_i$ and $\mathcal{B}_j$ changes by less than $\varepsilon_h \stackrel{\text{def}}{=} n^{-n^4}$ from time $t-1$ to time t. (In the case of model **K**, we add an extra rule to simplify the proofs.)

Remark 1. What does it mean for the bird group to converge? The standard definition is that each bird's velocity should converge to a fixed (perhaps distinct) vector. It seems difficult to get a handle on this question, however, without also considering the convergence of the flocking network. So, we say that the bird group has reached *steady state* if the flocking network no longer changes. Network convergence easily implies velocity convergence. The converse is not true: velocities might reach steady state while the network does not. This is an interesting but somewhat peripheral issue that it is best to bypass, as is done in [6], by injecting a minute amount of hysteresis into the system. This is necessary: flocking networks may not always converge without hysteresis. In Section 3, indeed, we specify a group of birds

---

[1] Adding or removing subscripts is how we distinguish between a vector and its coordinates; eg, $x(t) = (x_i(t))$. We denote the elements of a matrix $P(t)$ by $p_{ij}(t)$ and write $e^{i\theta(t)}$ to refer to the column vector of $\mathbb{C}^n$ with coordinates $e^{i\theta_i(t)}$; judging the disambiguation obvious, we use i both as an index and as the imaginary unit.

[2] The Kronecker product of two matrices A and B is the matrix we get if we replace each a_{ij} by the block $a_{ij}B$. Formally, if A is m-by-n and B is p-by-q, then the product $A \otimes B$ is the mp-by-nq matrix C such that $c_{ip+j,kq+l} = a_{i,k}b_{j,l}$. We will often use, with no further mention, the tensor identity $(A \otimes B)(C \otimes D) = AC \otimes BD$.

that alternates forever between one and two flocks without ever converging. Hysteresis prevents edge breaks based on microscopic changes. Note that if n^{-n^4} does not look microscopic enough for small n, then we can always set $\varepsilon_h = 2^{-100} n^{-n^4}$ without changing any of our results. While one might debate the best choice of hysteresis—our rule is one of many possible variants—it is important to choose a *sound* rule, where soundness is defined for our purposes by the guarantee that any two birds at distance ever so slightly away from 1 should have the right pairing status (ie, as determined by the unit distance rule). In that sense, hysteresis is, indeed, a tie-breaking rule.

Remark 2. In model **K**, the transition matrix $P(t)$ can also be expressed as $I_n - C_t L_t$ if we set $c_i(t) = (d_i(t) + 1)^{-1}$. Our investigation into $P(t)$ will thus apply equally to both models. The condition $0 \leq c_i(t) d_i(t) < 1$ enforces two desirable properties: $P(t)$ is row-stochastic and its diagonal is strictly positive. A natural question thus arises. If we ignore the control parameters, can't we solve the problem of convergence by simple linear algebra? Let's sketch an approach. Cucker and Smale [3] diagonalize the Laplacian and note that, since only differences are of interest, the vectors might as well be assumed to lie in the space $\mathbf{1}^\perp$. Not only is that space invariant under the Laplacian but it contracts at an exponential rate set by the Fiedler number (the second eigenvalue). From this, a quadratic Lyapunov function quickly emerges. When the graph is connected, the Fiedler number is bounded away from 0 by an inverse polynomial, so differences between velocities decay to 0 at a rate of $2^{t/n^c}$ for some constant $c > 0$. Of course, unlike Cucker and Smale, who assume the complete graph, we would have to track the connectivity of the flocking network. But, that aside, we could simply let "Fiedler do the work." Alas, such an approach is doomed in our models, both for deep and for obvious reasons. The deep reason is that, in general, the dynamical systems under consideration do not admit of any suitable quadratic Lyapunov function [5, 14]. Cucker and Smale's transition matrices are symmetric (and not stochastic) so this objection does not hold. It is the reverse in our case: our matrices are stochastic but not symmetric. This comes at a price. Even though $P(t)$ is diagonalizable, the right eigenspace for the subdominant eigenvalues is not orthogonal to $\mathbf{1}$. So, $\mathbf{1}^\perp$ is not invariant and unfortunately the map $P(t)$ might not be contractive. For example, suppose that the matrix in model **D** is of the form

$$P(t) = \tfrac{1}{15} \begin{pmatrix} 12 & 3 \\ 10 & 5 \end{pmatrix}.$$

The two eigenvalues are 1 and 0.133; yet $P(t)$ stretches the unit vector $(1, 0)$ to one of length 1.041. In other words, even though its spectrum is confined to $[0, 1]$, the map $P(t)$ increases some Euclidean distances. Linear algebra alone seems unable to prove convergence; we complement it with combinatorial arguments.

Remark 3. The initial angle between two birds can be made arbitrarily small, thus delaying their interaction as long as we wish. If we want to bound the convergence time, therefore, we need to take into account the encoding lengths of the input positions, velocities, and angles. We assume that all these inputs are rationals with $O(\log n)$-bit long numerators and denominators.[3] We use this particular bound only for convenience. A mere glance at our astronomical bounds shows that much longer encodings would not make a dent into our results.

To express these results, we need to define the third level of the Ackermann hierarchy, the so-called "tower-of-twos" function: $2 \uparrow\uparrow 1 = 2$ and, for $n > 1$, $2 \uparrow\uparrow n = 2^{2\uparrow\uparrow(n-1)}$. We mention our result for the simple case: $\kappa = 0$ and $\mathcal{K}(t) = I_{dn}$.

THEOREM 1.1. *In both models* **K** *and* **D**, *a group of* n *birds reaches steady state in at most* $2 \uparrow\uparrow O(n)$ *steps. The total number of network switches is* $n^{O(n^3)}$.

This result is extremely robust in terms of hysteresis and initial conditions: drastic changes there would still lead to the same bounds. Our hysteresis rule is sound in a very strong sense, too: (i) any two birds within unit distance of each other at time t are always joined by an edge of the flocking network G_t; (ii) no two birds at distance greater than $1 + n^{-n^3}$ are ever adjacent in G_t. We note that, with this particular choice of hysteresis, the flocking network can lose edges only during the first n^{2n^4} steps. After that, birds that join together into a flock can never split up.

It is quite surprising that, on the one hand, the network stabilizes in "only" an exponential number of steps (in the sense that adjacent birds stay joined forever), but on the other hand if a few more pairs of birds want to be joined later into a common flock, they may have to wait an extraordinarily long time. Another surprise is that the astronomical "tower-of-twos" growth rate is intrinsic to flocking. Indeed, we provide initial conditions for the birds, using only $O(\log n)$ bits per bird, that lead to steady state only after $2 \uparrow\uparrow \Omega(\log n)$ steps. The proof is technical (and

---

[3]All logarithms are taken to the base 2. Throughout this paper, n is assumed to be large enough.

omitted from this extended abstract), but the initial configuration is straightforward. The flocks are single paths, with the matrix $P(t)$ corresponding to a lazy random walk with probability $\frac{1}{3}$ of staying in place. The birds fly in two dimensions at constant y-speed 1. Their projections onto the X-axis have the following initial conditions; c is a constant:

$$
\begin{cases}
x(0) = \left(0, \frac{2}{3}, \ldots, 2l, 2l + \frac{2}{3}, \ldots, n-2, n - \frac{4}{3}\right)^T; \\
v(1) = \left(\underbrace{n^{-c}, 0, -n^{-c}, 0, \ldots, n^{-c}, 0, -n^{-c}, 0}_{n}\right)^T.
\end{cases}
$$

In the remainder of this abstract, we will focus on model **D** and the case $\mathcal{K}(t) = I_{dn}$. Because of space limitations, many of the proofs will be omitted.

2 Preliminaries

We begin with some simple geometry relating the relative displacement of birds with the transition matrix. Define $\Delta_{ij}(t) = |\text{DIST}_t(\mathcal{B}_i, \mathcal{B}_j) - \text{DIST}_{t-1}(\mathcal{B}_i, \mathcal{B}_j)|$.

LEMMA 2.1. *For $t \geq 1$, $\Delta_{ij}(t) \leq \|v_i(t) - v_j(t)\|_2$.*

Numerical Complexity. We mention a few basic facts, most of them known, about products of stochastic matrices. The *footprint* of a matrix A is the matrix $\underline{A}$ derived from A by replacing each nonzero entry by 1. For $t \geq s$, we use $P(t,s)$ as shorthand for $P(t)P(t-1)\cdots P(s)$. This matrix plays a crucial rule in both models: indeed, as can be seen from equations (1.1, 1.2): $\theta(t+1) = P(t,1)\,\theta(1)$ and $v(t+1) = (P(t,1) \otimes I_d)v(1)$. A bird may influence another one over a period of time without the converse being true; in other words, the matrices $P(t,s)$ and $\underline{P}(t,s)$ are in general not symmetric; the exception is $\underline{P}(t)$, which not only is symmetric but has its diagonal full of ones. Before we get to the structural properties of $P(t,s)$, we need to answer two basic questions: how small can the nonzero entries be and how many bits do we need to represent them? As was shown in [4, 9], nonzero elements of $P(t,s)$ can be bounded uniformly, ie, independently of t. Note that this relies critically on the positivity of the diagonals. Indeed, without the condition $c_i(t)d_i(t) < 1$, we could choose $P(t) = A$ for even t and $P(t) = B$ for odd t, where

$$
A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix}.
$$

For even $t > 0$,

$$
P(t,1) = (AB)^{t/2} = \begin{pmatrix} 2^{-t/2} & 1 - 2^{1-t/2} & 2^{-t/2} \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}.
$$

LEMMA 2.2. *For any $1 \leq s \leq t$, the elements of $P(t,s)$ are same-denominator rational coordinates over $O((t-s+1)n\log n)$ bits. Its nonzero elements are in $n^{-O(n^2)}$.*

LEMMA 2.3. *For any $t \geq 1$, the vectors $v(t)$ and $x(t)$ have same-denominator rational coordinates over $O(tn\log n)$ bits.*

Velocities are polynomially bounded at the outset. They remain so all the time since the transition matrices can only average out their coordinates. This simple bound will prove very useful. For any $t \geq 1$,

$$
(2.3) \qquad \|v(t)\|_2 = n^{O(1)}.
$$

Ergodicity. Let $\tau_p(A)$ denote the ℓ_p-diameter of the convex hull formed by the rows of a matrix A, ie, $\tau_p(A) = \max_{i,j} \|a_{i*} - a_{j*}\|_p$, where a_{i*} denotes the i-th row of A. In the case $p = 1$, for reasons soon to be apparent, we divide the diameter by two. To understand why $\tau_p(A)$ relates to ergodicity (and why we divide by 2), assume that A is row-stochastic. We observe then that

$$
0 \leq \tau_1(A) = 1 - \min_{i,j} \sum_k \min\{a_{ik}, a_{jk}\} \leq 1.
$$

There are many fascinating relations between these diameters [16]. For our purposes, the following submultiplicativity result will suffice [9].

LEMMA 2.4. *Given two row-stochastic matrices A, B that can be multiplied, $\tau_2(AB) \leq \tau_1(A)\tau_2(B)$.*

3 Steady-State Flocking

We leave the convergence time for later and try, first, to bound the number of changes in the flocking network. We also focus on the time-invariant case and establish the soundness of the hysteresis rule.

Counting Network Switches. Let $N(n)$ denote the maximum number of switches in the flocking network, ie, the number of times t such that $P(t) \neq P(t+1)$. (Note that, by our requirement that C_t may vary only when G_t does, we could use footprints equivalently in the definition.) The function $N(n)$ maximizes the number of switches uniformly over all initial conditions. Bounding this function involves putting a quantitative framework around the existential analyses in [4, 8–10] We prove the network switching bound of Theorem 1.1.

LEMMA 3.1. *The number of flocking network switches is bounded by $N(n) = n^{O(n^3)}$.*

Proof. Fix $s > 0$ once and for all. For $t > s$, let $N(t, s)$ be the number of network changes between times s and t, ie, the number of integers u ($s < u \leq t$) such that $\underline{P}(u) \neq \underline{P}(u - 1)$. Since the diagonal of each $P(t)$ is positive, $\underline{P}(t, s)$ can never lose a 1 as t grows, so there exists a smallest T_1 such that $\underline{P}(t, s) = \underline{P}(T_1, s)$ for all $t > T_1$. Consider the first column and let $1 \leq n_1 < \cdots < n_l \leq n$ be its successive Hamming weights (ie, number of ones). How many switches can take place between the time t_k when the column acquires weight n_k and the time t_{k+1} when the weight reaches n_{k+1}? Note that $t_l \leq T_1$ and, for consistency, put $t_0 = s$. In other words, how large can $N(t_{k+1}, t_k)$ be for $0 \leq k < l$? Let H denote the subgraph of G_{t_k+1} consisting of the connected components (ie, flocks) that include the n_k birds indexed by the first column of $\underline{P}(t_k, s)$. If H contains more than n_k birds then, at time $t_k + 1$, at least one of these extra birds, $\mathcal{B}_i$, is adjacent in G_{t_k+1} to one of the n_k birds, say, $\mathcal{B}_j$. Then, $p_{ij}(t_k + 1) > 0$ and $p_{j1}(t_k, s) > 0$; hence $p_{i1}(t_k + 1, s) \geq p_{ij}(t_k + 1)p_{j1}(t_k, s) > 0$. Since $p_{i1}(t_k, s) = 0$, the first column of $\underline{P}(t, s)$ acquires a new 1, so $t_{k+1} = t_k + 1$ and $N(t_{k+1}, t_k) = 1$. Assume now that H has exactly n_k vertices. The flocking network G_{t_k+1} consists of a set of flocks totalling n_k birds and a separate set of flocks including the $n - n_k$ others. After at most $N(n_k) + N(n - n_k)$ network switches and no contact between the two sets, the only possible further change is an interaction between two birds, one from each set. It follows by monotonicity of $N(n)$ that

$$N(t_{k+1}, t_k) \leq 1 + N(n_k) + N(n - n_k) \leq 2N(n-1) + 1$$

and

$$N(t_l, s) = \sum_{k=1}^{l} N(t_k, t_{k-1}) \leq 2nN(n-1) + n.$$

Of course, there is nothing special about bird $\mathcal{B}_1$. We can apply the same argument for each column and conclude that the time T_1 when the matrix $\underline{P}(t, s)$ has finally stabilized satisfies

$$(3.4) \qquad N(T_1, s) \leq 2nN(n-1) + n.$$

A technical subtlety is that the recursive definition of $N(n)$ must treat the $O(\log n)$-bit requirement on initial velocities as a global condition, ie, independent of the argument of the function N. For this reason, we will use ν to denote the global value of n. The index set V_1 corresponding to the ones in the first column of $\underline{P}(T_1, s)$ is called the first *stabilizer*. For $t > T_1$, no edge of G_t can join V_1 to its complement, since this would immediately add more ones to the first column of $\underline{P}(t, s)$. Relabel the

rows and columns so that all the ones in $\underline{P}(T_1, s)$'s first column appear on top. Then, for any $t > T_1$, $P(t)$ is a diagonal 2-block matrix with the top left block, indexed by $V_1 \times V_1$, providing the transitions among the vertices of V_1 at time t. (Note that a set defined similarly to V_1 for another column might partly overlap with V_1.) This property is invariant under composition, so $P(t, T_1 + 1)$ is also a diagonal block matrix of the same type. Let $A_{|V \times W}$ denote the submatrix of A with rows indexed by V and columns by W. Writing $V_0 = [n]$,

$$P_{|V_1 \times V_0}(t, s) = P_{|V_1 \times V_1}(t, T_1 + 1)P_{|V_1 \times V_0}(T_1, s).$$

By setting s to $T_1 + 1$ we can repeat the same argument, the only difference being that the transition matrices are now $|V_1|$-by-$|V_1|$. This leads to the second stabilizer $V_2 \subseteq V_1$, which, by relabeling, can be assumed to index the top of the subsequent matrices. We define T_2 as the smallest integer such that $\underline{P}_{|V_1 \times V_1}(t, T_1 + 1) = \underline{P}_{|V_1 \times V_1}(T_2, T_1 + 1)$ for all $t > T_2$. The set V_2 indexes the ones in the first column of $\underline{P}_{|V_1 \times V_1}(T_2, T_1 + 1)$. Iterating in this fashion leads to an infinite sequence of time $T_1 < T_2 < \cdots$ and stabilizers $V_1 \supseteq V_2 \supseteq \cdots$ such that, for any $t > T_k$,

$$(3.5)$$
$$P_{|V_k \times V_0}(t, s) = P_{|V_k \times V_k}(t, T_k+1)P_{|V_k \times V_{k-1}}(T_k, T_{k-1}+1)$$
$$\cdots P_{|V_2 \times V_1}(T_2, T_1 + 1)P_{|V_1 \times V_0}(T_1, T_0 + 1),$$

where $P_{|V_i \times V_{i-1}}(T_i, T_{i-1} + 1)$ is a $|V_i|$-by-$|V_{i-1}|$ matrix and $T_0 = s - 1$. What is the benefit of rewriting the top rows of $P(t, s)$ in such a complicated way? The first column of each $P_{|V_i \times V_{i-1}}(T_i, T_{i-1} + 1)$ consists entirely of positive entries. By Lemma 2.2, they are at least $\nu^{-O(\nu^2)}$, so half the ℓ_1-distance between any two rows is at most $1 - \nu^{-O(\nu^2)} \leq e^{-\nu^{-O(\nu^2)}}$ and

$$\tau_1(P_{|V_i \times V_{i-1}}(T_i, T_{i-1} + 1)) \leq e^{-\nu^{-O(\nu^2)}}.$$

Lemma 2.4 implies that

$$\tau_2(A) \leq \tau_1(A)\tau_2(I) \leq \sqrt{2}\,\tau_1(A),$$

and

$$(3.6) \quad \tau_2(P_{|V_k \times V_0}(t, s)) \leq \sqrt{2}\,\tau_1(P_{|V_k \times V_k}(t, T_k + 1))$$
$$\prod_{i=1}^{k} \tau_1(P_{|V_i \times V_{i-1}}(T_i, T_{i-1} + 1)) \leq \sqrt{2}\,e^{-k\nu^{-O(\nu^2)}}.$$

Let $\chi(i, j)$ denote the n-dimensional vector with all coordinates equal to 0, except for $\chi(i, j)_i = 1$ and $\chi(i, j)_j = -1$. Note that

$$v_i(t) - v_j(t) = ((\chi(i, j)P(t - 1, 1)) \otimes I_d)v(1).$$

Set $k = \nu^{b\nu^2}$ for a large enough constant $b > 0$. By the initial conditions, $\|v(1)\|_2 = \nu^{O(1)}$ and so by (3.6) and Cauchy-Schwarz, for any $t > T_k + 1$,

$$\max_{i,j \in V_k} \|v_i(t) - v_j(t)\|_2 \le \tau_2(P_{|V_k \times V_0}(t-1,1))\nu^{O(1)} < e^{-\nu^2}.$$

By Lemma 2.1, it then follows that

$$\Delta_{ij}(t) \le \|v_i(t) - v_j(t)\|_2 < \nu^{-\nu^4} = \varepsilon_h.$$

By the hysteresis rule, this means that if birds $\mathcal{B}_i$ and $\mathcal{B}_j$ are joined after time $T_k + 1$, they will always remain so. This leaves at most $\binom{|V_k|}{2}$ extra network changes (final pairings), so the total number is conservatively bounded by

$$N(T_k, T_{k-1}) + \cdots + N(T_1, 1) + \binom{|V_k|}{2}.$$

But (3.4) holds for any pair $(T_i, T_{i-1} + 1)$, so

$$N(n) < n^2 + k(2nN(n-1) + n) < \nu^{O(n\nu^2)}.$$

$\square$

Time-Invariant Flocking. There will be times, for example, post-convergence, when flocks are governed by a fixed network. We investigate this case separately. In this section, $G_t = G$ is time-invariant; for notational convenience, we assume there is a single flock, ie, G_t is connected. In both models **K** and **D**, we can express the stochastic matrix P as $I_n - CL$: it has the simple dominant eigenvalue 1 with right and left eigenvectors $\mathbf{1}$ and $\pi = (\operatorname{tr} C^{-1})^{-1} C^{-1} \mathbf{1}$, respectively. Lack of symmetry does not keep P from being diagonalizable, though it denies us eigenvector orthogonality. Define

$$(3.7) \qquad M = C^{-1/2} P C^{1/2} = I_n - C^{1/2} L C^{1/2}.$$

Being symmetric, M can be diagonalized as $\sum_{k=1}^n \lambda_k u_k u_k^T$, where the u_k's are orthonormal eigenvectors and the eigenvalues are real. By the connectivity of G_t, stochasticity of P, and other standard properties [2], $1 = \lambda_1 > \lambda_2 \ge \cdots \ge \lambda_n \ge -1$ and $u_1 = (\sqrt{\pi_1}, \ldots, \sqrt{\pi_n})^T$. It follows that P can be diagonalized as well, with the same eigenvalues. Since $\sum_k u_k u_k^T = I_n$, the following identity holds for all nonnegative s, including $s = 0$:

$$(3.8) \qquad P^s = \mathbf{1}\pi^T + \sum_{k>1} \lambda_k^s C^{1/2} u_k u_k^T C^{-1/2}.$$

The left and right eigenvectors of P for λ_k are given (in column form) by $C^{-1/2} u_k$ and $C^{1/2} u_k$ and form inverse matrices; in general, neither group forms an orthogonal basis. We can bound the second largest eigenvalue by standard algebraic graph theory.

LEMMA 3.2. *If $\mu \stackrel{\text{def}}{=} \max_{k>1} |\lambda_k|$, then $\mu = 1 - n^{-O(1)}$.*

By (3.8), for all $i, j, s > 0$, $(P^s)_{ij} \ge \pi_j - \sum_{k>1} |\lambda_k|^s \sqrt{c_i/c_j} \, |(u_k)_i (u_k)_j| \ge \pi_j - n^{O(1)} \mu^s$. A similar derivation gives us the corresponding upper bound; so, by Lemma 3.2,

$$(3.9) \qquad |(P^s)_{ij} - \pi_j| \le n^{O(1)} e^{-sn^{-O(1)}}.$$

Similarly, for $s > n^{c_0}$, for a constant c_0 large enough,

$$(3.10) \quad \tau(P^s) \le 1 - \sum_{k=1}^n (\pi_k - n^{O(1)} e^{-sn^{-O(1)}}) < \tfrac{1}{2}.$$

The Rationality of Model D. Unlike in model **K**, the locations of the birds in model **D** remain rational at all times. Does that mean their limit remains so? We prove that this is, indeed, the case. For $t > 0$, define

$$(3.11) \qquad \Gamma_t = -\mathbf{1}\pi^T t + \sum_{s=0}^{t-1} P^s.$$

We begin with a simple characterization of the limit of Γ_t, using the notation $(Y \mid y)$ to refer to the n-by-n matrix derived from Y by replacing its last column with the vector y.

LEMMA 3.3. *As $t \to \infty$, Γ_t converges to*

$$\Gamma = (I_n - \mathbf{1}\pi^T \mid \mathbf{0})(I_n - P \mid \mathbf{1})^{-1}.$$

The motion equation (1.2) becomes, for $t \ge 1$,

$$(3.12) \qquad x(t) = x(0) + \Big(\sum_{s=0}^{t-1} P^s \otimes I_d \Big) v(1)$$

or, equivalently, by (3.11),

$$(3.13) \ \ x(t) = x(0) + t((\mathbf{1}\pi^T) \otimes I_d)v(1) + (\Gamma_t \otimes I_d)v(1).$$

We call $\mathbf{m}_\pi[x(t)] = (\pi^T \otimes I_d)x(t)$ the *mass center* of the flock. It moves at constant speed along a fixed line in d-space: Indeed, by (3.12),

$$(\pi^T \otimes I_d)x(t) = (\pi^T \otimes I_d)x(0) + \sum_{s=0}^{t-1}(\pi^T \otimes I_d)(P^s \otimes I_d)v(1)$$

$$= (\pi^T \otimes I_d)x(0) + t(\pi^T \otimes I_d)v(1).$$

The vector $\mathbf{m}_\pi[v(1)] = (\pi^T \otimes I_d)v(1)$ is called the *stationary velocity* of the flock. Moving the origin to the mass center of the birds, we express $x(t)$, relative to this moving frame, as

$$x^r(t) = x(t) - (\mathbf{1} \otimes I_d)\mathbf{m}_\pi[x(t)].$$

By (3.13) and simple tensor manipulation, we find that

$$x^r(t) = ((I_n - \mathbf{1}\pi^T)\otimes I_d)x(0) + (\Gamma_t \otimes I_d)v(1)$$

and, by Lemma 3.3,

LEMMA 3.4. *If G is connected, the relative flocking configuration $x^r(t)$ converges to the limit*

$$\begin{aligned} x^r = {}&((I_n - \mathbf{1}\pi^T)\otimes I_d)x(0) \\ &+ ((I_n - \mathbf{1}\pi^T \,|\, \mathbf{0})\,(I_n - P\,|\,\mathbf{1})^{-1} \otimes I_d)v(1). \end{aligned}$$

The mass center of the configuration moves in $\mathbb{R}^d$ at constant speed in a fixed direction.

LEMMA 3.5. *The coordinates of the limit configuration x^r as well as the elements of Γ are same-denominator rationals over $O(n \log n)$ bits.*

Soundness of Hysteresis Rule. We begin with a proof that hysteresis is required to ensure convergence. We build a 4-bird flock in one dimension in model **D**, whose network cannot converge without a hysteresis rule. The speed of the birds will decay exponentially. In real life, the birds would stall. It is immediate to lift our construction in two dimensions, however, to ensure that the birds' speeds never fall below a constant, thus making stalling a nonissue. These are the initial conditions:

$$\begin{cases} x(0) = \frac{1}{16}(0, 8, 21, 29); \\ v(1) = \frac{1}{8}(1, -1, 1, -1). \end{cases}$$

The flocking network alternates between a pair of 2-bird edges and a single 4-bird path, whose respective transition matrices are:

$$\frac{1}{3}\begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{pmatrix} \quad \text{and} \quad \frac{1}{3}\begin{pmatrix} 1 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

The beauty of the initial velocity $v(1)$ is that it is a right eigenvector for both flocking networks for the same eigenvalue $-\frac{1}{3}$; therefore, $v(s) = (-3)^{1-s}v(1)$ and, by (1.2),
(3.14)

$$x(t) = x(0) + \sum_{s=1}^{t} v(s) = x(0) + \frac{3}{4}\Big(1 - (-\tfrac{1}{3})^t\Big)v(1).$$

Thus, for $t \geq 1$,

$$x_{i+1}(t) - x_i(t) = \begin{cases} \frac{1}{16}\Big(5 - (-\tfrac{1}{3})^{t-1}\Big) & \text{if } i = 1, 3; \\ 1 + \frac{1}{16}(-\tfrac{1}{3})^{t-1} & \text{if } i = 2. \end{cases}$$

The distance between the first and second birds stays comfortably between $\frac{1}{4}$ and $\frac{1}{2}$; same with birds $\mathcal{B}_3$ and $\mathcal{B}_4$. The distance between the middle birds $\mathcal{B}_2$ and $\mathcal{B}_3$ oscillates around 1, so the network forever alternates between one and two connected components.

LEMMA 3.6. *The hysteresis rule is sound: (i) any two birds within unit distance of each other at time t are always joined by an edge of the flocking network G_t; (ii) no two birds at distance greater than $1 + n^{-n^3}$ are ever adjacent in G_t.*

The Geometry of Flocking. Can birds fly in giant loops and come back to their point of origin? Are there constraints on their trajectories? We show that, after enough time has elapsed, two birds can be newly joined only if they fly almost parallel to each other. We assume model **D**, but a similar reasoning leads to the same results in model **K**. It is convenient to lift the birds into $\mathbb{R}^{d+1}$ by adding time as an extra dimension: $x(t) \mapsto (x_1(t), \ldots, x_d(t), t); v(t) \mapsto (v_1(t), \ldots, v_d(t), 1)$. Since $\mathbf{1}$ is a right eigenvector, this lifting still satisfies the equation of motion. The hysteresis rule kicks in at precisely the same time and in the same manner as before. The behavior of the birds is unchanged but it lends itself to a better geometric interpretation.

We begin with a simple observation. Define $w = \frac{1}{t}x_i(t)$; note that w depends on both i and t (we omit these arguments for notational simplicity). Consider the triangle formed by having the vectors w and $v_i(t)$ pointing away from the same corner A of the triangle. By (2.3) and the lifting, the two sides incident to A have length between 1 and $n^{O(1)}$; therefore, the smaller angle β not at A satisfies $n^{-O(1)} < \beta < \pi/2$. By the law of sines, it follows easily that, if $v_i(t) \neq w$,

$$(3.15) \qquad n^{-O(1)} < \frac{\angle(x_i(t), v_i(t))}{\|v_i(t) - w\|_2} = O(1).$$

LEMMA 3.7. *For any bird $\mathcal{B}_i$, at any time $t > n^{bn^3}$, for a large enough constant b,*

$$\angle(x_i(t), v_i(t)) = \frac{\log t}{t}\, n^{O(n^3)}.$$

Suppose that birds $\mathcal{B}_i$ and $\mathcal{B}_j$ are joined in G_t, for some $t > n^{bn^3}$, where b is defined in Lemma 3.7. By (3.15) and Lemma 2.1,

$$\Delta_{ij}(t) \leq \|v_i(t) - v_j(t)\|_2 = \frac{\log t}{t}\, n^{O(n^3)}.$$

By hysteresis, edges of G_t can break only if $t < n^{2n^4}$. This means that, past that time, flocks can only merge.

LEMMA 3.8. *After time n^{2n^4}, the flocking network G_t can only gain new edges and never lose any.*

The remaining number of network switches after time n^{2n^4} can be at most $\binom{n}{2}$. We now examine how long one must wait for all these switches to be completed.

LEMMA 3.9. *Assume that the flocking network G_t stays invariant during the period $[t_1, t-1]$, where $t_1 > n^{2n^4}$. If two birds are adjacent in G_t but not in G_{t-1}, then either the birds are in distinct flocks at time $t-1$, in which case $t = n^{O(t_1 n)}$, or they belong to the same flock, in which case $t = t_1 2^{n^{O(1)}}$.*

Proof. Assume that the flocking network G_t stays invariant during the period $[t_1, t-1]$. Consider two birds $\mathcal{B}_i$ and $\mathcal{B}_j$ that are adjacent in G_t but not during $[t_1, t-1]$. Note that the two birds may or may not be in the same flock. Let the flock for $\mathcal{B}_i$ (resp. $\mathcal{B}_j$) consist of m (resp. m') birds; $m + m' \leq n$. By abuse of notation, we use the terminology of (3.8), ie, P, π, C, u_k, λ_k, to refer to the flock of m birds, and we add primes to refer to the flock of $\mathcal{B}_j$. We wish to place an upper bound on $t - t_1$. Let $\chi(i)$ denote the n-dimensional vector with all coordinates equal to 0, except for $\chi(i)_i = 1$. By (3.8, 3.12), for $t \geq t_1$,

$$x_i(t) = x_i(t_1) + \left(\chi(i)^T \sum_{s=0}^{t-t_1-1} P^s \otimes I_d \right) v(t_1+1)$$

$$= x_i(t_1) + (t-t_1)y + \sum_{k=2}^{m} \frac{1-\lambda_k^{t-t_1}}{1-\lambda_k} \Phi_k,$$

where
(3.16)
$$\begin{cases} y = (\pi^T \otimes I_d)v(t_1+1); \\ \Phi_k = ((\chi(i)^T C^{1/2} u_k u_k^T C^{-1/2}) \otimes I_d)v(t_1+1). \end{cases}$$

Note that, by (3.11), $\sum_{k=2}^{m} \frac{1}{1-\lambda_k}\Phi_k = ((\chi(i)^T\Gamma) \otimes I_d)v(t_1+1)$; therefore,

$$x_i(t) = x_i(t_1) + ((\chi(i)^T\Gamma) \otimes I_d)v(t_1+1)$$
$$+ (t-t_1)y - \sum_{k=2}^{m} \lambda_k^{t-t_1} \frac{\Phi_k}{1-\lambda_k}.$$

Adding primes to distinguish between the flocks of $\mathcal{B}_i$ and $\mathcal{B}_j$, we find that

$$(3.17) \quad x_i(t) - x_j(t) = A + B(t-t_1) - \sum_{k=2}^{m_0} \Psi_k \mu_k^{t-t_1},$$

where

(i) $A = x_i(t_1) - x_j(t_1) + ((\chi(i)^T\Gamma - \chi(j)^T\Gamma') \otimes I_d)v(t_1 + 1)$: By Lemma 2.3, the vectors $v(t_1+1)$, $x_i(t_1)$, and $x_j(t_1)$ have same-denominator rational coordinates over $O(t_1 n \log n)$ bits. In view of Lemma 3.5, this implies that the same is true of the vector A.

(ii) $B = y - y'$: The stationary distribution $\pi = (\operatorname{tr} C^{-1})^{-1}C^{-1}\mathbf{1}$ is a same-denominator rational vector over $O(n \log n)$ bits. Together with Lemma 2.3, this implies that B has same-denominator rational coordinates over $O(t_1 n \log n)$ bits; hence either $B = 0$ or $\|B\|_2 > n^{-O(t_1 n)}$.

(iii) $\mu_2 \geq \cdots \geq \mu_{m_0}$: Each μ_k is an eigenvalue λ_l or λ'_l ($l, l' > 1$) and $|\mu_k| < 1$.

(iv) Each Ψ_k is a d-dimensional vector of the form $\Phi_l/(1-\lambda_l)$ or $-\Phi'_l/(1-\lambda'_l)$. Since the eigenvalues are bounded away from 1 by $n^{-O(1)}$ (Lemma 3.2), the obvious bounds on C, $v(t_1+1)$ and u_k show that $\|\Psi_k\|_2 = n^{O(1)}$.

We distinguish among three cases:

Case I. $B \neq 0$: Note that the two flocks must be distinct. Indeed, having the two birds in the same flock implies that $\pi = \pi'$; hence $y = y'$. By (i, ii, iv) respectively, $\|A\|_2 = n^{O(t_1 n)}$, $\|B\|_2 > n^{-O(t_1 n)}$, and $\|\sum \Psi_k \mu_k^{t-t_1}\|_2 = n^{O(1)}$. It follows that, if the two birds are to be joined in G_t, then $t - t_1 = n^{O(t_1 n)}$. Beyond that, indeed, at least one coordinate of $x_i(t) - x_j(t)$ becomes too big for the unit-distance rule to kick in.

Case II. $B = 0$ and $\|A\|_2 \neq 1$: By (i), $\|A\|_2$ is bounded away from 1 by $n^{-O(t_1 n)}$. Since, by (iv), the length of any vector Ψ_k is $n^{O(1)}$ and the eigenvalue gap, $1 - \mu_k$, is at least $n^{-O(1)}$, $\|\sum_k \Psi_k \mu_k^{t-t_1}\|_2 = n^{O(1)}e^{-(t-t_1)n^{-O(1)}}$. It follows from (3.17) and the triangle inequality that

$$\big|\, \|x_i(t) - x_j(t)\|_2 - 1 \,\big| \geq n^{-O(t_1 n)} - n^{O(1)}e^{-(t-t_1)n^{-O(1)}}.$$

This implies that, for a large enough constant b_0, the distance between the two birds remains bounded away from 1 by $n^{-O(t_1 n)}$ for any $t > t_1 n^{b_0}$. Not only that, but the sign of $\|x_i(t)-x_j(t)\|_2 - 1$ can no longer change after $t_1 n^{b_1}$, for constant b_1 large enough, since the distance can vary by increments of at most $\|\sum_k \Psi_k \mu_k^{t-t_1}\|_2 = n^{O(1)}e^{-(t-t_1)n^{-O(1)}}$. We conclude that if the two birds must join in G_t then $t = t_1 n^{O(1)}$.

Case III. $B = 0$ and $\|A\|_2 = 1$: The distance between the two birds tends toward 1. The issue is that the two birds might stay safely away from each other for a long period of time and then decide to join into an edge of the network by moving within distance 1 of each other. We show that this is impossible. In other words, if $\|x_i(t) - x_j(t)\|_2$ ever falls below 1 for $t > t_1$, this must happen relatively soon. Recall that, by (3.17),

$$x_i(t) - x_j(t) = A - \sum_{k=2}^{m_0} \Psi_k \mu_k^{t-t_1},$$

where A is a unit vector. We investigate the behavior of the birds' distance locally around 1.

$$\|x_i(t) - x_j(t)\|_2^2 = 1 - 2\sum_k A^T \Psi_k \, \mu_k^{t-t_1}$$
$$+ \sum_{k,k'} \Psi_k^T \Psi_{k'} \, (\mu_k \mu_{k'})^{t-t_1}.$$

Let $1 > \rho_1 > \cdots \rho_N > 0$ be all the distinct nonzero values among $\{|\mu_k|, |\mu_k \mu_{k'}|\|$ $(N < n^2)$. We can rewrite the expression above as

$$(3.18) \qquad \|x_i(t) - x_j(t)\|_2^2 - 1 = \sum_{k=1}^{N} \Upsilon_k \, \rho_k^{t-t_1},$$

where

$$\Upsilon_k = \Upsilon_k^+ + (-1)^t \, \Upsilon_k^-.$$

We distinguish between odd and even values of t so as to keep each Υ_k time-invariant. We assume that t is even; the odd case is similar and justifiably skipped. Of course we may also assume that each $\Upsilon_k = \Upsilon_k^+ + \Upsilon_k^-$ is nonzero. We know that $\sum_k \Upsilon_k \rho_k^{t-t_1}$ tends to 0, as t goes to infinity, but we need to find out if it oscillates or keeps a constant sign in the long run. For this we must derive bounds on eigenvalue gaps and on $|\Upsilon_k|$. Much tighter results can be obtained from current spectral technology, but they would not make any difference for our purposes, so we settle for simple, conservative estimates.

LEMMA 3.10. *For all $k > 1$ and $k \geq 1$, respectively,*

$$\rho_k < (1 - 2^{-n^{O(1)}})\rho_1 \ \ and \ \ 2^{-t_1 2^{n^{O(1)}}} < |\Upsilon_k| = n^{O(1)}.$$

By (3.18), it follows from the lemma that

$$\|x_i(t) - x_j(t)\|_2^2 - 1 = (1 + o(1))\Upsilon_1 \, \rho_1^{t-t_1},$$

for $t > t_1 2^{n^{b_2}}$, where b_2 is a large enough constant. The same argument for odd values of t shows that after $t_1 2^{n^{b_2}}$, either $\|x_i(t) - x_j(t)\|_2^2$ stays on one side of 1 forever or it constantly alternates. In both cases, the permanent edge status between the two birds is settled by that time. This concludes Case III.

Putting all three cases together, we find that the bound from Case I is the most severe: $t = n^{O(t_1 n)}$. When the two birds are in the same flock at time $t - 1$, however, the bound from Case III takes precedence. □

Lemmas 3.8 and 3.9 show that all network switches take place within the first $2 \uparrow\uparrow O(n)$ steps. This completes the proof of Theorem 1.1.

References

[1] Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., Lecomte, V., Orlandi, A., Parisi, G., Procaccini, A., Viale, M., Zdravkovic, V. *Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study*, Proc. National Academy of Sciences 105 (2008), 1232-1237.

[2] Chung, F. *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics, Vol. 92, Amer. Math. Soc., Providence, 1997.

[3] Cucker, F., Smale, S. *Emergent behavior in flocks*, IEEE Trans. Automatic Control 52 (2007), 852–862.

[4] Hendrickx, J.M., Blondel, V.D. *Convergence of different linear and non-linear Vicsek models*, Proc. 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS2006), Kyoto (Japan), July 2006, 1229–1240.

[5] Jadbabaie, A., Lin, J., Morse, A.S. *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Trans. Automatic Control 48 (2003), 988–1001.

[6] Ji, M., Egerstedt, M. *Distributed coordination control of multi-agent systems while preserving connectedness*, IEEE Transactions on Robotics 23 (2007), 693–703.

[7] Landau, H.J., Odlyzko, A.M. *Bounds for eigenvalues of certain stochastic matrices*, Linear algebra and Its Applications 38 (1981), 5–15.

[8] Li, S., Wang, H. *Multi-agent coordination using nearest neighbor rules: revisiting the Vicsek model*, 2004, arXiv:cs/0407021v2

[9] Lorenz, J. *A stabilization theorem for dynamics of continuous opinions*, Physica A: Statistical Mechanics and its Applications 355 (2005), 217–223.

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

[11] Moshtagh, N., Jadbabaie, A., Daniilidis, K. *Distributed geodesic control laws for flocking of nonholonomic agents*, Proc. ECC-CDC 2005, Seville, Spain.

[12] Olfati-Saber, R. *Flocking for multi-agent dynamic systems: algorithms and theory*, IEEE Transactions on Automatic Control 51 (2006), 401-420.

[13] Olshevsky, A., Tsitsiklis, J.N. *Convergence speed in distributed consensus and averaging*, arXiv:math/0612682, 2006.

[14] Olshevsky, A., Tsitsiklis, J.N. *On the nonexistence of quadratic Lyapunov functions for consensus algorithms*, IEEE Transactions on Automatic Control, to appear, 2008.

[15] Reynolds, C.W. *Flocks, herds, and schools: A distributed behavioral model*, Computer Graphics 21 (1987), 25–34.

[16] Seneta, E. Non-Negative Matrices and Markov Chains, Springer, 2nd ed., 2006.

[17] Shi, H., Wang, L., Chu, T. *Coordination of multiple dynamic agents with asymmetric interactions*, Proc.

2005 IEEE International Symposium on Intelligent Control, Cyprus, June 27–29, 2005.

[18] Tahbaz-Salehi, A., Jadbabaie, A. *On recurrence of graph connectivity in Vicsek's model of motion coordination for mobile autonomous agents*, American Control Conference, ACC 2007.

[19] Tanner, H.G., Jadbabaie, A., Pappas, G.J. *Stable flocking of mobile agents, Part I: fixed topology*, Proc. IEEE Conference on Decision and Control, Maui, Hawaii (2003), 2010–2015.

[20] Tanner, H.G., Jadbabaie, A., Pappas, G.J. *Stable flocking of mobile agents, Part II: dynamic topology*, Proc. IEEE Conference on Decision and Control, Maui, Hawaii (2003), 2016–2021.

[21] Tanner, H.G., Jadbabaie, A., Pappas, G.J. *Flocking in fixed and switching networks*, IEEE Trans. Automatic Control 52, (2006), 8630-868.

[22] Vicsek, T., Czirok, A., Ben-Jacob, E., Cohen, I., Shochet, O. *Novel type of phase transition in a system of self-driven particles*, Physical Review Letters 75 (1995), 1226–1229.

[23] Yap C. *Fundamental Problems of Algorithmic Algebra*, Oxford University Press, Inc., New York, NY, 2000.