The Total s-Energy of a Multiagent System *

Bernard Chazelle †

Abstract

We introduce the *total s-energy* of a multiagent system and bound its maximum asymptotic value. This offers a new analytical lens on bidirectional agreement dynamics. We use our results to bound the convergence rates of dynamical systems for synchronization, flocking, opinion dynamics, and social epistemology.

1 Introduction

We introduce an analytical device for the study of multiagent agreement systems. The *total s-energy* of an infinite sequence of graphs $(G_t)_{t\geq 0}$ embedded in Euclidean d-space is a Dirichlet series that encodes all of the edge lengths. If $x_i(t)$ denotes the position in $\mathbb{R}^d$ of node i of G_t, then the *total s-energy* is defined as

$$E(s) = \sum_{t\geq 0} \sum_{(i,j)\in G_t} \|x_i(t) - x_j(t)\|_2^s \,. \tag{1}$$

The definition generalizes both the graph Laplacian and the Riesz s-energy of points on a sphere. The total s-energy may diverge everywhere. For bidirectional agreement systems, however, we show that it converges precisely for $s > 0$. We establish asymptotic bounds for any s between 0 and 1, which we then use to bound the convergence rates of classical systems in synchronization, flocking, opinion dynamics, and social epistemology. We also establish a new bound for products of certain stochastic matrices. Most of our proofs are algorithmic and bypass algebraic graph theory. This work shows the benefits of approaching multiagent dynamical systems algorithmically [5, 7].

1.1 Multiagent Dynamics

Moreau [29] introduced a geometric framework for multiagent agreement dynamics of appealing generality. He established qualitative convergence criteria but no quantitative results. The introduction of the total s-energy is an attempt to fill this gap.

---

*A preliminary version of this work appeared in *Proc. 26th Annual ACM Symposium on Computational Geometry*, 2010. This work was supported in part by NSF grant CCF-0634958 and NSF CCF-0832797.

†Department of Computer Science, Princeton University, chazelle@cs.princeton.edu

Bidirectional agreement systems. Our model is a quantitative version of Moreau's. It consists of n agents located at the points $x_1(t), \ldots, x_n(t)$ in $\mathbb{R}^d$ at any time $t \geq 0$, together with an infinite sequence $(G_t)_{t\geq 0}$ of undirected graphs over $n \geq 2$ nodes (corresponding to the agents) with $d = O(1)$; each node has a self-loop. These graphs represent the various configurations of a communication network changing over time. The neighbors of i form the set $N_i(t) = \{\, j \mid (i,j) \in G_t \,\}$, which includes i. At time t, each agent i moves anywhere within the convex hull

$$C_i(t) = \text{conv}\{\, x_j(t) \mid j \in N_i(t) \,\},$$

but not too close to the boundary. Formally, we pick a point $\omega_i(t)$ centrally located within $C_i(t)$: our choice is the *Löwner-John center* $\omega_i(t)$ of $C_i(t)$, ie,[1] the center of the minimum-volume ellipsoid $\mathcal{E}$ that encloses $C_i(t)$—minimum relative to the dimension of the affine hull of $C_i(t)$. It is well known that $\mathcal{E}$, and hence $\omega_i(t)$, are unique and that the ellipsoid derived from $\mathcal{E}$ by scaling it by a factor of $1/d$ about $\omega_i(t)$ is entirely contained inside $C_i(t)$: in other words, $C_i(t)$ is sandwiched between two concentric ellipsoids centered at $\omega_i(t)$ and differing by a factor of d [11]. Fix a parameter $0 < \rho \leq 1$. At time t, the next position of agent i is subject to the constraint:

$$x_i(t+1) \in (1-\rho)C_i(t) + \rho\, \omega_i(t). \tag{2}$$

Informally, this forces agent i to stay within a slightly shrunken version of its neighbors' convex hull (Figure 1). The freedom of an agent decreases as ρ increases; when $\rho = 1$, the system is deterministic. For this reason, we may always assume that ρ is smaller than a suitable constant. All the agents are updated in parallel at each step $t = 0, 1, 2$, etc. We show that this process always converges and we bound how long it takes (§1.3). The model defines a *bidirectional agreement system*. We conclude the presentation with a few general remarks.

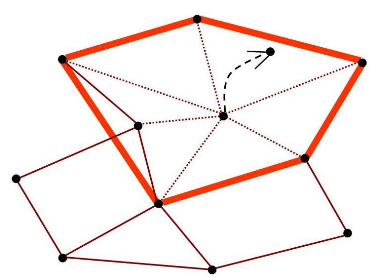

Figure 1: The agent can move anywhere inside the pentagon but may not touch the thick boundary.

- The model is affine-invariant (ie, the dynamics commutes with affine transformations). It is heavily nondeterministic: both the choice of communication graphs and the motion of the agents are left in the hands of an adversary. In applications, of course, the "adversary" is often endogenous.
- Choosing the mass center (or any other interior point) instead of the Löwner-John center is an option—in fact, it is our option for reversible systems—but in general it leads to bounds that are not quite as good. The choice of convex hulls is made partly for convenience. We could use smallest enclosing boxes, for example, albeit at the price of losing affine invariance; see [1, 24] for generalizations of Moreau's model to different shapes.

[1] We must emphasize that $C_i(t)$ and $\omega_i(t)$ are only analytical devices and that their computational complexity is irrelevant.

- The condition $\rho > 0$ is essential. Without it, a 2-agent system with a single edge could see the agents swap places forever without converging. Obviously, any convergence bound must grow inversely with ρ.

Much of the previous work on agreement systems has been concerned with conditions for consensus (ie, for all agents to come together), beginning with the pioneering work of [36, 37] and then [1, 3, 4, 15, 16, 20, 29, 30]. Bounds on the convergence rate have been obtained under various connectivity assumptions [4, 30] and for specialized closed-loop systems [6, 26]. The convergence of unrestricted bidirectional agreement systems can be derived from the techniques in [15, 23, 29]. Bounding the convergence rate, however, has been left open. This is the main focus of this paper.

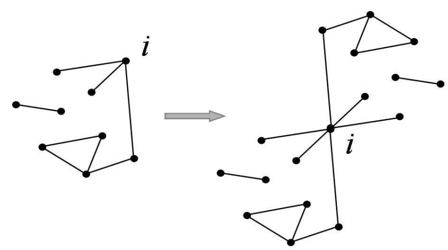

Figure 2: Reflecting the system about the agent i that we wish to fix.

The fixed-agent agreement model. We can fix one agent i if we so desire. By this, we mean skipping the update rule at an agent i selected once and for all. To see why, create the point reflection of the $n-1$ mobile agents about the fixed agent i to create a bidirectional system of $2n-1$ agents (Figure 2). Every mobile agent (and its reflected copy) mimics the behavior of its counterpart in the original n-agent system while respecting (2). The fixed agent now lies at the (Löwner-John) center of the centrally symmetric convex hull $C_i(t)$ and hence does not need to move, irrespective of the value of ρ. To summarize, any n-agent agreement system with one fixed agent can be simulated with a $(2n-1)$-agent bidirectional agreement system with the same value of ρ and at most twice the diameter. We apply this result to truth-seeking systems in §2.2.

Why not stochastic matrices? It is customary to model agreement systems by using products of (row-)stochastic matrices. The total s-energy suggests an alternative. Consider a Cartesian system of reference and let $y_i(t)$ denote the first coordinate of agent i at time t. Fix t, i and let y_l (resp. y_r) denotes the minimum (resp. maximum) of $y_j(t)$ over all $j \in N_i(t)$; note that both l and r depend on t, i. By definition of the Löwner-John center $\omega_i(t)$,

$$\Big(1 - \frac{\rho}{2d}\Big) y_l + \frac{\rho}{2d}\, y_r \leq y_i(t+1) \leq \frac{\rho}{2d}\, y_l + \Big(1 - \frac{\rho}{2d}\Big) y_r \,. \tag{3}$$

These conditions are necessary in general and also sufficient if $d = 1$. Let us now reverse our perspective and, instead of (2), assume that the dynamics of the agents is given by

$$x_i(t+1) = \sum_{j=1}^{n} p_{ij}(t) x_j(t), \tag{4}$$

where the matrix $P(t)$ with entries $p_{ij}(t)$ is stochastic and $x_i(t) \in \mathbb{R}^d$. When can we say that the coordinates $y_i(t)$ form an agreement system in one dimension? (Note that we do not attempt to interpret (4) as an agreement system in $\mathbb{R}^d$ but only, by projection, in $\mathbb{R}$.) First, we need a communication graph G_t: we define it by including the edge (i,j) if $p_{ij}(t) > 0$; we also add a self-loop at i, whether $p_{ii}(t)$ is positive or not. This, in turn, defines $N_i(t)$. The convex hull $C_i(t)$ is now the interval $[y_l, y_r]$; likewise, the Löwner-John center is the midpoint $\frac{1}{2}(y_l + y_r)$. Setting $d = 1$ in (3), the system $y_i(t)$ obeys the dynamics of a one-dimensional bidirectional agreement system with parameter ρ as long as the following two conditions are satisfied: for all t,

$$\begin{cases} \textit{Mutual confidence}: \text{ No pair } p_{ij}(t), p_{ji}(t) \text{ has exactly one zero;} \\ \textit{No extreme influence}: \text{ For all } i, \, \min\{\, p_{il}(t),\, p_{ir}(t)\,\} \geq \rho/2\,. \end{cases} \tag{5}$$

Condition (5) is weaker than the usual set of three constraints found in the literature [15, 23], which, besides mutual confidence, include: *self-confidence* (nonzero diagonal entries) and *nonvanishing confidence* (lower bound on all nonzero entries). Our model requires lower bounds on *only* two entries per matrix row. Previous work [5, 23] highlighted the importance of self-confidence for the convergence of agreement systems. Our results refine this picture, the moral of the story being: *To reach harmony in a group, individuals may be influenced extremely by non-extreme positions but must be influenced non-extremely by extreme positions (y_l or y_r).* In the case of a two-agent system, this maxim coincides with the need for self-confidence; in general, the latter is not needed.

Reversible agreement systems. We mention an important special case, which captures the notion of *reversibility* geometrically. Assume that each G_t is connected. Choose ρ no greater than the minimum degree ratio at each node, ie, $\rho \leq N_i(t)/N_i(t')$ for all i, t, t'. This allows each agent to pick a time-invariant "motion parameter" q_i between $|N_i(t)|$ and $\frac{1}{\rho}|N_i(t)|$, for all $t \geq 0$. Let m_i be the mass center of $C_i(t)$, ie,

$$m_i(t) = \frac{1}{|N_i(t)|} \sum_{j \in N_i(t)} x_j(t)\,.$$

The next position of agent i is given by

$$x_i(t+1) = x_i(t) + \frac{|N_i(t)|}{q_i}\Big(m_i(t) - x_i(t)\Big).$$

This fits within the previous model because agent i moves toward $m_i(t)$ in a straight line by a distance at least a fraction ρ of its distance to $m_i(t)$, now playing the role of $\omega_i(t)$. We call this model a *reversible agreement system.* The agents obey the dynamics (4), written in vector notation as

$$x_i(t+1) = \sum_{j \in N_i(t)} p_{ij}(t) x_j(t),$$

where

$$p_{ij}(t) = \begin{cases} 1 - (|N_i(t)| - 1)/q_i & \text{if } i = j; \\ 1/q_i & \text{if } i \neq j \in N_i(t); \\ 0 & \text{else.} \end{cases} \tag{6}$$

We take note of the identity $q_i p_{ij}(t) = q_j p_{ji}(t)$. This is the standard balanced condition of a reversible Markov chain, with (q_i) in the role of the stationary distribution (up to scaling). One can check that a lazy random walk in a connected graph can be modeled by a reversible agreement system. The difference is that we do not require the graph to be fixed. We bound the convergence time of reversible systems in §1.3.

1.2 The Total s-Energy

There is no obvious reason why the total s-energy, as defined in (1), should *ever* converge, so we treat it as a formal series for the time being. We prove that it converges for any $s > 0$ and we bound its maximum value, $E_n(s)$, over all moves and n-node graph sequences. We may assume unit initial diameter throughout, $D = \text{diam}\,\{x_1(0), \ldots, x_n(0)\} = 1$, since the total s-energy obeys the power-law $D^s E_n(s)$.

THEOREM 1.1. *The maximal total s-energy of an n-agent bidirectional agreement system with unit initial diameter satisfies*[2]

$$E_n(s) \leq \begin{cases} \rho^{-O(n)} & \textit{for } \; s = 1; \\ s^{1-n}\rho^{-n^2-O(1)} & \textit{for } \; 0 < s < 1. \end{cases}$$

There is a lower bound of $O(\rho)^{-\lfloor n/2 \rfloor}$ on $E_n(1)$ and of $s^{1-n}\rho^{-\Omega(n)}$ on $E_n(s)$ for n large enough, any $s \leq s_0$, and any fixed $s_0 < 1$.

Since $E_n(s) \leq E_n(1)$ for $s \geq 1$, the theorem proves the convergence of the total s-energy for all $s > 0$. When the graph remains connected at all times, it is sometimes more useful to define the total s-energy as the sum of the s-th powers of the diameters. It maximum value, for unit initial diameter, is denoted by

$$E_n^D(s) = \sum_{t \geq 0} \Big(\text{diam}\,\{x_1(t), \ldots, x_n(t)\} \Big)^s.$$

THEOREM 1.2. *The maximal diameter-based total s-energy of an n-agent reversible agreement system with unit initial diameter satisfies*

$$n^{-2} E_n(s) \leq E_n^D(s) \leq \frac{n}{s} \Big(\frac{2\delta n}{\rho} \Big)^{s/2+1},$$

for all $0 < s \leq 1$, where δ is the maximum degree of any node in the graph sequence.

We proceed with general remarks about the function $E(s)$. All of the terms in the series are nonnegative, so we can assume them rearranged in nonincreasing order. This allows us to express the total s-energy as a general Dirichlet series:

$$E(s) = \sum_{k \geq 1} n_k e^{-\lambda_k s}, \tag{7}$$

where $\lambda_k = -\ln d_k$ and n_k is the number of edges of length d_k. Thus, $E(s)$ is the Laplace transform of a sum of scaled Dirac delta functions centered at $x = \lambda_k$. This implies that the total s-energy can be inverted and, hence, provides a lossless encoding of the edge lengths. We show that $E(s)$ converges for any real $s > 0$. By the theory of Dirichlet series [12], it follows that $E(s)$ is uniformly convergent over any finite region $\mathcal{D}$ of the complex plane within $\Re(s) \geq r$, for any $r > 0$; furthermore, the series defines an analytic function over $\mathcal{D}$. It is not hard to determine the maximum s-energy of a 2-agent system with unit initial diameter. For $\rho = 1 - 1/e$ and $s = x + iy$, it satisfies (Figure 3):

$$|E(s)| = 1/\sqrt{1 - 2e^{-x}\cos y + e^{-2x}}\,.$$

[2]The constants implied in the asymptotic notation are absolute, so that, with our assumption that ρ is small enough, expressions such as $\rho^{O(n)}$, $O(\rho)^{-n}$, and $\rho^{-\Omega(n)}$ all denote lower bounds.

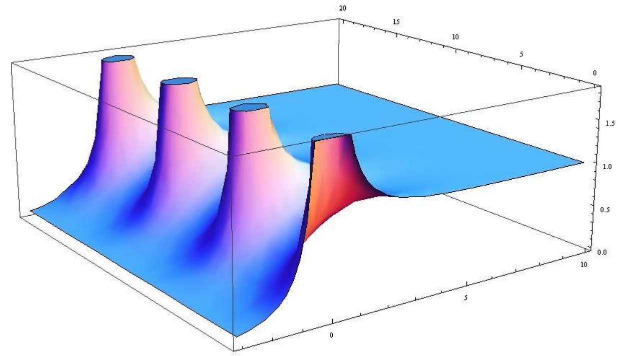

Figure 3: By analytic continuation, the maximum total s-energy of a two-agent system is a meromorphic function over the whole complex plane; the function is depicted in absolute value.

The singularities are the simple poles $s = 2\pi ik$, for all k. The total s-energy can be continued meromorphically over the whole complex plane. We conjecture that the same holds true for $E_n(s)$ in general.[3]

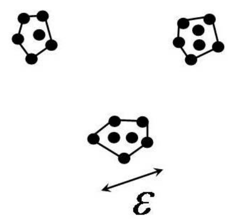

Figure 4: Convergence is reached when the agents fall within groups with disjoint convex hulls of diameter at most ε and no further interaction ever takes place between the groups.

1.3 Convergence

Given $0 < \varepsilon < 1/2$, we say that a step t is *trivial* (where ε is understood) if G_t has no edges or all of them are of length at most ε. The *communication count* C_ε is defined as the total number of nontrivial steps: roughly speaking, it is the number of nonmicroscopic moves. The system is said to *ε-converge* if the n agents can be partitioned at some time t into subsets with disjoint convex hulls, each one of diameter at most ε, so that no edge in $G_{t'}$ ever joins two of these subsets for any $t' \geq t$ (Figure 4). *Consensus* refers to the case of a one-set partition. Informally, convergence means that, visually, the system is forever frozen. Of course, an adversary can always delay convergence by throwing in the graph G_t consisting of n isolated nodes, so in our model's full generality, with notable exceptions, some of which we discuss below, the convergence time cannot be bounded.

[3] This is obviously false for *nonmaximal* s-energies: for example, the function $\sum_k e^{-sk!}$ is a valid total s-energy, but its singularities form a dense subset of its line of convergence (the imaginary axis), hence an impassable barrier for any analytic continuation into $\Re(s) < 0$.

Theorem 1.3. *For any $0 < \varepsilon < 1/2$, an n-agent bidirectional agreement system εn-converges by the time its last nontrivial step has elapsed. If the communication network remains connected at all times, then the system εn-converges to consensus within C_ε time.*

Theorem 1.4. *For any $0 < \varepsilon < 1/2$, the maximum communication count $C_\varepsilon(n)$ of any n-agent bidirectional agreement system with unit initial diameter satisfies*

$$O(\rho)^{-\lfloor n/2 \rfloor} \log \tfrac{1}{\varepsilon} \le C_\varepsilon(n) \le \min\Big\{ \tfrac{1}{\varepsilon}\, \rho^{-O(n)} \,,\, (\log \tfrac{1}{\varepsilon})^{n-1} \rho^{-n^2 - O(1)} \Big\}.$$

If the initial diameter D is not 1, then we must replace ε by ε/D in the bounds for $C_\varepsilon(n)$. We easily check that the bound is tight as long as ε is not unreasonably small (ie, $1/\varepsilon$ is not superexponential).

Corollary 1.5. *If $\varepsilon \ge \rho^{O(n)}$, then $C_\varepsilon(n) = \rho^{-\Theta(n)}$.*

Remark 1.6. In the upper bounds of the theorem and its corollary we can replace $C_\varepsilon(n)$ by "maximum number of nontrivial steps before the system ε-converges." Indeed, Theorem 1.3 shows that ε-convergence is reached after at most $C_{\varepsilon/n}$ steps have witnessed edges longer than ε/n; meanwhile, substituting ε/n for ε does not change the asymptotic bounds.

Theorem 1.7. *For any $0 < \varepsilon < \rho/n$, an n-agent reversible agreement system ε-converges to consensus in time $O(\frac{1}{\rho}\, \delta n^2 \log \frac{1}{\varepsilon})$, where δ is the maximum degree in the graph sequence.*

The theorem implies the polynomial mixing time of lazy random walks in connected graphs: we get the usual upper bound on the mixing time, with ρ controlling the laziness factor. The communication count is related to the total s-energy via the obvious inequality:

$$C_\varepsilon \le \varepsilon^{-s} E(s).$$

In view of this relation, the two upper bounds in Theorem 1.4 follow directly from those in Theorem 1.1: simply set $s = 1$ and $s = (n-1)/\ln \frac{1}{\varepsilon}$, respectively. Note that the second assignment can be assumed to satisfy $s < 1$, since it only concerns the case where $\frac{1}{\varepsilon}\rho^{-O(n)}$ is the dominant term in the right-hand side of the expression in Theorem 1.4. For reversible systems, we set $s = 1/\ln \frac{1}{\varepsilon}$, and observe that the number of steps witnessing a diameter in excess of ε is at most $\varepsilon^{-s} E_n^D(s) = O(\frac{1}{\rho}\, \delta n^2 \log \frac{1}{\varepsilon})$, hence Theorem 1.7. □

Communication and consensus. It has often been observed that if some agent communicates with all others infinitely often then the system evolves to consensus [29]. We provide a quantitative version of this fact. Agent i is said to communicate with agent j if there exists a sequence of times $t_0 < \cdots < t_k$ and agents $i_1, \ldots, i_k$ such that the edges $(i, i_1), (i_1, i_2), \ldots, (i_{k-1}, i_k), (i_k, j)$ belong to $G_{t_0}, G_{t_1}, \ldots, G_{t_k}$. Two agents i and j *communicate with each other* if i communicates with j or the other way around. Communication is not quite a symmetric relation but almost. Indeed, if i communicates with j at least $n-1$ times during nonoverlapping time intervals, then agent j must communicate with i at least once. To see why, observe that, until j communicates with i, whenever i communicates with j, the set of agents with whom j communicates increases by at least one compared with the previous communication. Indeed, because of self-loops, this set cannot decrease. Moreover, any further communication from i to j must penetrate that set, hence involve an edge with exactly one endpoint in it; this can happen at most $n-1$ times. If any two agents at distance greater than ε always communicate in the future, the system ε-converges to consensus. This follows directly from this result:

THEOREM 1.8. *Given any $0 < \varepsilon < 1/2$, consider two agents who always communicate at some point in the period following a separation by ε or more. The agents are then guaranteed to be at most ε apart when the system ε-converges. The total number of communications between them, counting only one after each such separation, is bounded by*

$$\min\Big\{\, \tfrac{1}{\varepsilon}\,\rho^{-O(n)}\,,\ (\log\tfrac{1}{\varepsilon})^{n-1}\rho^{-n^2-O(1)}\,\Big\}.$$

2 Applications

We highlight the utility of the total s-energy by looking at five examples: opinion dynamics; social epistemology; synchronization; flocking; and products of stochastic matrices.

2.1 Opinion Dynamics

The Krause opinion dynamics model [13, 18] is a sociological framework for tracking opinion polarization in a population. In its d-dimensional version, the *bounded-confidence model*, as it is often called, sets a parameter $0 < \delta < 1$ and, at time 0, specifies the opinions of n agents as n points in the unit cube $[0,1]^d$, for $d = O(1)$. At time $t \geq 0$, each opinion x moves to the position given by the average of every opinion in the Euclidean ball centered at x of radius δ. Viewed as a multiagent agreement system, G_t consists of n nodes (the agents) with edges joining any two of them within distance δ of each other. The time transition is specified by

$$x_i(t+1) = \frac{1}{|N_i(t)|}\sum_{j\in N_i(t)} x_j(t), \tag{8}$$

where $N_i(t)$ is the set of neighbors of node i in G_t, which includes i itself. The system is known to converge [2, 18, 22, 23]. Theorem 1.4 allows us to bound how long it takes to reach equilibrium. Consider a Cartesian coordinate system. In view of (5, 8), we may set $p_{ij}(t) = 1/|N_i(t)|$ and $\rho = 2/n$ to make the opinion dynamics system along each coordinate axis conform to a one-dimensional multiagent agreement model (2). We can assume that the diameter D along each axis is at most δn. Indeed, by convexity, if along any coordinate axis the n opinions have diameter greater than δn, then they can be split into two subsets with no mutual interaction now and forever. Set $\varepsilon = \delta/2$ and let t_ε be the smallest t such that G_t consists only of edges in $\mathbb{R}^d$ of length at most ε. During the first $dC_{\varepsilon/\sqrt{d}}(n)+1$ steps, it must be the case that, at some time t, the graph G_t contains only edges of length at most ε. By Theorem 1.4, therefore

$$t_\varepsilon \leq d^{3/2}\frac{D}{\varepsilon}\,n^{O(n)} = n^{O(n)}. \tag{9}$$

Each connected component of G_{t_ε} is a complete graph. To see why, observe that if opinion x is adjacent to y in G_{t_ε} and the same is true of y and z, then x and y are at a distance at most $2\varepsilon = \delta$, hence are connected and therefore at distance at most ε at time t_ε. This "transitive closure" argument proves our claim. This implies that the opinions within any connected component end up at the same position at time $t_\varepsilon + 1$. Of course, when two opinions are joined together they can never get separated. The argument is now easy to complete. Either G_{t_ε} consists entirely of isolated nodes, in which case the system is frozen in place, or it consists of complete subgraphs that collapse into single points. The number of distinct opinions decreases by at least one, so this process can be repeated at most $n-2$ times. By (9), this proves that Krause opinion dynamics converges in $n^{O(n)}$ time. We summarize our result.

THEOREM 2.1. *Any initial configuration of n opinions in the bounded-confidence Krause model with equal-weight averaging converges to a fixed configuration in $n^{O(n)}$ time.*

Martinez et al [26] have established a polynomial bound for the one-dimensional case, $d = 1$. While extending their proof to higher dimension might be difficult, a polynomial bound could well hold for any constant d.

2.2 Truth-Seeking Systems

In their pioneering work on computer-aided social epistemology, Hegselmann and Krause considered a variant of the bounded-confidence model that assumes a *cognitive division of labor* [14]. The idea is to take the previous model and fix one agent, *the truth*, while keeping the $n-1$ others mobile. A *truth seeker* is a mobile agent joined to the truth in every G_t. All the other mobile agents are *ignorant*, meaning that they never connect to the truth via an edge, although they might indirectly communicate with it via a path. Any two mobile agents are joined in G_t whenever their distance is less[4] than δ. Hegselmann and Krause [14] showed that, if all the mobile agents are truth seekers, they eventually reach consensus with the truth. Kurz and Rambau [19] proved that the presence of ignorant agents cannot prevent the truth seekers from converging toward the truth. The proof is quite technical and the authors leave open the higher-dimensional case. We generalize their results to any dimension and, as a bonus, bound the convergence rate.

THEOREM 2.2. *Any initial configuration of n opinions in the truth-seeking model converges, with all the truth seekers coalescing around the truth. If, in addition, we assume that the initial coordinates of each opinion as well as the radius δ are encoded as $O(n)$-bit rationals then, after $n^{O(n)}$ time, all the truth seekers lie within a ball of radius 2^{-cn^n} centered at the truth, for any arbitrarily large constant $c > 0$. Ignorant agents either lie in that ball or are frozen in place forever. This holds in any fixed dimension.*

Proof. Along each coordinate axis, a truth-seeking system falls within the fixed-agent agreement model and, as we saw in §1.1, can be simulated by a $(2n-1)$-agent one-dimensional bidirectional agreement system with at most twice the initial diameter. Convergence follows from Theorem 1.3. As we observed in the previous section, restricting ourselves to the equal-weight bounded confidence model allows us to set $\rho = 2/(2n-1)$. (We could easily handle arbitrary weights but this complicates the notation without adding anything of substance to the argument.) That all truth seekers reach consensus with the truth is a consequence of Theorem 1.8. Kurz and Rambau [19] observed that the convergence rate cannot be bounded as a function of n and ρ alone because it also depends on the initial conditions (hence the need to bound the encoding length of the initial coordinates).

Set $\varepsilon = 2^{-bn}$ for some large enough constant $b > 0$, and define t_ε as the smallest t such that G_t consists only of edges not longer than ε. The same argument we used in (9) shows that $t_\varepsilon = n^{O(n)}$. The subgraph of G_{t_ε} induced by the mobile agents consists of disjoint complete subgraphs. Indeed, the transitive closure argument of the previous section shows that the distance between any two

[4]We follow [2] in linking mobile agents less than δ apart: using open balls makes the proofs a little easier.

agents within the same connected component is at most $2\varepsilon = 2^{1-bn} < \delta$ (the inequality following from the $O(n)$-bit encoding of δ), hence at most ε. For similar reasons, because $4\varepsilon < \delta$, the truth agent cannot join more than one of these complete subgraphs; therefore, all the subgraphs consist of ignorant agents, except for one of them, which contains all the truth seekers and to which the truth agent is joined. This *truth group* might contain some ignorant agents as well, ie, mobile agents not connected to the truth. For that reason, the truth group might not form a complete subgraph. At time $t_\varepsilon + 1$, the truth group has collapsed into either a single edge with the truth at one end or a collinear 3-agent system consisting of the truth, a truth seeker, and an ignorant agent. (We refer to a single agent or truth seeker although it may be a collection of several of them collapsed into one.) All the other complete subgraphs collapse into *all-ignorant* single agents. By Theorem 2.1, there is a time

$$t_0 = t_\varepsilon + n^{O(n)} = n^{O(n)} \tag{10}$$

by which the all-ignorant agents will have converged into frozen positions unless they get to join with agents in the truth group at some point.

Case I. Assume that the all-ignorant agents do not join with any agent in the truth group at any time $t > t_\varepsilon$: The truth group then behaves like a one-dimensional fixed-agent system with 2 or 3 agents embedded in $\mathbb{R}^d$. We assume the latter, the other case being similar, only easier. We saw in §1.1 how such a system can be simulated by a one-dimensional 5-agent bidirectional system of at most twice the diameter. Recall that agents may represent the collapse of several of them, so we must keep the setting $\rho = 2/(2n-1)$. The 5-agent system remains connected at all times; therefore, by Theorems 1.3 and 1.4, it β-converges to consensus in $t_0 + n^{O(1)}(\log \frac{1}{\beta})^4$ time. By (10), this implies that, for any fixed $c > 0$, the agents of the truth group are within distance 2^{-cn^n} of the truth after $n^{O(n)}$ time.

Case II. Assume now that an all-ignorant agent z joins with an agent y of the truth group at time t_1 but not earlier in $[t_\varepsilon, t_1)$. That means that the distance $\|y(t_1)z(t_1)\|_2$ dips below δ for the first time after t_ε. We want to show that $t_1 \leq t_0 + n^{O(n)}$, so we might as well assume that $t_1 > t_0$. Recall that t_0 is an upper bound on the time by which the all-ignorant agents would converge if they never interacted again with the truth group past t_ε. Let L be the line along which the truth group evolves and let σ be its (nonempty) intersection with the open ball of radius δ centered at $z(t_1) = z(t_0)$. Note that σ cannot be reduced to a single point. This implies that the shortest nonzero distance Δ between the truth and the endpoints of σ is well-defined. We claim that $\Delta \geq 2^{-n^{O(n)}}$. Here is why. It is elementary to express Δ as a feasible value of a variable in a system of m linear and quadratic polynomials over m variables, where m is a relatively small constant (depending on d); the details are unnecessary. The coefficients of the polynomials can be chosen to be integers over $\ell = n^{O(n)}$ bits. (We postpone the explanation.) We need a standard root separation bound [41]. Given a system of m integer-coefficient polynomials in m variables with a finite set of complex solution points, any nonzero coordinate has modulus at least $2^{-\ell\gamma^{O(m)}}$, where $\gamma - 1$ is the maximum degree of any polynomial and ℓ is the number of bits needed to represent any coefficient. This implies our claimed lower bound of $2^{-n^{O(n)}}$ on Δ.

Why is $\ell = n^{O(n)}$? At any given time, consider the rationals describing the positions of the n agents and put them in a form with one common denominator. At time 0, each of the initial positions now requires $O(n^2)$ bits (instead of just $O(n)$ bits). A single time step produces new rationals whose common denominator is at most $n!$ times the previous one, while the numerators are sums of at most n previous numerators, each one multiplied by an integer at most $n!$. This means that, at time t, none of the numerators and denominators require more than $O(n^2 + tn\log n)$ bits. The system of equations expressing Δ can be formulated using integer coefficients with $O(n^2 + t_0 n \log n)$ bits, hence the bound of $\ell = n^{O(n)}$. Next, we distinguish between two cases.

- The truth is not an endpoint of σ: Then there is a closed segment of L centered at the truth that lies either entirely outside σ or inside it. Because $\Delta \geq 2^{-n^{O(n)}}$, the segment can be chosen of length at least $2^{-n^{O(n)}}$. Setting $\beta = 2^{-n^{c_0 n}}$, for c_0 large enough, shows that $t_1 \leq t_0 + n^{O(1)}(\log \frac{1}{\beta})^4 = n^{O(n)}$.
- The truth is an endpoint of σ: Quite clearly, β-convergence alone does not suffice to bound t_1; so we reason as follows. When the truth group has β-converged (for the previous value of β), the only way its mobile agents avoided falling within σ (in which case the previous bound on t_1 would hold) is if the truth group ended up separated from σ by the truth (lest one of the mobile agents lay in σ). By convexity, however, this property remains true from then on, and so z can never join y, which contradicts our assumption.

When agents y and z join in G_t at time $t = t_1$, their common edge is of length at least $\delta/3$ unless y or z has traveled a distance at least $\delta/3$ between t_ε and t_1. In all cases, the system must expend 1-energy at least $\delta/3$ during that time interval. By Theorem 1.1, this can happen at most $n^{O(n)}(3/\delta) = n^{O(n)}$ times. (We easily verify that the same argument can be repeated safely even though the bit lengths will increase.) At the completion of this process (should it happen), we are back to Case I. □

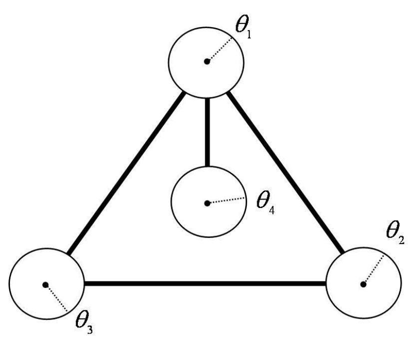

Figure 5: Four coupled oscillators connected by four edges.

2.3 Kuramoto Synchronization

The *Kuramoto model* is a general framework for nonlinear coupled oscillators, with a dazzling array of applications: circadian neurons, chirping crickets, microwave oscillators, yeast cell suspensions, pacemaker cells in the heart, etc. Winfree's pioneering work on the subject led Kuramoto to formulate the standard sync model for coupled oscillators [34, 40]. The system consists of n oscillators: the i-th one has phase θ_i and natural frequency ω_i. In its original formulation, the model is a mean-field approximation that assumes all-pair coupling. A more realistic assumption is to use a time-varying network to model communications. Considerable work has been done on this problem; see [10, 17, 21, 25, 27, 31, 32, 39, 42] for a small sample. Further research introduced a time-1 discretization of the continuous model [25, 29, 33, 35]. Assuming all oscillators share the same natural frequency, a fixed phase shift yields the dynamics:

$$\theta_i(t+1) = \theta_i(t) + \frac{K\Delta T}{|N_i(t)|} \sum_{j \in N_i(t)} \sin(\theta_j(t) - \theta_i(t)),$$

where $|N_i(t)|$ is the degree of i in the communication graph G_t, which, as always, counts the self-loop at i (Figure 5). As in [29], we also assume that all the agents start in the same open half-circle. By shifting the origin, we express this condition as $\alpha - \pi/2 \leq \theta_i(0) \leq \pi/2$, for some arbitrarily small positive constant α. This implies that

$$\sin(\theta_j(0) - \theta_i(0)) = a_{ij}(\theta_j(0) - \theta_i(0)),$$

for $\alpha/4 \leq a_{ij} \leq 1$. By (5); therefore, to make the dynamics conform to a bidirectional multiagent agreement system at time 0, it suffices to enforce the constraints:

$$\frac{2n\rho}{\alpha} \leq K\Delta T \leq 1 - \frac{\rho}{2}\,.$$

Choosing $\rho = b\alpha/n$ for a small enough constant $b > 0$, we note that the constraints are roughly equivalent to $0 < K\Delta T < 1$. By convexity, the angles at time 1 remain within $[\alpha - \pi/2, \pi/2]$; therefore, our previous argument can be repeated to show that the synchronization dynamics fits within the bidirectional agreement model. The result below follows from Corollary 1.5 and Remark 1.6.

Theorem 2.3. *Any Kuramoto synchronization system with n agents initialized in an open half-circle ε-converges after $n^{O(n)}$ nontrivial steps, for any $\varepsilon > n^{-cn}$ and any constant $c > 0$.*

2.4 Flocking

Beginning with Reynolds's pioneering work in the mid-eighties, bird flocking has generated an abundant literature, with a sudden flurry of interest in the last few years. Mathematically, flocking appears more complex than the previous agreement systems because the communication network is specified not by the variables the system seeks to "agree upon" but by their integrals. Specifically, the graphs G_t are a function of the birds' *positions* while the agreement dynamics averages their *velocities*. Many models have been studied in the literature but most of them are variants of the following [5, 9, 16, 38]: given the initial conditions $z(0)$ and $z(1)$, for any $t > 0$,

$$\begin{cases} z(t) = z(t-1) + v(t); \\ v(t+1) = (P(t) \otimes I_3)v(t). \end{cases} \tag{11}$$

The vectors $z(t), v(t)$ encode the positions and velocities of the n birds in 3-space: each vector is formed by stacking the n relevant triplets into one vector in $\mathbb{R}^{3n}$; $P(t)$ is an n-by-n stochastic matrix whose nonzero entries correspond to the edges of G_t; the communication graph G_t links any two birds within a fixed distance of each other. The tensor product $P(t) \otimes I_3$ is the $3n$-by-$3n$ matrix formed by replacing each entry of $P(t)$ by a 3-by-3 diagonal matrix with copies of that entry along the diagonal; it is a common notational device to indicate that the averaging should be carried out separately along each coordinate axis. Intuitively, each bird averages out its own velocity with those of its neighbors in G_t: all of its neighbors weigh equally in the average except perhaps for itself, ie, for fixed i, all nonzero $p_{ij}(t)$'s are equal, with the possible exception of $p_{ii}(t)$; all nonzero entries in $P(t)$ are assumed to be at least $n^{-O(1)}$, as are the rational initial conditions; finally, all the diagonal entries are strictly positive.

By (5), it suffices to set $\rho = n^{-b}$, for a large enough constant $b > 0$, to make flocking conform to the bidirectional multiagent agreement model, with $v(t)$ encoding into a single vector the n points $(x_1(t), \ldots, x_n(t))$. By Corollary 1.5 and Remark 1.6, the system ε-converges within $n^{O(n)}$ nontrivial steps for $\varepsilon \geq n^{-cn}$ and any constant $c > 0$. We showed in [6] that the sequence G_t always converges to a fixed graph G, but that the number of steps to get there can be astronomical: it can be as high as a tower-of-twos of height on the order of $\log n$, which, amazingly, is tight.

THEOREM 2.4. *The velocities of n birds ε-converge after $n^{O(n)}$ nontrivial steps, for any $\varepsilon > n^{-cn}$ and any constant $c > 0$.*

2.5 Products of Stochastic Matrices

Let $\mathcal{P}$ be the family of n-by-n stochastic matrices such that each $P \in \mathcal{P}$ satisfies the three standard constraints: (i) *self-confidence* (nonzero diagonal entries); (ii) *mutual confidence* (no pair p_{ij}, p_{ji} with exactly one 0); and (iii) *nonvanishing confidence* (each positive entry being at least p). Lorenz [23] and Hendrickx and Blondel [15] independently proved the following counterintuitive result: in any finite product of matrices in $\mathcal{P}$, each nonzero entry is at least $p^{O(n^2)}$. What is surprising is that this lower bound is uniform, ie, does not depend on the number of multiplicands in the product. Not only does stochasticity play a key role, but so do conditions (i–iii). We improve the lower bound to its optimal value.

THEOREM 2.5. *Let $\mathcal{P}$ be the family of n-by-n real stochastic matrices such that each $P \in \mathcal{P}$ satisfies: each diagonal entry is nonzero; each positive entry is at least p; and no pair p_{ij}, p_{ji} contains exactly one zero. In any finite product of matrices in $\mathcal{P}$, each nonzero entry is at least p^{n-1}. The bound is optimal.*

3 The Proofs

It remains for us to prove the upper bounds of Theorem 1.1 (see §3.2 for the case $s = 1$ and §3.3 for the case $s < 1$), Theorem 1.2 (see §3.1), Theorem 2.5 (see §3.3), Theorems 1.3 and 1.8 (see §3.5), as well as the lower bounds of Theorems 1.1 and 1.4 (see §3.4). We begin with Theorem 1.2.

3.1 The Reversible Case

Let $\pi_i = q_i / \sum_j q_j$. We easily verify that $\pi = (\pi_1, \ldots, \pi_n)$ is the (time-invariant) stationary distribution of the stochastic matrix $P = P(t)$ specified by (6):

$$p_{ij} = \begin{cases} 1 - (|N_i| - 1)/q_i & \text{if } i = j; \\ 1/q_i & \text{if } i \neq j \in N_i; \\ 0 & \text{else.} \end{cases}$$

The discussion focuses on a fixed step so we drop the argument t from the notation for simplicity. Let $y = (y_1, \ldots, y_n)$, where y_i is the first coordinate of $x_i(t)$, and, for $u, v \in \mathbb{R}^n$, let $\langle u, v \rangle_\pi = \sum \pi_i u_i v_i$. We can always choose the origin of our Cartesian coordinate system so that $\langle y, \mathbf{1} \rangle_\pi = 0$. Because π is the stationary distribution, this property is time-invariant; in particular, $\langle Py, \mathbf{1} \rangle_\pi = 0$. The following derivations on the Dirichlet form $\langle y, (I - P)y \rangle_\pi$ are standard [8, 28]. Because P is reversible, we can decompose $y = a_i v_i$ in an eigenbasis $\{v_i\}$ for P orthonormal with respect to $\langle \cdot \rangle_\pi$. Any positive p_{ij} is at least $1/q_i \geq \rho/\delta$, where $\delta > 1$ is the maximum degree in any G_t (including self-loops). Let $1 = \lambda_0 > \lambda_1 \geq \cdots \geq \lambda_{n-1} \geq 2\rho/\delta - 1$ be the eigenvalues of P, with the labeling matching the v_i's: briefly, the gap is strict between the two largest eigenvalues because the graph is connected; the smallest eigenvalue is separated from -1 by at least $2\rho/\delta$ because $(\delta P - \rho I)/(\delta - \rho)$ is itself a reversible Markov chain (with the same eigenvectors), hence with spectrum in $[-1, 1]$. If $\mu = \max\{\lambda_1^2, \lambda_{n-1}^2\}$

then, by reversibility, $\pi_i p_{ij} = \pi_j p_{ji}$, and

$$\begin{aligned}\langle y, y\rangle_\pi - \langle Py, Py\rangle_\pi = \langle y, (I-P^2)y\rangle_\pi = \sum_{i,j} a_i a_j \langle v_i, (I-P^2)v_j\rangle_\pi = \sum_i a_i^2(1-\lambda_i^2) \\ \geq (1-\mu)\sum_i a_i^2 \geq (1-\mu)\sum_{i,j} a_i a_j \langle v_i, v_j\rangle_\pi = (1-\mu)\langle y, y\rangle_\pi.\end{aligned} \tag{12}$$

Because P is reversible and $\pi_i p_{ij} \geq \rho/\delta n$, for any vector z,

$$\langle z, (I-P)z\rangle_\pi = \frac{1}{2}\sum_{i,j} \pi_i p_{ij}(z_i - z_j)^2 \geq \frac{\rho}{\delta n}\sum_{(i,j)\in G_t}(z_i - z_j)^2.$$

Set $z = v_1$; by orthonormality, $\langle z, \mathbf{1}\rangle_\pi = 0$ and $\langle z, z\rangle_\pi = 1$; therefore z must contain a coordinate z_a such that $|z_a| \geq 1$ and another one, z_b, of opposite sign. If L is a simple path in G_t connecting nodes a and b, then, by Cauchy-Schwarz,

$$\begin{aligned}1 - \lambda_1 = \langle z, (I-P)z\rangle_\pi \geq \frac{\rho}{\delta n}\sum_{(i,j)\in L}(z_i - z_j)^2 \geq \frac{\rho}{\delta n^2}\Big(\sum_{(i,j)\in L}|z_i - z_j|\Big)^2 \\ \geq \frac{\rho}{\delta n^2}(z_a - z_b)^2 \geq \frac{\rho}{\delta n^2}\,.\end{aligned}$$

Since $\rho n^{-2}/\delta \leq 2\rho/\delta \leq 1 + \lambda_{n-1}$, it follows that

$$\mu \leq \Big(1 - \frac{\rho}{\delta n^2}\Big)^2 \leq 1 - \frac{\rho}{\delta n^2},$$

and, by (12),

$$\langle Py, Py\rangle_\pi \leq \mu\langle y, y\rangle_\pi \leq \Big(1 - \frac{\rho}{\delta n^2}\Big)\langle y, y\rangle_\pi.$$

By analogy, for $x = (x_1(t), \ldots, x_n(t))$, we define $\langle x, x\rangle_\pi = \sum_i \pi_i \|x_i(t)\|_2^2$ and again we may assume that $\sum_i \pi_i x_i(t)$ lies at the origin in $\mathbb{R}^d$. It then follows that

$$\langle Px, Px\rangle_\pi \leq \Big(1 - \frac{\rho}{\delta n^2}\Big)\langle x, x\rangle_\pi.$$

Let $E_n^D(L, s)$ be the maximum value of the (diameter-based) total s-energy of an n-agent reversible agreement system such that $\langle x, x\rangle_\pi = L$ at time 0. By the triangle inequality, the distance between any two agents is at most

$$2\max\|x_i\|_2 \leq 2\sqrt{L/\min \pi_i} \leq \sqrt{2L\delta n/\rho}\,;$$

therefore,

$$E_n^D(L, s) \leq E_n^D((1 - \rho n^{-2}/\delta)L, s) + (2L\delta n/\rho)^{s/2}.$$

The total s-energy obeys the scaling law $E_n^D(\alpha L, s) = \alpha^{s/2}E_n^D(L, s)$. The definition of $E_n^D(s)$ assumes unit initial diameter, which implies that $\langle x, x\rangle_\pi \leq 1$, hence $E_n^D(s) \leq E_n^D(1, s)$ and

$$E_n^D(s) \leq \frac{(2\delta n/\rho)^{s/2}}{1 - (1 - \rho n^{-2}/\delta)^{s/2}} \leq \frac{n}{s}\Big(\frac{2\delta n}{\rho}\Big)^{s/2+1},$$

which proves Theorem 1.2. We used an inequality that is worth mentioning for later use: for any $0 \leq x, y \leq 1$,

$$(1-x)^y \leq 1 - xy. \tag{13}$$

□

3.2 The General Case: $s = 1$

We prove the upper bound of Theorem 1.1 for $s = 1$. We show that $E_n(1) \leq \rho^{-O(n)}$, with the constant in the exponent being roughly 2. To do that, we introduce a *wingshift system.* Since we will focus on a single transition at a time, we write a_i, b_i instead of $x_i(t), x_i(t+1)$ for notational convenience, and we relabel the agents so that $0 \leq a_1 \leq \cdots \leq a_n \leq 1$. Given $a_1, \ldots, a_n$, the agents move to their next positions $b_1, \ldots, b_n$ adversarially and then repeat this process endlessly in the manner described below. Let $\ell(i)$ and $r(i)$ be indices satisfying the following inequalities:

RULE 1: $1 \leq \ell(i) \leq i \leq r(i) \leq n$ and $(\ell \circ r)(i) \leq i \leq (r \circ \ell)(i)$;

RULE 2: $a_{\ell(i)} + \delta_i \leq b_i \leq a_{r(i)} - \delta_i$, where $\delta_i = \rho(a_{r(i)} - a_{\ell(i)})$.

Each agent i picks an *associate* to its left and one to its right, $\ell(i)$ and $r(i)$, respectively. It then shifts anywhere in the interval $[a_{\ell(i)}, a_{r(i)}]$, though keeping away from the endpoints by a small distance δ_i. This process is repeated forever, with each agent given a chance to change associates at every step. Wingshift systems can be used to simulate projections of agreement systems. Indeed, if we project a d-dimensional bidirectional agreement system, with parameter ρ', onto a line then, by (3), once we identify y_l with $a_{\ell(i)}$ and y_r with $a_{r(i)}$, we get a wingshift system satisfying

$$a_{\ell(i)} + \frac{\rho'}{2d}(a_{r(i)} - a_{\ell(i)}) \leq b_i \leq a_{r(i)} - \frac{\rho'}{2d}(a_{r(i)} - a_{\ell(i)}),$$

hence a wingshift system with $\rho = \rho'/(2d)$. In other words, projecting a bidirectional agreement system onto a line gives us a wingshift system for a value of ρ equal to a fraction $1/(2d)$ of that parameter's value in the agreement system. Rule 1 says that the interval $[\ell(i), r(i)]$ should contain i as well as all agents j pointing to i. This is also necessary. Consider three agents $a_1 = 0$, $a_2 = \frac{1}{2}$, $a_3 = 1$, with $\ell(1) = r(1) = \ell(2) = 1$ and $r(2) = \ell(3) = r(3) = 3$. Agents 1 and 3 are stuck in place while agent 2 can move about almost anywhere: convergence cannot be assured. Intuitively, iterating the functions ℓ and r in alternation should produce an outward-growing spiral ending in a cycle. Specifically, if we follow the path $i, r(i), l(r(i)), r(l(r(i)))$, etc, we spiral around i along edges of nondecreasing length until we eventually fall into a two-edge cycle of the form $l(r(j)) = j$. The *wingshift graph* consists of the nodes $1, \ldots, n$ and the directed edges $(i, \ell(i))$ and $(i, r(i))$.

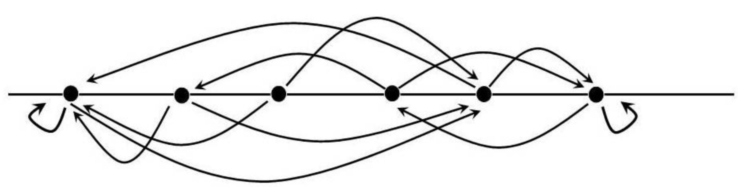

Figure 6: A six-node wingshift graph.

By analogy, we define the total 1-energy of the wingshift system as $V = \sum_{t \geq 0} V_t$, where

$$V_t = \sum_{i=1}^{n} (a_{r(i)} - a_{\ell(i)}).$$

THEOREM 3.1. *The maximal total* 1*-energy of an* n*-agent wingshift system with unit initial diameter and parameter* ρ *is at most* $\rho^{-O(n)}$.

Summing the total 1-energy of the wingshift system formed along each coordinate axis by a bidirectional agreement system gives us an upper bound of $\rho^{-O(n)}$ on the total 1-energy of the latter, which establishes the upper bound of Theorem 1.1 for $s = 1$ and arbitrary d. We omitted a number of scaling factors of no asymptotic significance: (i) a factor of $2d$ for ρ to account for the switch from agreement to wingshift; (ii) a factor of d to account for all d projections; and (iii) a factor of n to account for the fact that the energy is computed over all the edges in an agreement system but only over the two edges from each node in a wingshift system. □

Recall that $\rho > 0$ is smaller than a suitable constant, a fact we shall assume throughout. We need some notation to describe rightward paths in the wingshift graph: $r(i,0) = i$ and $r(i,k) = r(r(i,k-1))$ for $k > 0$. We define the distance between an agent and its right associate, $\Delta_i = a_{r(i)} - a_i$. When traversing a rightward path $i = r(i,0), r(i,1), \ldots, r(i,k)$ etc, a sudden drop in $\Delta_{r(i,k)}$ is of particular interest; so we define

$$k_i = \min\Big\{\, k \geq 1 \,|\, \Delta_{r(i,k)} \leq \tfrac{\rho}{2}\, \Delta_{r(i,k-1)} \,\Big\}.$$

Since, for k large enough, $r(i,k) = r(i,k+1)$, hence $\Delta_{r(i,k)} = 0$, a drop by a factor of at least $\rho/2$ is certain to occur and thus k_i is always well-defined (even if $r(i) = i$, in which case $k_i = 1$). We now show that agent $r(i,k_i)$, which we call a *stopper*, must move left by an amount proportional to Δ_i. What makes this result interesting is that it points to some obligatory motion that the system must undergo; in other words, staying in place is not an option, unless the system consists entirely of self-loops.

Lemma 3.2. *For any agent i, its stopper $s = r(i,k_i)$ satisfies $a_s - b_s \geq (1+\rho)(\frac{\rho}{2})^{k_i}\Delta_i$.*

Proof. Note that the index s has nothing to do with the s-energy—no risk of confusion since we focus here on the 1-energy. Recall that b_s is the position at time $t+1$ of the agent associated with a_s at time t; unlike the latter, b_s may not be the s-th agent from the left. By Rule 1,

$$a_{r(r(i,k_i))} - a_{\ell(r(i,k_i))} \geq \Delta_{r(i,k_i-1)} + \Delta_{r(i,k_i)}\,.$$

By definition, $\Delta_{r(i,k_i)} \leq \frac{\rho}{2}\,\Delta_{r(i,k_i-1)}$; therefore, by Rule 2,

$$\begin{aligned} b_{r(i,k_i)} &\leq a_{r(r(i,k_i))} - \rho\big(a_{r(r(i,k_i))} - a_{\ell(r(i,k_i))}\big) \leq a_{r(i,k_i)} + \Delta_{r(i,k_i)} - \rho\big(\Delta_{r(i,k_i-1)} + \Delta_{r(i,k_i)}\big) \\ &\leq a_{r(i,k_i)} + ((1-\rho)\rho/2 - \rho)\Delta_{r(i,k_i-1)} \leq a_{r(i,k_i)} - (1+\rho)(\rho/2)^{k_i}\Delta_i\,. \end{aligned}$$

The last inequality follows from the fact that $\Delta_{r(i,k_i-1)} \geq (\rho/2)^{k_i-1}\Delta_i$. The reason the inequality is not strict is that both sides are equal if $k_i = 1$. □

We bound V by tallying the *total motion* of the system, $M = \sum_{t\geq 0} M_t$, where $M_t = \sum_{i=1}^n |a_i - b_i|$. Recall that a_i and b_i are the positions of agent i at times t and $t+1$, respectively.

Lemma 3.3. *If M is finite, then $V \leq n2^n\rho^{1-n}M$.*

Proof. Obviously,

$$V = \sum_{t\geq 0}\sum_{i=1}^n (a_{r(i)} - a_{\ell(i)}) \leq \sum_{t,i}\Big\{(a_i - a_{\ell(i)}) + (a_{r(i)} - a_i)\Big\} \leq \sum_{t,i}\Delta_i + \sum_{t,i}\Delta'_i\,,$$

where $\Delta_i' = a_i - a_{\ell(i)}$. By Lemma 3.2,

$$\sum_{t\geq 0}\sum_{i=1}^{n}\Delta_i \leq \sum_{t\geq 0}\sum_{i=1}^{n}\Big(\frac{2}{\rho}\Big)^{k_i}(a_{r(i,k_i)} - b_{r(i,k_i)}) \leq n\Big(\frac{2}{\rho}\Big)^{n-1}\sum_{t\geq 0}\sum_{i=1}^{n}|a_i - b_i|\,.$$

A mirror-image argument yields the same upper bound on $\sum_{t,i}\Delta_i'$. □

By symmetry, we can assume that at least half of the contribution to M is provided by rightward motions, ie, $\frac{1}{2}M \leq \sum_{t,i}\{\, b_i - a_i \,|\, b_i > a_i \,\}$. Thus we can conveniently ignore all leftward travel for accounting purposes. We use an amortization technique that involves assigning a *credit account* to each agent and requiring them to spend an amount equal to the distance they travel to the right. We also have a *bank* where *credits* can be deposited, withdrawn, or borrowed. In the end, we will ensure that bank loans do not exceed deposits, so that the initial credit accounts provide an upper bound on $\frac{1}{2}M$. The power of the method is that agents can trade credits among one another. This creates an "economy" of credits, regulated by rules, which we embed within an *algorithmic proof.*

Let $rk(a_i)$ be the rank of agent i. Initially, the agents are labeled from left to right, $a_1 \leq \cdots \leq a_n$, so that $rk(a_i) = i$; ties are broken arbitrarily. When agent i is moved from a_i to b_i, the ranks are updated in the obvious way; ties are broken by minimizing the number of agents whose ranks are changed. So, for example, if $a_1 = a_2 = 0$ and $a_3 = a_4 = \frac{1}{2}$, then moving a_1 to $b_1 = \frac{1}{2}$ changes the ranking $(1, 2, 3, 4)$ into $(2, 1, 3, 4)$; or, say, moving a_4 to $b_4 = 0$ produces the ranking $(1, 2, 4, 3)$. We call this *lazy tie-breaking.* We denote by $a_i(t)$ the position of agent i at time t. For convenience, we focus on the time interval $[0, 1]$ and use a continuous time scale, with $a_i(0) = a_i$ and $a_i(1) = b_i$. The notation $rk(a_i(t))$ refers to the time-t rank of agent i within the set of all agents at time $0 \leq t \leq 1$. We schedule the motion of each agent i from a_i to b_i one at a time in a specific fashion described below.

> Invariant: The credit account of each agent i is supplied at time t with at least $a_i(t)\alpha^{rk(a_i(t))}$ credits, where $\alpha = 9/\rho$.

The cost of a right shift $a_i \rightarrow b_i$ ($a_i < b_i$) is $b_i - a_i$ plus whatever is needed to update the agents' accounts. Note that moving agent i to the right affects not only its own account but also those of the agents it passes over, whose ranks decrease by one, and who may therefore release credits and deposit them into the bank. The cost attached to shifting agent i is paid for by borrowing from the bank, a loan that will be paid back later by the stopper $r(i, k_i)$, ie, *not* by the borrower. To explain how this works, we partition the set of all agents into equivalence classes consisting of all the agents i with the same stopper $s = r(i, k_i)$. We define the set P_s of the immediate predecessors of s, in the wingshift graph, that claim s as their stopper; the set H_j^s consists of all the nodes that lead rightward to $j \in P_s$: in other words, for any stopper s,

$$\begin{cases} P_s &= \{\, j \neq s \,|\, r(j) = s \text{ and } k_j = 1 \,\}; \\ H_j^s &= \{\, i \neq j \,|\, r(i, k_i - 1) = j \text{ and } j \in P_s \,\}. \end{cases}$$

Note that some P_s or H_j^s might be empty. First, for each stopper s from right to left, we move all the right-shifting agents in each H_j^s and pass on the cost to agent j. Second, we move each such agent j to b_j, charging the bank for the cost. Finally, we move the stopper s to the left and use the credits released to pay back the bank. All remaining left moves, which are self-paying, are then carried out. Why don't we account for all the nodes claiming s as a stopper uniformly? We distinguish between P_s and $\{H_j^s\}$ because their relations to the stopper are fundamentally different: the defining feature of P_s is to witness a big drop in the length Δ_k, whereas the opposite is true for H_j^s. The accounting mechanism will vary accordingly, with H_j^s charging its cost to P_s, and then P_s passing on its costs to s via a bank loan.

- For each stopper s in decreasing rank order, do:
  - **[1]** For each agent $j \in P_s$ ordered by nonincreasing value of $\max\{a_j, b_j\}$, do:
    - **[1.1]** Move from a_i to b_i any agent $i \in H_j^s$ such that $a_i < b_i$. Charge agent j for all the costs.
    - **[1.2]** If $a_j < b_j$ then move agent j from a_j to b_j. In both cases, borrow from bank to cover all charges incurred by agent j.
  - **[2]** Move agent s left from a_s to b_s. Use released credits to clear all outstanding bank loans.
- Move from a_i to b_i any agent i such that $b_i < a_i$.

Step [1] Fix s. We process each agent $j \in P_s$ in the order $j_1, \ldots, j_\nu$ such that $c_{j_1} \geq \cdots \geq c_{j_\nu}$, where $c_k = \max\{a_k, b_k\}$. We resolve ties in decreasing rank order of a_k; in other words, if $c_{j_k} = c_{j_l}$ and $rk(a_{j_k}(t)) > rk(a_{j_l}(t))$, then we process j_k before j_l (where t is the time right before step [1] for the current stopper s). Recall that lazy tie-breaking ensures that each agent has a distinct rank at any time. Fix $j = j_l$. The execution of step [1.1] occurs in the period denoted by $[t_{l,1}, t_{l,2})$, while that of step [1.2] takes place during $[t_{l,2}, t_{l+1,1})$. No agent moves left in step [1]. Agents of H_j^s may right-shift during $[t_{l,1}, t_{l,2})$ but agent j itself remains in place in that time interval. The sets P_s and H_j^s are disjoint over all j, s, so no agent in H_j^s has moved prior to $t_{l,1}$. But couldn't a stopper in H_j^s (indeed, H_j^s may include stoppers) have moved left in earlier executions of step [2]? No. The reason is that we process stoppers from right to left.

Step [1.1] Consider the motion of an agent $i \in H_j^s$ ($j = j_l$) during the period $[t_{l,1}, t_{l,2})$; the order in which we schedule the processing of these agents is immaterial. We use summation by parts to do the accounting. Picture a slow, continuous motion. An infinitesimal motion δ at any time requires $(\alpha^q + 1)\delta$ credits, with q the current rank of agent i: $\alpha^q\delta$ to maintain the credit invariant and δ to pay for the actual move. Whenever i passes over another agent m because $a_i(t)$ reaches a value $x = a_m(t)$ along the way and $b_i > x$, we switch the identities of agents i and m. This is done only for accounting purposes in the analysis of step [1.1]. This way, agent m can then proceed to the right to complete the journey of agent i if need be. Further swaps may then occur before the motion $a_i \to b_i$ is over. The benefit of this perspective is that the ranks of the new, so-called *virtual*, agents no longer change: only their positions do. Instead of tallying the individual costs of each right shift $a_i \to b_i$, we simply add up the costs of the fixed-ranked virtual agents.

Each one of the n virtual agents has a distinct rank k at time $t_{l,1}$: call it v_k. Processing H_j^s may move some of them to the right (perhaps more than once in fact), so let $D_k \geq 0$ be the total distance traveled by agent v_k while processing H_j^s. Any shift in v_k at time $t \in [t_{l,1}, t_{l,2})$ implies the existence of an interval $[a_i, b_i] \subset [a_i, a_{r(i)})$, where $i \in H_j^s$, that contains the position of v_k at time t. Being in H_j^s or P_s, the $k_i - 1$ agents $r(i), r(i,2), \ldots, r(i, k_i - 1) = j$ did not move between during $[0, t_{l,1})$. (Note that $k_i > 1$.) They may have moved to the right between $t_{l,1}$ and t but they always remained between the moving virtual agent v_k and the fixed position a_j: the reason is that, by definition, v_k

cannot overtake any agent, while no agent in H_j^s can right-shift past a_j. As we mentioned earlier, they cannot have moved left either. It follows that $k_i \le rk(a_j(t)) - k + 1$. (Observe that $a_j(t) = a_j$ but $rk(a_j(t))$ may be different from the original rank of a_j at time 0.) Of all the a_i's used in relation to v_k over the course of step [1.1], we take the index i corresponding to the smallest a_i. We derive an upper bound on the distance traveled by v_k in the time interval $[t_{l,1}, t_{l,2})$. For any k, there exist i and $t \in [t_{l,1}, t_{l,2})$ such that

$$D_k \le a_j - a_i = \sum_{l=0}^{k_i-2} \Delta_{r(i,l)} \le \sum_{l=0}^{k_i-2} \Delta_j \Big(\frac{\rho}{2}\Big)^{l+1-k_i} \le \frac{2\Delta_j}{2-\rho}\Big(\frac{\rho}{2}\Big)^{1-k_i} \le 2\Delta_j \Big(\frac{\rho}{2}\Big)^{k-rk(a_j(t))}.$$

Summing over all virtual agents v_k allows us to bound the number A_l of credits that agent $j = j_l$ needs to cover the costs incurred in processing H_j^s in the time period $[t_{l,1}, t_{l,2})$. No agent of H_j^s can move to the right of j during $[t, t_{l,2})$ and, of course, none can pass j from right to left in that time interval; therefore,

$$rk(a_j(t)) = rk(a_j(t_{l,2})) > 1,$$

and agent j_l is charged

$$A_l = \sum_{k=1}^{rk(a_j(t_{l,2}))-1} (\alpha^k + 1) D_k \le \sum_{k=1}^{rk(a_j(t_{l,2}))-1} 2\Delta_j (\alpha^k + 1)\Big(\frac{\rho}{2}\Big)^{k-rk(a_j(t_{l,2}))} \le \Delta_j \alpha^{rk(a_j(t_{l,2}))}, \tag{14}$$

where, we recall, $\alpha = 9/\rho$ and $t_{l,2}$ is the time at which step [1.1] ends for fixed s and $j = j_l$.

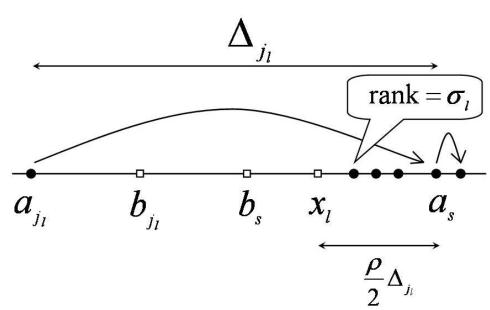

Figure 7: Covering the charges of agent $j = j_l$.

Step [1.2] We keep the values of s and $j = j_l \in P_s$ set above. By Lemma 3.2, stoppers can only move left so, by scheduling stoppers from right to left, we ensure that s has not moved yet. Recall that, by definition of $j_l \in P_s$, we have $k_{j_l} = 1$ and $s = r(j_l)$. As illustrated in Figure 7, define

$$x_l = a_s - \tfrac{1}{2}\rho\Delta_{j_l}\,. \tag{15}$$

By definition, $j_l \ne s$; therefore $\Delta_{j_l} > 0$; by Rule 2, the larger of a_{j_l} and b_{j_l} does not exceed $a_s - \rho\Delta_{j_l} < x_l$. Applying Lemma 3.2 to agent j_l shows that, in step [2], the stopper s will move left past x_l, as, indeed, $b_s = b_{r(j_l,k_{j_l})} = b_{r(j_l,1)} \le a_s - \frac{1}{2}(1+\rho)\rho\Delta_{j_l} < x_l$. Summarizing these bounds,

$$\begin{cases} \max\{\, a_{j_l},\, b_{j_l} \,\} < x_l\,, \quad \text{for each } \ l\,; \\ b_s < \min\{\, x_l \,|\, 1 \le l \le \nu \,\}\,. \end{cases} \tag{16}$$

We introduce

$$\sigma_l = \min\Big\{ rk(a_i(t_{l,2})) \,|\, x_l \le a_i(t_{l,2}) \text{ and } 1 \le i \le n \Big\}. \tag{17}$$

Informally, σ_l would be the rank of x_l, were it the position of an actual agent, measured once we are done handling $H^s_{j_l}$ in step [1.1]. Since $a_s(t_{l,2}) = a_s > x_l$, the agent of rank σ_l exists and, by definition, is unique. Furthermore, by (16), it will not change even if agent j_l is shifted to b_{j_l} during $[t_{l,2}, t_{l+1,1})$ in step [1.2]. We define μ_l inductively by $\mu_1 = \sigma_1$ and

$$\mu_l = \min\{\sigma_l, \mu_{l-1} - 1\},$$

for $1 < l \le \nu$. We show that agent j_l needs to borrow no more than B_l credits from the bank, where

$$B_l = |x_l - a_s|\Big(1 - \frac{1}{\alpha}\Big)\alpha^{\mu_l}. \tag{18}$$

The key fact driving our analysis is that

$$\mu_l \ge rk(a_{j_l}(t_{l+1,1})) + 1\,. \tag{19}$$

We prove this by induction, beginning with the case $l = 1$. By (16), $x_l > \max\{\, a_{j_l},\, b_{j_l}\,\} = a_{j_l}(t_{l+1,1})$. (Recall that $t_{l+1,1}$ is the time when agent j_l has just been moved to b_{j_l} if it is to the right.) This shows that agent j_l cannot reach across to the right of x_l; therefore, the agent i of rank σ_l at time $t_{l,2}$ still has that same rank at time $t_{l+1,1}$. Since $a_i(t_{l+1,1}) = a_i(t_{l,2}) \ge x_l > a_{j_l}(t_{l+1,1})$, it then follows that

$$\mu_l = \sigma_l = rk(a_i(t_{l,2})) = rk(a_i(t_{l+1,1})) > rk(a_{j_l}(t_{l+1,1}))\,;$$

hence (19). Suppose now that $l > 1$. The argument we just gave can be repeated verbatim for the case $\mu_l = \sigma_l$, so we can assume that $\sigma_l > \mu_{l-1} - 1$, and that $\mu_l = \mu_{l-1} - 1$. The key observation is that agents j_l are processed by nonincreasing value of $\max\{a_{j_l}, b_{j_l}\}$. By lazy tie-breaking, therefore, the value of $rk(a_{j_l}(t_{l+1,1}))$ strictly decreases as l goes from 1 to ν. (This is not simply a result of agent j_l landing at or to the left of agent j_{l-1}, but also of the absence of leftward motion; indeed, note that the strict monotonicity of $rk(a_{j_l}(t_{l+1,1}))$ is a claim about ranks at *different* times.) This immediately implies (19) by induction. By (14), agent j_l was charged in step [1.1] a number of credits

$$A_l \le \Delta_{j_l}\alpha^{rk(a_{j_l}(t_{l,2}))} \le \Delta_{j_l}\alpha^{rk(a_{j_l}(t_{l+1,1}))}$$

credits. If step [1.2] moves the agent j_l to b_{j_l} (ie, $b_{j_l} > a_{j_l}$) then this debt will increase by no more than $(1 + \alpha^{rk(a_{j_l}(t_{l+1,1}))})\Delta_{j_l}$: the term Δ_{j_l} is to pay for the motion itself; the other term is a conservative upper bound easily derived from the "continuous motion" argument. In both cases, therefore, agent j_l is charged $A'_l \le 3\alpha^{rk(a_{j_l}(t_{l+1,1}))}\Delta_{j_l}$. By (15, 18, 19) and $\alpha = 9/\rho$,

$$B_l \ge \frac{\rho}{2}\Big(1 - \frac{1}{\alpha}\Big)\alpha^{rk(a_{j_l}(t_{l+1,1}))+1}\Delta_{j_l} \ge \frac{\rho}{6}(\alpha - 1)A'_l \ge A'_l\,,$$

which proves that agent j_l has enough credits borrowed from the bank to pay for step [1].[5]

[5] The reader may have noticed that we did not use the factor of $1 - 1/\alpha$ in the expression for B_l given in (18). Why then do we have it? This is a technical trick to make the geometric-like sum in (21) "telescope."

Step [2] We need to tally the release of credits past time $t_{\nu+1,1}$ produced by moving the stoppers to the left and show there are at least $B_1 + \cdots + B_\nu$ of them. Turning once again to the summation-by-parts argument, we track the leftward motion of s and rewrite the sum as an integral,

$$\sum_{l=1}^{\nu} B_l = -\int_{a_s}^{b_s} \sum_{l=1}^{\nu} (\alpha^{\mu_l} - \alpha^{\mu_l - 1}) I_l(u)\, du,$$

where $I_l(u) = 1$ if $x_l \leq u \leq a_s$ and 0 otherwise. We need the minus sign because the integral runs in the negative direction. At time $t_{\nu+1,1}$, the stopper s proceeds to the left continuously until it reaches b_s. Since leftward travel costs nothing, an infinitesimal motion δ at position u releases $\delta\alpha^{q(u)}$ credits, where $q(u)$ is the lowest rank of any agent, at time $t_{\nu+1,1}$, at or to the right of u (or n if there is no such thing). (Note that $q(u)$ compares the moving agent s against a snapshot of the agents, excluding s, taken at the beginning of step [2].) We can also express the release as the integral $R = -\int_{a_s}^{b_s} \alpha^{q(u)}\, du$. It suffices to prove that

$$\alpha^{q(u)} \geq \sum_{l=1}^{\nu} (\alpha^{\mu_l} - \alpha^{\mu_l - 1}) I_l(u). \tag{20}$$

If $l_1 < \cdots < l_k$ are the indices l such that $I_l(u) = 1$ then, by the strict monotonicity of μ_l,

$$\sum_{l=1}^{\nu} (\alpha^{\mu_l} - \alpha^{\mu_l - 1}) I_l(u) = \sum_{j=1}^{k} (\alpha^{\mu_{l_j}} - \alpha^{\mu_{l_j} - 1}) \leq \alpha^{\mu_{l_1}}. \tag{21}$$

Because of the ordering of j_l, by (16), the rank of agent j_l remains unchanged between $t_{l,2}$ and $t_{\nu+1,1}$, so we could replace $t_{l,2}$ by $t_{\nu+1,1}$ in the definition of σ_l in (17) and get the same value of σ_l. This shows that μ_{l_1}, being at most σ_{l_1}, does not exceed the rank of the smallest-ranked $a_i(t_{\nu+1,1}) \geq x_{l_1}$. Given $I_{l_1}(u) = 1$, we have $u \geq x_{l_1}$; hence $q(u) \geq \mu_{l_1}$ and (20). This establishes the soundness of our credit invariant: allocating each agent i at time 0 with $a_i \alpha^{rk(a_i)}$ credits is sufficient to pay for all the costs, while maintaining the credit invariant. Since each $a_i \leq 1$, this implies that

$$\tfrac{1}{2} M \leq \sum_{i=1}^{n} a_i \alpha^{rk(a_i)} \leq \sum_{i=1}^{n} \alpha^i < 2\Big(\frac{9}{\rho}\Big)^n.$$

The proof of Theorem 3.1 (and hence the upper bound of Theorem 1.1 for $s = 1$) follows from Lemma 3.3. □

3.3 The General Case: $s < 1$

We prove the upper bound of Theorem 1.1 for $0 < s < 1$. Instead of assuming unit diameter, we allow any initial configuration within the unit cube $[0,1]^d$; the notation $F_n(s)$ refers to the maximum total s-energy under these conditions. Obviously, $E_n(s) \leq F_n(s)$. We show that the total s-energy satisfies the recurrence: $F_1(s) = 0$ and, for $n \geq 2$,

$$F_n(s) \leq 2n F_{n-1}(s) + (1 - (\rho/4d)^n)^s F_n(s) + d^{s/2} n^3. \tag{22}$$

Consider a Cartesian system of reference and let $y_i(t)$ again denote the first coordinate of agent i at time t. If the neighbors of agent i span an interval $[\alpha, \beta] \subseteq [0,1]$ then, by (3),

$$y_i(t+1) \in (1-\gamma)[\alpha, \beta] + \tfrac{1}{2}(\alpha + \beta)\gamma, \tag{23}$$

where $\gamma = \rho/d$. All agents are initially "dry," except for a selected agent, denoted by its index 1, which is wet and will spread its "wetness" from one agent to the next, causing the geometry to change in the process. Once wet, an agent always remains so. Again, we use an algorithmic proof:

[1] Initially, all agents are dry except for agent 1. Set $S(0) = \{y_1(0)\}$.

[2] For $t = 0, 1, \ldots, \infty$:

- **[2.1]** Declare wet any agent adjacent to a wet agent in G_t.
- **[2.2]** $S^*(t) \leftarrow S(t) \cup \{$ positions at time t of dry agents just turned wet $\}$.
- **[2.3]** Move each agent i from $y_i(t)$ to $y_i(t+1)$. [*If no newly wet agent, then we may carry all motion within* $S(t) = S(t^*)$ *in isolation from the* $n - |S(t)|$ *other agents.*]
- **[2.4]** $S(t+1) \leftarrow \{$ positions at time $t+1$ of agents corresponding to $S^*(t)$ $\}$.

The sets $S^*(t)$ track wetness propagation. We interpret both $S(t)$ and $S^*(t)$ as multisets, for example by labeling each position with the agent's index. Let $\|S(t)\|$ denote the length of the smallest interval enclosing $S(t)$ and let $\{t_k\}_{k\geq 1}$ be the times $t \geq 0$, in chronological order, at which $|S^*(t)| > |S(t)|$. Recall that ρ is smaller than a suitable constant. We show that:

$$\|S(t_k)\| \leq 1 - \left(\frac{\gamma}{4}\right)^k. \tag{24}$$

The smallest interval $[a, b]$ defining $\|S(t_k)\|$ is in $[0, 1]$. By flipping the interval $[a, b]$ defining if necessary, we can assume that $a + b \geq 1$. Because $\|S(t_1)\| = 0$, we can safely assume by induction that (24) holds up to t_k; hence $a \geq \frac{1}{2}(\gamma/4)^k$. Since $\|S(t)\|$ can increase only when at least one dry agent becomes wet, ie, at times of the form $t = t_l$, we can prove (24) for t_{k+1} by showing that $\|S(t_k+1)\| \leq 1 - (\gamma/4)^{k+1}$, which follows from $[0, \frac{1}{2}a\gamma) \cap S(t_k+1) = \emptyset$. We proceed by contradiction. Consider an agent i contributing to $S(t_k + 1)$ with $y_i(t_k + 1) < \frac{1}{2}a\gamma$. In G_{t_k}, the set $N_i(t_k)$ includes at least one (not newly) wet agent; therefore, in (23), $\beta \geq a$ and hence $y_i(t_k + 1) \geq \frac{1}{2}a\gamma$, which is impossible and proves (24).

The set $S(t_k)$ can only gain agents, as k grows, but the set may stop growing before it absorbs all of them. Note this has to do with the graphs G_t and not with the positions of the agents, so it can be used uniformly along each coordinate axis. When t is not of the form t_k, step [2.3] indicates that the adversary can act on $S(t)$ in isolation from the rest. Therefore, the s-energy expended during steps $t_{k-1}, \ldots, t_k - 1$ is bounded by $F_{|S(t_k)|}(s) + F_{n-|S(t_k)|}(s)$. At time t_k, the extra energy involved is

$$\sum_{(i,j)\in G_{t_k}} \|x_i(t) - x_j(t)\|_2^s \leq \binom{n}{2} d^{s/2}.$$

Using obvious monotonicity properties, it follows that, up to the highest value of t_k, the s-energy is bounded by

$$\sum_{l=1}^{n-1} \left\{ F_l(s) + F_{n-l}(s) + \binom{n}{2} d^{s/2} \right\} \leq 2nF_{n-1}(s) + n^3 d^{s/2}.$$

When t_k reaches its highest value t, if $|S(t+1)| < n$ then all the energy has been accounted for above. Otherwise, we must add the energy expended by the n agents past t. By (24), however, at time $t+1$, the n agents fit within a cube of side length $1-(\rho/4d)^n$. So, all we need to do is add $(1-(\rho/4d)^n)^s F_n(s)$ inductively to the s-energy; hence (22). The case $n=2$ is worthy of attention because it is easy to solve exactly. The problem is inherently one-dimensional, so we can assume that the two agents start at opposite corners of the unit d-cube and move toward each other by the minimum allowed distance. This gives us the equation $F_2(s) = d^{s/2} + (1-\rho)^s F_2(s)$; hence, by (13),

$$F_2(s) = \frac{d^{s/2}}{1-(1-\rho)^s} \leq \frac{d^{s/2}}{s\rho}\,. \tag{25}$$

We now consider the case $n > 2$. By (13, 22),

$$F_n(s) \leq \frac{2nF_{n-1}(s) + n^3 d^{s/2}}{s(\rho/4d)^n}\,.$$

By (25) and the monotonicity of $F_n(s)$, we verify that the numerator is less than $3n^3 F_{n-1}(s)$; therefore, for $n > 2$,

$$F_n(s) \leq \frac{3n^3 F_{n-1}(s)}{s(\rho/4d)^n} \leq s^{1-n} \rho^{-n^2 - O(1)}.$$

This proves the upper bound of Theorem 1.1 for $s < 1$. □

We now turn to Theorem 2.5. Recall that $\mathcal{P}$ is the family of n-by-n stochastic matrices such that each $P \in \mathcal{P}$ satisfies: (i) each diagonal entry is nonzero; each positive entry is at least p; no pair p_{ij}, p_{ji} contains exactly one zero. The entry (i,j) of a product of t such matrices can be viewed as the position of agent i after t iterations of a bidirectional system with $\rho = 2p$, by (5), and an initial vector consisting of 0 everywhere, except for 1 at position j. Referring back to the boxed algorithm, we designate agent j as the one initially wet, with all the others dry. Let $m(t)$ be the minimum value in $S(t)$. At every time t_k when $S(t)$ grows in size, the minimum $m(t)$ cannot approach 0 closer than $\rho(m(t)/2)$. Since $|\{t_k\}| < n$, agent j cannot be at a position smaller than p^{n-1}. The lower bound proof suggests a trivial construction that achieves the very same bound and therefore proves its optimality. This completes the proof of Theorem 2.5. □

3.4 The Lower Bounds

We prove the lower bounds in Theorems 1.1 and 1.4.

The case $s < 1$**.** We describe an algorithm $\mathcal{A}_n(a,b)$ that moves n agents initially within $[a,b]$ toward a single point $a + (b-a)x(n)$ while producing a total s-energy equal to $(b-a)^s E(n,s)$. Clearly, $E(1,s) = 0$, so assume $n > 1$. We specify $\mathcal{A}_n(0,1)$ as follows. Place $n-1$ agents at position 0 and one at position 1. The graph G_0 consists of a single edge between agent 1 at position 1 and agent 2 at position 0. At time 0, agent 2 moves to position $\alpha = \rho/2$ while agent 1 shifts to $1-\alpha$. The $n-2$ other agents stay put. Next, apply $\mathcal{A}_{n-1}(0,\alpha)$ to the set of all agents except 1. This brings them to position $\alpha x(n-1)$. Finally, apply $\mathcal{A}_n(\alpha x(n-1), 1-\alpha)$ to all the agents. The operations of $\mathcal{A}_n$ leave the center of mass invariant, so if $x(n)$ exists it must be $1/n$. Here is a formal argument. The attractor point $x(n)$ satisfies the recurrence

$$x(n) = \alpha x(n-1) + (1 - \alpha x(n-1) - \alpha)x(n),$$

where, for consistency, $x(1) = 1$. This implies that

$$\frac{1}{x(n)} = 1 + \frac{1}{x(n-1)};$$

therefore $x(n) = 1/n$, as claimed. The total s-energy $E(n,s)$ satisfies the relation: $E(1,s) = 0$; and, for $n > 1$,

$$\begin{aligned} E(n,s) &= \alpha^s E(n-1,s) + (1 - \alpha x(n-1) - \alpha)^s E(n,s) + 1 \\ &\geq \frac{\alpha^s E(n-1,s) + 1}{1 - (1-2\alpha)^s} \geq \frac{\alpha^{(n-2)s}}{(1-(1-2\alpha)^s)^{n-1}} \,. \end{aligned}$$

Since $\alpha = \rho/2$ is small enough, $(1-2\alpha)^s \geq 1 - 3\alpha s$ and $E(n,s) \geq s^{1-n}\rho^{-\Omega(n)}$, for any n large enough, $s \leq s_0$, and fixed $s_0 < 1$ (with ρ going to 0 as s_0 tends to 1). We observe that Algorithm $\mathcal{A}_n$ cannot start the second recursive call before the first one is finished, which literally takes forever. This technicality is easily handled, however, and we skip the discussion. This completes the proof of the lower bound of Theorem 1.1 for $s < 1$. □

The case $s = 1$. Suppose that each G_t joins the two nodes of a 2-agent system. The length of the edge can be made to shrink exponentially fast at a rate of $1-\rho$. We show that having n agents allows us to mimic the behavior of a 2-agent system with ρ replaced by $\rho^{\Theta(n)}$. Without loss of generality, we assume that n is an even integer $2m \geq 4$. Our construction is symmetric by reflection along the X-axis about the origin, so we label the agents $-m, \ldots, -1, 1, \ldots, m$ from left to right, and restrict our discussion to the m agents with positive coordinates. (Equivalently, we could fix one agent.) The evolution of the system consists of phases denoted by $\theta = 0, 1, 2$, etc. At the beginning of phase θ, agent i lies at $x_1(\theta) = (1-\rho^m)^\theta$ for $i = 1$ and at[6]

$$x_i(\theta) = x_{i-1}(\theta) + \rho^{i-1}(1-\rho^m)^\theta,$$

for $2 \leq i \leq m$. As usual, we assume that $\rho > 0$ is small enough. The system includes a mirror image of this configuration about the origin at all times. Note that all the agents are comfortably confined to the interval $[-2, 2]$, so the diameter D is at most 4.

We now describe the motion at phase θ in chronological order, beginning with agent m. During phase θ, the first graph G_t ($t = \theta m$) consists of exactly two edges: one joining m and $m-1$ (with its mirror image across $x = 0$); the graph G_{t+1} joins $m-1$ with $m-2$ (and its mirror image); etc. The last graph in phase θ, G_{t+m-1}, follows a different pattern: it joins the two agents indexed 1 and -1. Except for m, all of these agents (to right of the origin) are moved twice during phase θ: first to the right, then to the left. Specifically, agent $1 \leq i < m$ moves right at time $t+m-i-1$ and left at time $t+m-i$. We use barred symbols to denote the intermediate states, ie, the location after the rightward moves. At phase θ,

$$G_t : \begin{cases} x_m(\theta+1) = \alpha_m x_{m-1}(\theta) + (1-\alpha_m)x_m(\theta) = (1-\rho^m)x_m(\theta); \\ \bar{x}_{m-1}(\theta) = \tfrac{1}{2}x_{m-1}(\theta) + \tfrac{1}{2}x_m(\theta) = x_{m-1}(\theta) + \tfrac{1}{2}\rho^{m-1}(1-\rho^m)^\theta \,. \end{cases}$$

We easily verify the identities above for $\alpha_m = (\rho - \rho^{m+1})/(1-\rho)$. For $i = m-1, m-2, \ldots, 2$, with G_{t+m-i} joining agent $i-1$ and i, the two moves are specified by:

$$G_{t+m-i} : \begin{cases} x_i(\theta+1) = \alpha_i x_{i-1}(\theta) + (1-\alpha_i)\bar{x}_i(\theta) = (1-\rho^m)x_i(\theta)\,; \\ \bar{x}_{i-1}(\theta) = (1-\beta_i)x_{i-1}(\theta) + \beta_i \bar{x}_i(\theta) = x_{i-1}(\theta) + \tfrac{1}{2}\rho^{i-1}(1-\rho^m)^\theta \,, \end{cases}$$

[6] We deviate from our usual notation by letting the argument of $x_i(\theta)$ refer to the phase of the construction and not the time t.

where $\beta_i = 1/(2+\rho)$ and

$$\alpha_i = \frac{\rho}{2+\rho} + \frac{2(1-\rho^i)\rho^{m-i+1}}{(1-\rho)(2+\rho)}\,.$$

Finally, at time $t+m-1$, choosing $\alpha_1 = (\rho + 2\rho^m)/(4+2\rho)$ allows us to write

$$G_{t+m-1}:\ x_1(\theta+1) = -\alpha_1 \bar{x}_1(\theta) + (1-\alpha_1)\bar{x}_1(\theta) = (1-\rho^m)x_1(\theta)\,.$$

All the coefficients α_i are $\Theta(\rho)$, so we can rescale ρ by a constant factor to make the dynamics conform to a standard one-dimensional bidirectional agreement system with parameter ρ. Obviously the system converges to consensus. In each phase θ, the union of the intervals formed by the edges of all of that phase's graphs G_t covers $[-x_m(\theta), x_m(\theta)]$; therefore, the total 1-energy is at least

$$2\sum_{\theta=0}^{\infty} x_m(\theta) = \frac{2(1-\rho^m)}{1-\rho}\sum_{\theta=0}^{\infty}(1-\rho^m)^\theta > \rho^{-m}.$$

This proves the lower bound of Theorem 1.1 for $s=1$. For any positive $\varepsilon < 1/2$, the length of the edge in G_{t+m-1}, which is $2x_1(\theta)$, does not fall below ε until θ is on the order of $\rho^{-m}\log\frac{1}{\varepsilon}$, which establishes the lower bound of Theorem 1.4. □

3.5 The Remaining Proofs

Proof of Theorem 1.3. Let $t_0 - 1$ be the time of the last (ie, C_ε-th) nontrivial step. Iterate on the following process: initially make each agent at time t_0 its own one-vertex convex hull; then as long as two convex hulls are within distance ε, remove them and replace them by the convex hull of their union. (Pick any pair if there are several candidates.) By construction, no two of the final hulls are within ε of each other. But, by convexity, each one of them can only evolve, at time $t > t_0$, within the region it occupies at time t_0, so the distance between any two of them cannot decrease ever after; hence it remains greater than ε. Since all subsequent steps at and following t_0 are trivial, further interaction and motion is confined to within each hull. By induction, we see that if any of these final hulls at time t_0 has k points then its diameter cannot exceed $(k-1)\varepsilon$. This proves that the system has εn-converged at time t_0. Assume now that the communication network remains connected at all times. Let $t = t_\varepsilon$ be the first time at which all the edges of G_t are of length at most ε. Obviously, $t_\varepsilon \leq C_\varepsilon$. Since G_{t_ε} is connected, the diameter of the set of n agents is less than εn: this means that the system has εn-converged to consensus. □

Proof of Theorem 1.8. We know by Theorem 1.3 that the system ε-converges eventually. Suppose that agents i and j are more than ε apart at time t_1. By assumption, i and j communicate during some time interval $[t_1, t_2]$. (It does not matter who communicates with whom.) Picture all agents having dry hands at time t_1, except for i and j, whose hands are respectively white and red. At any time $t \geq t_1$, each pair of neighbors in G_t shake hands. Wetness propagates according to the obvious rules: (i) two wet (shaking) hands remain wet; (ii) two dry hands remain dry; (iii) two hands of mixed status both become wet. Because of the self-loops, once wet, an agent remains wet forever. An agent can be dry or wet and the latter category has three types: white; red; white and red.

Consider the smallest ball W_t centered at $x_i(t_1)$ (ie, at the position of agent i at time t_1) that encloses all the white agents at time t; similarly, R_t is the smallest ball centered at $x_j(t_1)$ that encloses all the red agents at time t. Suppose that, between t_1 and t_2, no edge of G_t exceeds $\varepsilon/2n$ in length. Then, by convexity, no agent can move by more than $\varepsilon/2n$ in one step. Since the edges along which the white water flows are of length at most $\varepsilon/2n$, the radius of W_t can grow by at most

$\varepsilon/2n + \varepsilon/2n$ in a single step: $\varepsilon/2n$ for the acquisition of any new agent and another $\varepsilon/2n$ for its displacement. Note that it can also shrink. By convexity, however, the ball W_t can grow only if it absorbs at least one new white agent; therefore, it can grow at most $l-1$ times if l is the number of white agents absorbed by it (including i). The same is true of R_t. This shows that, as long as they do not meet, the balls have a combined radius of at most $(n-2)\varepsilon/n$, hence remain more than $2\varepsilon/n$ apart. By our previous upper bound on the single-step growth of these balls, they can therefore never meet; it follows that communication between i and j is impossible. This implies the existence of an edge in G_t ($t_1 \leq t \leq t_2$) of length greater than $\varepsilon/2n$. The number of such times t is bounded by $C_{\varepsilon/2n}$. After all steps witnessing an edge of length greater than $\varepsilon/2n$ have elapsed, agents i and j stay permanently within ε of each other (else they could never communicate again, which would contradict our assumption). Since we count communications only over nonoverlapping time intervals, the number of them following a separation between agents i and j in excess of ε is at most $C_{\varepsilon/2n}$. The proof follows then from Theorem 1.4. □

Acknowledgments

I wish to thank Ali Jadbabaie and Peter Sarnak for helpful discussions.

References

[1] Angeli, D., Bliman, P.-A. *Stability of leaderless multi-agent systems: extension of a result by Moreau*, Mathematics of Control, Signals & Systems 18 (2006), 293–322.

[2] Blondel, V.D., Hendrickx, J.M., Tsitsiklis, J.N. *On Krause's multi-agent consensus model with state-dependent connectivity*, IEEE Transactions on Automatic Control 54, 11 (2009), 2586–2597.

[3] Cao, M., Morse, A.S., Anderson, B.D.O. *Coordination of an asynchronous multi-agent system via averaging*, Proceedings 2005 IFAC Congress, 2005.

[4] Cao, M., Spielman, D.A. Morse, A.S. *A lower bound on convergence of a distributed network consensus algorithm*, 44th IEEE Conference on Decision and Control, and the European Control Conference 2005, Seville, Spain, 2005.

[5] Chazelle, B. *Natural algorithms*, Proc. 20th Annual ACM-SIAM SODA 2009, 422–431.

[6] Chazelle, B. *The convergence of bird flocking*, arXiv:0905.4241v1, 2009.

[7] Chazelle, B. *Analytical tools for natural algorithms*, Proc. 1st Symposium on Innovations in Computer Science (2010), Beijing, 32–41.

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

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

[10] Earl, M.G., Strogatz, S.H. *Synchronization in oscillator networks with delayed coupling: a stability criterion*, Phys. Rev. E, 67 (2003), 036204(1–4).

[11] Grötschel, M., Lovász, L., Schrijver, A. Geometric Algorithms and Combinatorial Optimization, Springer-Verlag, 1988.

[12] Hardy, G.H., Riesz, M. The General Theory of Dirichlet's Series, Cambridge University Press, 1915.

[13] Hegselmann, R., Krause, U. *Opinion dynamics and bounded confidence models, analysis, and simulation*, J. Artificial Societies and Social Simulation 5, 3 (2002).

[14] Hegselmann R, Krause U. *Truth and cognitive division of labor: first steps towards a computer aided social epistemology*, J. Artificial Societies and Social Simulation 9 (2006).

[15] 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.

[16] 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.

[17] Jadbabaie, A., Motee, N., Barahona, M. *On the stability of the Kuramoto model of coupled nonlinear oscillators*, Proc. American Control Conference 5 (2004), 4296–4301.

[18] Krause, U. *A discrete nonlinear and non-autonomous model of consensus formation*, Communications in Difference Equations (2000), 227236.

[19] Kurz, S., Rambau, J. *On the Hegselmann-Krause conjecture in opinion dynamics*, Journal of Difference Equations and Applications (2009).

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

[21] Li, C., Xu, H., Liao, X., Yu, J. *Synchronization in small-world oscillator networks with coupling delays*, Physica A 335 (2004), 359–364.

[22] Lorenz, J. *Multidimensional opinion dynamics when confidence changes*, Economic Complexity, Aix-en-Provence, May 2003.

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

[24] Lorenz, D., Lorenz, J. *Convergence to consensus by general averaging*, in R. Bru and S. Romero-Vivo (eds), Positive Systems (2009), LNCS Springer 91–99.

[25] Martinez, S. *A convergence result for multiagent systems subject to noise*, Proc. American Control Conference (2009), St. Louis, Missouri.

[26] Martinez, S., Bullo, F., Cortés, J., Frazzoli, E. *On synchronous robotic networks – Part II: Time complexity of rendezvous and deployment algorithms*, IEEE Transactions on Automatic Control 52 (2007), 2214–2226.

[27] McGraw, P.N., Menzinger, M. *Analysis of nonlinear synchronization dynamics of oscillator networks by Laplacian spectral methods*, Phys. Rev. E 75 (2007), 027104(1–4).

[28] Montenegro, R., Tetali, P. Mathematical Aspects of Mixing times in Markov Chains, Foundations and Trends in Theoretical Computer Science 1, 3 (2006), now Publishers Inc., 237–354.

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

[30] Olshevsky, A., Tsitsiklis, J.N. *Convergence speed in distributed consensus and averaging*, SIAM Journal on Control and Optimization 48 (2009), 33–55.

[31] Papachristodoulou, A., Jadbabaie, A. *Synchonization in oscillator networks: switching topologies and non-homogeneous delays*, Proc. IEEE CDC-ECC (2005), 5692–5697.

[32] Papachristodoulou, A., Jadbabaie, A. *Synchonization in oscillator networks with heterogeneous delays, switching topologies and nonlinear dynamics*, Proc. IEEE CDC-ECC (2006), 4307–4312.

[33] Scardovi, L., Sarlette, A., Sepulchre, R. *Synchronization and balancing on the N-torus*, Systems & Control Letters 56 (2007), 335–341.

[34] Strogatz, S.H. *From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators*, Physica D 143 (2000), 1–20.

[35] Triplett, B.I., Klein, D.J., Morgansen, K.A. *Discrete time Kuramoto models with delay*, Proc. Networked Embedded Sensing and Control Workshop, Lecture Notes in Control and Information Sciences, Springer, 2005.

[36] Tsitsiklis, J.N., *Problems in decentralized decision making and computation*, PhD thesis, MIT, 1984.

[37] Tsitsiklis, J.N., Bertsekas, D.P., Athans, M. *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Transactions on Automatic Control 31 (1986), 803–812.

[38] Vicsek, T., Czirók, 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.

[39] Wang, Q., Duan, Z., Chen, G., Feng, Z. *Synchronization in a class of weighted complex networks with coupling delays*, Physica A 387 (2008), 5616–5622.

[40] Winfree, A.T. *Biological rhythms and the behavior of populations of coupled oscillators*, J. Theoret. Bio. 16, 1 (1967), 15–42.

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

[42] Zanette, D.H. *Propagation of small perturbations in synchronized oscillator networks*, Europhys. Lett. 68 (2004), 356.