Swarming on Random Graphs

James von Brecht   Theodore Kolokolnikov   Andrea L. Bertozzi   Hui Sun

August 18, 2012

Abstract

We consider a compromise model in one dimension in which pairs of agents interact through first-order dynamics that involve both attraction and repulsion. In the case of all-to-all coupling of agents, this system has a lowest energy state in which half of the agents agree upon one value and the other half agree upon a different value. The purpose of this paper is to study the behavior of this compromise model when the interaction between the N agents occurs according to an Erdős-Rényi random graph $\mathcal{G}(N, p)$. We study the effect of changing p on the stability of the compromised state, and derive both rigorous and asymptotic results suggesting that the stability is preserved for probabilities greater than $p_c = O(\frac{\log N}{N})$. In other words, relatively few interactions are needed to preserve stability of the state. The results rely on basic probability arguments and the theory of eigenvalues of random matrices.

Keywords: swarm equilibria, biological aggregations, random graphs

1   Introduction

Swarming is an ubiquous natural phenomenon which occurs at all levels of animal kingdom, from bacterial colonies to schools of fish and human crowds. One of the simplest ways to reproduce swarming mathematically is an aggregation model. In this model, each individual is modelled by a particle that typically interacts with all other particles according to a specified potential function. Typically it is assumed that the particles repel each other at short distances and attract each other at longer distances. In many cases this leads to the formation of swarms. This model and its variants are also often used in robotics literature. As has been observed recently in the literature, the behavior of the interaction potential can lead to highly complex patterns [14, 23, 26, 25]. Most notably we find that the potential can be classified as one in which the system is "confining" in the large N limit (where N is the number of particles) or non-confining. For second-order systems the language used is H-stability vs. catastrophic to describe the energy per particle in the limit of large N [9, 5]. For this paper we are interested in the confining case such that the $N \to \infty$ limit has stable solutions that are confined to some finite region, typically approximating a continuum density (possibly concentrated as a measure). The paper focuses specifically on a simple model in one dimension in which the stable solution has two clusters of points regardless of the number of agents or the underlying interaction structure.

In practice, and especially for large system sizes, the assumption of all-to-all interaction is often unrealistic. In robotics, for instance, all-to-all communication can prove prohibitively expensive

for a large number of robots. Thus it is quite relevant to understand the stability of such systems with a relatively sparse network interaction topology. In this paper we study what happens to the system when the assumption of all-to-all communication is relaxed. Specifically, we consider the aggregation model on a network, where the particles are represented by vertices on a graph that interact only if there is an edge between them. The basic particle model then becomes
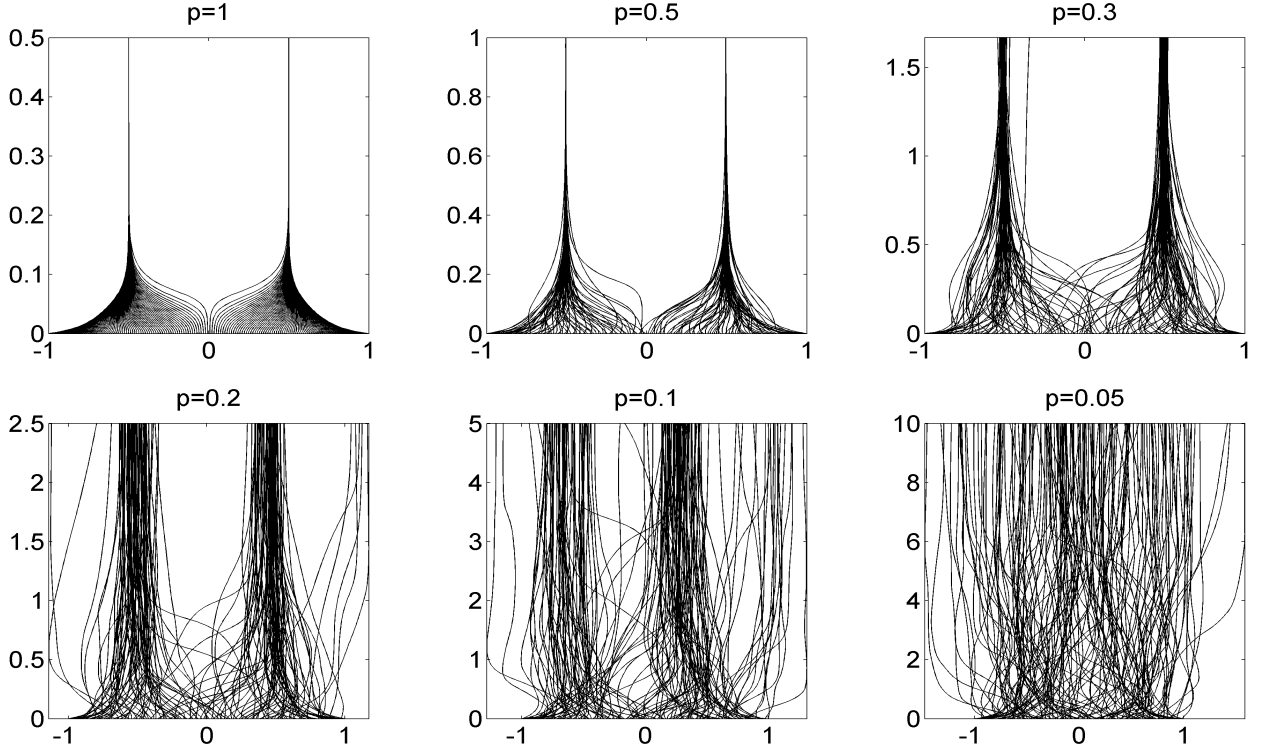

Figure 1: Dynamics of the aggregation model (1) for a p-random graph (i.e. with e_{ij} given by (4)) for a range of different p with $N = 150$ and $\kappa = 0.5$. The y-axis denotes time.

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{1}{N}\sum_{k=1}^{N} e_{ij}F(|x_i - x_j|)\frac{x_i - x_j}{|x_i - x_j|}, \quad i = 1\ldots N. \tag{1}$$

We take $F(r)$ to be a repulsive-attractive force, i.e. positive for small r and negative for large r. The coefficients e_{ij} encode the connectivity between the particles x_i and x_j, so that $e_{ij} = 1$ if the vertices i and j interact and zero otherwise. We assume that $e_{ij} = e_{ji}$, so that the underlying graph is undirected, and in addition that $F(0) = 0$ so that the repulsion is "weak" at the origin. An example of a simple continuous force satisfying these conditions is

$$F(r) = \min\left(\kappa r, 1 - r\right), \quad \kappa > 0. \tag{2}$$

For such a force, it was shown in [11, 13, 23] that in the case of a fully connected graph, the equilibrium state can consist of several clusters. For two clusters, this equilibrium is simply

$$x_1\ldots x_n = -1/2, \qquad x_{n+1}\ldots x_N = 1/2, \qquad n = N/2. \tag{3}$$

2

An example of the evolution to such an equilibrium is shown in the top-left corner of Figure 1.

We refer to this system as a *compromise model* because the agents in each group prefer to remain a fixed distance away from all other agents; however, their attraction to the other group forces them to coexist at the same location with half of the total agents. Under all-to-all interactions, this particular steady state is stable provided that $0 < \kappa < 1$. The main issue that concerns us is how the lack of full connectivity affects the stability of this two-cluster solution. As a toy model in this direction, we consider an Erdős-Rényi random graph model $\mathcal{G}(N, p)$ for the interaction structure, for which

$$e_{ij} = \begin{cases} 1 \text{ with probability } p \\ 0 \text{ with probability } q = 1 - p, \end{cases} \tag{4}$$

and investigate how the stability of this equilibrium changes as p decreases. It is well-known that such a random graph $\mathcal{G}(N, p)$ undergoes a phase transition from connectedness to disconnectedness when when $p = \log N / N$. Roughly speaking, for the critical scaling $p = p_0 \frac{\log N}{N}$ and in the limit of large N, the graph is disconnected with high probability provided that $p_0 < 1$ and is connected when $p_0 > 1$. While the connectivity of the underlying graph is a necessary condition for the stability of the two-cluster steady state, as we shall show it is certainly not sufficient.

Figure 1 illustrates how the fully non-linear dynamics change as p is decreased. Note that the long-time dynamics and equlibrium remain the same as the fully connected case, even for relatively small p. When p becomes too small (around $p \approx 0.3$ for parameters of figure 1), the system undergoes a phase change and the two-cluster steady state appears to lose its stability. This observation naturally leads to the following question: *how small (resp. big) can we take p and still guarantee stability (resp. instability) of the two-cluster steady state?* To formulate this question more precisely, we focus on linear stability theory as $N \to \infty$. As a result of translation invariance of the steady state, the linearized problem always has at least one zero eigenvalue. We shall therefore call the two-cluster solution *stable asymptotically almost surely* if the probability that all remaining eigenvalues are negative tends to one as $N \to \infty$. Similarly, we call the two-cluster solution *unstable asymptotically almost surely* if the probability that one of the remaining eigenvalues is non-negative tends to one as $N \to \infty$.

The main results of the paper follow from a linear stability analysis of the equilibrium solution (3). We write the $N \times N$ adjacency matrix E of the random graph in terms of $n \times n$ blocks,

$$E = \begin{pmatrix} A & B \\ B^t & C \end{pmatrix},$$

to emphasize the distinction between the two groups. Evaluating the Jacobian of (1) at the equilibrium (3) leads to the linearized system for $\mathbf{\Phi} = (\Phi_1, \dots, \Phi_n)^t \in \mathbb{R}^n$ and $\mathbf{\Psi} = (\Psi_1, \dots, \Psi_n)^t \in \mathbb{R}^n$

$$\begin{cases} \lambda \Phi_i = \sum_{j=1}^{n} \kappa a_{ij} (\Phi_i - \Phi_j) + \sum_{j=1}^{n} b_{ij} (\Psi_j - \Phi_i) \\ \lambda \Psi_i = \sum_{j=1}^{n} b_{ji} (\Psi_j - \Psi_i) + \sum_{j=1}^{n} \kappa c_{ij} (\Psi_i - \Psi_j) \end{cases}. \tag{5}$$

We can re-write this in the form

$$\lambda \begin{pmatrix} \mathbf{\Phi} \\ \mathbf{\Psi} \end{pmatrix} = L \begin{pmatrix} \mathbf{\Phi} \\ \mathbf{\Psi} \end{pmatrix}$$

where L denotes the $N \times N$ matrix

$$L = \kappa \begin{pmatrix} D_A - A & 0 \\ 0 & D_C - C \end{pmatrix} - \begin{pmatrix} D_B & -B \\ -B^t & D_{B^t} \end{pmatrix} := \kappa L_1 - L_2. \tag{6}$$

3

Here, D_A (respectively, D_B, D_C) denotes the $n \times n$ diagonal matrix formed from row sums of A, i.e.

$$(D_A)_{ii} = \sum_{j=1}^{n} a_{ij}.$$

The matrices L_1 and L_2 therefore equal the graph Laplacians formed from two subgraphs of E; the first subgraph contains only those edges that do not connect the two groups and the second subgraph contains only those edges that do connect the two groups. We can therefore interpret the original eigenvalue problem for L as a difference of positive semi-definite Laplacian matrices.

Our main goal is to characterize the stability regime and transition to instability for the two-cluster steady state. Our approach is two-fold. On one hand, we will apply rigorous estimates using random matrix theory to derive rigorous bounds on the transition regime. On the other hand, we will use some heuristical arguments, combined with formal asymptotics and extensive numerics to derive sharp estimates for the transition regime. We summarize our rigorous result as follows.

Theorem 1.1. (Rigorous bounds for stability) *Define F by (2) for $0 < \kappa < 1$ and let e_{ij} denote the $N \times N$ adjacency matrix of an Erdős-Rényi random graph $\mathcal{G}(N, p)$. If*

$$p \geq \epsilon \frac{\log^{3/2} N}{N} \tag{7}$$

for some $\epsilon > 0$ then the steady state (3) is stable asymptotically almost surely. If

$$p \leq 2(1 - \epsilon) \frac{\log(N)}{N} \tag{8}$$

for some $\epsilon > 0$ then the steady state (3) is unstable asymptotically almost surely.

Using heuristic methods, we make the following conjecture based on asympotics of the transition regime:

Conjecture 1.1. (Heuristic asymptotics of the transition regime) *Define F by (2) for $0 < \kappa < 1$ and let e_{ij} denote the $N \times N$ adjacency matrix of an Erdős-Rényi random graph $\mathcal{G}(N, p)$. There exists a constant p_{0c} (independent of N) with the following property. If*

$$p \geq (p_{0c} + \epsilon) \frac{\log N}{N}$$

for some $\epsilon > 0$ then the steady state (3) is stable asymptotically almost surely. If

$$p \leq (p_{0c} - \epsilon) \frac{\log N}{N}$$

then the steady state (3) is unstable asymptotically almost surely.

The asymptotic calculations in section 3, suggest that in fact

$$p_{0c} = \frac{2}{2 - \kappa^{-\kappa/(\kappa+1)} (1 + \kappa)}. \tag{9}$$

We compare this formula and the simplifying assumptions used to derive it with numerical experiments in section 3.2.

4

1.1 Connection to Consensus Algorithms

Equation (6) generalizes the classical consensus algorithm on an undirected graph. Consensus occurs when all the agents approach a common value in the long time limit. This problem has attracted scholars from many areas such as physics, control theory and biology. Specifically, the case of single-integrator, linear dynamics taking the form

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \sum_{j \neq i}^{N} e_{ij}(x_i - x_j), \quad x_i \in \mathbb{R}^m, \tag{10}$$

is especially well-studied in control theory. Here $E = \{e_{ij}\}_{i,j=1}^{N}$ denotes the weighted adjacency matrix of the underlying graph $\mathcal{G}(N)$: $e_{ij} > 0$ if two nodes are connected and $e_{ij} = 0$ otherwise. The above equation can be written in matrix form,

$$\dot{\mathbf{x}} = -[L \otimes \mathrm{Id}]\mathbf{x},$$

where $\mathbf{x} = (x_1, \ldots, x_N)$ is the vector of all nodes and L is the Laplacian matrix. While in general the consensus problem is asymmetric, we consider the symmetric case $e_{ij} = e_{ji}$ to emphasize the similarity with our compromise problem.

Definition 1.1. *The system* (10) *reaches consensus if, for all initial conditions $x_i(0)$ and all $1 \leq i, j \leq N$, it holds that $\|x_i(t) - x_j(t)\| \to 0$ as $t \to \infty$.*

As shown in Ren [22, 19, 21], Olfati-Saber [18, 17], Moreau [15, 16], and Cao [4], the role of connectivity and stability are highly relevant for cooperative control algorithms. For directed graphs with a fixed topology (i.e. the e_{ij} are constant in time), the main theorem regarding consensus is the following [20]:

Theorem 1.2. *The system* (10) *reaches consensus if and only if the directed graph $\mathcal{G}(N)$ has a directed spanning tree. In this case, $x_i(t) \to \sum_{i=1}^{N} \nu_i x_i(0)$ as $t \to \infty$, where $\nu = [\nu_1, \ldots, \nu_n]^T \geq 0$, $\mathbf{1}^T \nu = 1$, and $L^T \nu = 0$.*

In the case of a fully connected, undirected graph this reduces to the well-known fact that the consensus is the center of mass of the initial data. The above theorem also has a natural simplification in the case of general undirected graphs:

Corollary 1.3. *For an undirected graph $\mathcal{G}(N)$ ($e_{ij} = e_{ji}$), the system* (10) *reaches consensus if and only if $\mathcal{G}(N)$ is connected.*

We now connect these results to well-known results for graph connectivity of Erdős-Rényi graphs in the large N limit [3, 10, 1]. These results are not new, but the connection between the two problems has not been emphasized in the literature. We state the main result in this direction to have a point of comparison for our results on the compromise model.

Theorem 1.4. *Consider equation* (10) *in the case of an Erdős-Rényi graph $\mathcal{G}(N, p)$. The graph is constructed by choosing $e_{ij} = e_{ji}$ equal to zero or one with some fixed probability $p \in (0, 1)$. If for some $\epsilon > 0$*

$$p \geq (1 + \epsilon)\frac{\log N}{N}$$

then the problem (10) *reaches consensus asymptotically almost surely. If for some* $\epsilon > 0$

$$p \leq (1 - \epsilon)\frac{\log N}{N}$$

then the problem (10) *fails to reach consensus asymptotically almost surely.*

In the body of this paper we find that the critical probability for convergence to the compromise solution, in the compromise model, differs from the critical probability for graph connectivity; the threshold has the same $\log N/N$ scaling but with a different coefficient.

1.2 Preliminary Material

Before we address our main task, we first introduce our notation along with the background material from linear algebra, probability theory and random matrix theory that we shall use. Capital roman letters such as A, B, C refer to matrices while the corresponding lower-case letters a_{ij}, b_{ij}, c_{ij} denote their respective entries. We reserve Id for the identity matrix and $\mathbf{1} = (1, \ldots, 1)^t$ for the constant vector. The size of both will be clear from context. For an $n \times n$ symmertic matrix A, we let $\lambda_i(A)$ denote the its ith eigenvalue sorted in decreasing order. In other words, we have

$$\lambda_1(A) \geq \lambda_2(A) \cdots \geq \lambda_n(A), \tag{11}$$

where each eigenvalue apprears according to its algebraic multiplicity. Under this convention, a standard identity from linear algebra provides a useful characterization of the ith eigenvalue, i.e. the Courant-Fischer formula. This formula yields a system of eigenvalue inequalities, Weyl's inequalities, that will also prove useful.

Lemma 1.5. *(Courant-Fischer Max-Min Formula) Let A denote an $n \times n$ symmetric matrix and V a subspace of $\mathbb{R}^n$. Then*

$$\lambda_i(A) = \max_{V:\dim(V)=i} \quad \min_{\mathbf{v}\in V:||v||_2=1} \langle \mathbf{v}, A\mathbf{v} \rangle \tag{12}$$

Lemma 1.6. *(Weyl's Inequalities) Let A and B denote symmetric $n \times n$ matrices. Then for any i, j such that $1 \leq i, j, i + j - 1 \leq n$,*

$$\lambda_{i+j-1}(A + B) \leq \lambda_i(A) + \lambda_j(B). \tag{13}$$

Given a sequence of measurable events $\{W_n\}_{n=1}^{\infty}$, each of which lies in some (possibly different) probability space, we say that W_n holds asymptotically almost surely (a.a.s) if

$$\mathbb{P}(W_n) \to 1 \quad \text{as} \quad n \to \infty.$$

Here and in what follows, $\mathbb{P}$ always denotes the measure on the probability space in which the relevant event lies. We denote by X_p a Bernoulli random variable with parameter p, i.e.

$$\mathbb{P}(X_p = 1) = p \quad \text{and} \quad \mathbb{P}(X_p = 0) = q := 1 - p, \tag{14}$$

whereas $\tilde{X}_p := X_p - p$ will denote a mean-zero Bernoulli random variable. We use $\mathbb{E}(X)$ to denote the mean or expectation of the random variable X, and the notation $X =_d Y$ to signify that the random variables X and Y have the same distribution.

Our arguments require probabilistic estimates of the form $\mathbb{P}(|X| \geq \lambda)$, where X will represent either a weighted sum of $\tilde{X}_p$ variables or the operator norm of a symmetric random matrix. For the first case it suffices to simply recall a variant of the well-known Chernoff bound (c.f. lemma 1.7). In lemma 1.8 we prove an operator norm estimate using a standard technique from random matrix theory. The proof of lemma 1.8 closely mirrors theorem 1.4 in [27]. Theorem 1.4 in [27] is stated without proof, so to obtain an explicit result we essentially reproduce the arguments from [27] while taking care to keep the estimates as concrete as we will need.

Lemma 1.7. *(Chernoff Bound) Let $X_1, \ldots, X_m$ denote discrete, independent random variables satisfying $\mathbb{E}(X_i) = 0$ and $|X_i| \leq 1$. If $\mathbb{E}(X_i^2) = \sigma_i^2$ and $\sigma^2 = \sum \sigma_i^2$, then for any $0 \leq \lambda \leq 2\sigma$*

$$\mathbb{P}\left(|\sum_{i=1}^{m} X_i| \geq \lambda\sigma\right) \leq 2\mathrm{e}^{-\lambda^2/4}. \tag{15}$$

Lemma 1.8. *Let $A = \{a_{ij}\}_{i,j=1}^{n}$ denote a symmetric random matrix with independent upper-triangular entries. Let $p_0, q_0 > 0$ denote arbitrary constants independent of n. If*

$$\text{(i) } \mathbb{E}(a_{ij}) = 0 \quad \text{(ii) } \mathbb{E}(a_{ij}^2) \leq 1 \quad and \quad \text{(iii) } |a_{ij}|^2 \leq \frac{n}{q_0 \log^{3/2} n} \quad a.s. \ ,$$

then

$$\mathbb{P}\left(||A||_2 \geq 4\sqrt{p_0 n \log^{3/2} n}\right) \to 0 \tag{16}$$

faster than n^{-M} for any $M > 0$ and all n sufficiently large.

Proof. We follow Wigner's trace method [28] as outlined in [27, 2, 24]. For any integer $k > 0$ and any $\lambda > 0$, the fact that $||A||_2^{2k} \leq \text{trace}(A^{2k})$ and Markov's inequality combine to show

$$\mathbb{P}(||A||_2 \geq \lambda) = \mathbb{P}(||A||_2^{2k} \geq \lambda^{2k}) \leq \mathbb{P}(\text{trace}(A^{2k}) \geq \lambda^{2k}) \leq \lambda^{-2k}\mathbb{E}(\text{trace}(A^{2k})).$$

Substituting the hypotheses $(\text{i} - \text{iii})$ into the estimate from [27] yields

$$\mathbb{E}(\text{trace}(A^{2k})) \leq \sum_{p=1}^{k+1} \left(\frac{n}{q_0 \log^{3/2} n}\right)^{k-p+1} n^p \binom{2k}{2p-2} 2^{4k-2p+3} p^{2k-2p+2}(2k-2p+4)^{k-p+1}.$$

After performing the change of variables $l = k + 1 - p$, this estimate reads

$$\mathbb{E}(\text{trace}(A^{2k})) \leq \sum_{l=0}^{k} \left(\frac{n}{q_0 \log^{3/2} n}\right)^{l} n^{k+1-l} \binom{2k}{2l} 2^{2l+2k+1}(k+1-l)^{2l}(2l+2)^{l}$$

$$= 2^{2k+1} n^{k+1} \sum_{l=0}^{k} \left(\frac{8(l+1)}{q_0 \log^{3/2} n}\right)^{l} (k+1-l)^{2l} \binom{2k}{2l}.$$

To estimate the sum, let $C(n,k) := 8(k+1)/(q_0 \log^{3/2} n)$ and $f(l) := [\sqrt{C}(k+1-l)]^{2l}$. Elementary calculus then shows that f attains its maximum when

$$\sqrt{C}(k+1-l)[\log \sqrt{C}(k+1-l) + 1] = \sqrt{C}(k+1).$$

7

This has a unique solution in terms of special functions,

$$\sqrt{C}(k + 1 - l) = e^{W[e\sqrt{C}(k+1)]-1},$$

where $W[x]$ denotes the product-log function. That is, $W[x]$ denotes the unique (for $x > 0$) solution to $W[x]e^{W[x]} = x$. As a consequence,

$$\mathbb{E}(\text{trace}(A^{2k})) \le 2^{2k+1}n^{k+1}\left(\frac{\sqrt{C}(k+1)}{W[e\sqrt{C}(k+1)]}\right)^{2(k+1)(W-1)/W}\sum_{l=0}^{k}\binom{2k}{2l}$$

$$= 4^{2k}n^{k+1}\left(\frac{\sqrt{C}(k+1)}{W[e\sqrt{C}(k+1)]}\right)^{2(k+1)(W-1)/W}.$$

Substituting this expression into the estimate for $\mathbb{P}(||A||_2 \ge \lambda)$ and simplifying demonstrates

$$\mathbb{P}(||A||_2 \ge \lambda) \le \lambda^{-2k}4^{2k}n^{k+1}\left[\frac{C^{1-1/W}(k+1)^{2-2/W}}{W^{2-2/W}}\right]^{k+1}.$$

Now set $\lambda = 4\sqrt{p_0 n \log^{3/2} n}$ and $k = \lfloor \log n \rfloor$. Then

$$\mathbb{P}\left(||A||_2 \ge 4\sqrt{p_0 n \log^{3/2} n}\right) \le n\log^{3/2} n\left[\left(\frac{8}{q_0 W^2}\right)^{1-1/W}\frac{(k+1)^{3(1-1/W)}}{p_0 \log^{3-3W/2} n}\right]^{k+1}.$$

With this choice of $k + 1$, it then follows that

$$\mathbb{P}\left(||A||_2 \ge 4\sqrt{p_0 n \log^{3/2} n}\right) \le n\log^{3/2} n\left[\left(\frac{8}{q_0 W^2}\right)^{1-1/W}\frac{1}{p_0 \log^{3/2W} n}\right]^{k+1}$$

and that $W[e\sqrt{C}(k+1)] \to \infty$ as well. Thus

$$\left[\left(\frac{8}{q_0 W^2}\right)^{1-1/W}\frac{1}{p_0 \log^{3/2W} n}\right] \to 0,$$

so that

$$\mathbb{P}\left(||A||_2 \ge 4\sqrt{p_0 n \log^{\beta} n}\right) \to 0$$

faster than n^{-M} for any $M > 0$ and all n sufficiently large. $\qquad\square$

We also need to establish the connectivity properties of a slight modification of the standard Erdős-Rényi random graph $\mathcal{G}(N,p)$ on N vertices. Given a parameter $p \in (0,1)$, we construct an undirected, bipartite graph on N vertices by assigning independent edges $e_{ij} = e_{ji} =_d X_p$ whenever $1 \le i \le n < j \le N$ and forcing $e_{ij} = e_{ji} = 0$ otherwise. We let $\mathcal{K}(N)$ denote the set of all possible bipartite graphs constructed in this manner. We write $(V, E) \in \mathcal{K}(N, p)$ or $(V, E) \in \mathcal{G}(N, p)$ to specify the parameter p when referring to a randomly sampled graph of either type. Here $V = \{v_1, \ldots, v_N\}$ is a set of N vertex labels and $E = \{e_{ij}\}_{ij=1}^{N}$ denotes the corresponding

$N \times N$ adjacency matrix. We let $\mathcal{K}_c(N) \subset \mathcal{K}(N)$ denote the subset of connected graphs and $\mathcal{K}_d(N) = \mathcal{K}(N) \setminus \mathcal{K}_c(N)$ the disconnected graphs.

By slightly modifying standard proofs from the literature regarding Erdős-Rényi graphs [3] we can readily prove the following lemma. While more sophisticated and more general results exist concerning random bipartite graphs [3], we include a proof below for the sake of completeness.

Lemma 1.9. *Given $p \in (0,1)$ let $(V, E) \in \mathcal{K}(N, p)$ denote a corresponding random graph. If*

$$p \geq 2(1 + \epsilon) \frac{\log N}{N} \tag{17}$$

for some constant $\epsilon > 0$ then (V, E) is connected with probability at least $1 - (N/2)^{-\epsilon/2}$, i.e. asymptotically almost surely. Conversely, if

$$p \leq 2(1 - \epsilon) \frac{\log N}{N} \tag{18}$$

for some constant $\epsilon > 0$ then (V, E) contains isolated vertices with probability at least $1 - \mathrm{e}^{-(N/2)^{\epsilon/2}}$.

Proof. For fixed $n := N/2$, let $W_{n,k}$ denote the event that there exist k vertices $\{v_{i_1}, \ldots, v_{i_k}\}$ with no edges connecting $\{v_{i_1}, \ldots, v_{i_k}\}$ and $V \setminus \{v_{i_1}, \ldots, v_{i_k}\}$. Note that

$$\mathbb{P}\left((V, E) \in \mathcal{K}_d(N)\right) = \mathbb{P}\left(\cup_{k=1}^n W_{n,k}\right) \leq \sum_{k=1}^n \mathbb{P}(W_{n,k}).$$

To estimate $\mathbb{P}(W_{n,k})$, for a fixed $\{v_{i_1}, \ldots, v_{i_k}\} \in W_{n,k}$ let j denote the number of indices i_l with $i_l \leq n$ and $k - j$ the number of indices with $i_l > n$. By independence of the edges e_{ij}, a straightforward computation shows that

$$\mathbb{P}\left(\{v_{i_1}, \ldots, v_{i_k}\} \in W_{n,k}\right) = (1 - p)^{(n-j)(k-j) + j(n+j-k)}.$$

Summing over the total number of possible choices for $\{v_{i_1}, \ldots, v_{i_k}\}$ yields as a consequence that

$$\mathbb{P}\left(W_{n,k}\right) \leq \sum_{j=0}^k \binom{n}{j}\binom{n}{k-j}(1-p)^{(n-j)(k-j)+j(n+j-k)} \leq (1-p)^{(k/2)(N-k)} \sum_{j=0}^k \binom{n}{j}\binom{n}{k-j}$$

$$= \binom{N}{k}(1-p)^{(k/2)(N-k)}.$$

The estimates $\binom{N}{k} \leq (N\mathrm{e}/k)^k$ and $(1-p) \leq e^{-p}$ therefore give

$$\mathbb{P}\left((V, E) \in \mathcal{K}_d(N)\right) \leq \sum_{k=1}^n \binom{N}{k}(1-p)^{(k/2)(N-k)} \leq \sum_{k=1}^n \exp(k(\log(N\mathrm{e}/k) - p(n - k/2))).$$

Without loss of generality, assume $0 < \epsilon < 1$. Then $\log(N\mathrm{e}/k) - p(n - k/2) \leq -\epsilon \log n + (1 + \epsilon)(\log n)/N + \log 2\mathrm{e}$. As $\epsilon > 0$, for all n sufficiently large it follows that

$$\mathbb{P}\left((V, E) \in \mathcal{K}_d(N)\right) \leq \sum_{k=1}^n n^{-k\epsilon/2} \leq n^{-\epsilon/2}\frac{1}{1 - n^{-\epsilon/2}} \to 0$$

9

as desired. To show the second statement, it suffices to show that one of the vertices $\{v_1, \ldots, v_n\}$ becomes isolated. Let $R_i = \{e_{i,n+1} = \ldots = e_{i,N} = 0\}$ denote the event that the vertex v_i is isolated. As $R_1 =_d \cdots =_d R_n$ and these events are independent for $1 \le i \le n$, it follows that

$$\mathbb{P}\left(\cup_{i=1}^n R_i\right) = 1 - \mathbb{P}(R_1^c)^n = 1 - [1 - \mathbb{P}(R_1)]^n = 1 - [1 - (1-p)^n]^n.$$

When $p \le (1-\epsilon)n^{-1}\log n$ it follows that $(1-p)^n \ge n^{\epsilon/2-1}$ for n sufficiently large. Thus $[1 - \mathbb{P}(R_1)]^n \le \exp(-n\mathbb{P}(R_1)) \le \exp(-n^{\epsilon/2})$, so that

$$\mathbb{P}\left(\cup_{i=1}^n R_i\right) \ge 1 - \exp(-n^{\epsilon/2}) \to 1$$

as desired. $\qquad\square$

2 Rigorous Estimates for Stability

We may now turn to our main task, i.e. the proof of theorem 1.1. Recall the eigenvalue problem reads

$$\lambda \begin{pmatrix} \mathbf{\Phi} \\ \mathbf{\Psi} \end{pmatrix} = L \begin{pmatrix} \mathbf{\Phi} \\ \mathbf{\Psi} \end{pmatrix} \tag{19}$$

Note that $\mathbf{\Phi} = \mathbf{\Psi} = \mathbf{1}$ always defines an eigenvector of (19) with eigenvalue zero. We therefore call a "two-group" solution stable when the second largest eigenvalue, $\lambda_2(L)$, of the system (19) is strictly negative. In crude analogy to the law of large numbers, we expect that L should concentrate around its mean, $L \approx \mathbb{E}(L)$, where the error becomes negligible in the limit of infinite system size. Taking the expectation $\mathbb{E}$ of (19) entriwise gives

$$\mathbb{E}(L) = pL_{\text{comp}},$$

where L_{comp} denotes the stability matrix when the underlying graph is complete ($e_{ij} \equiv 1$ in (19)). Thus

$$L = pL_{\text{comp}} + R, \tag{20}$$

where $R := L - \mathbb{E}(L)$ is a $N \times N$, symmetric matrix that has mean-zero entries. From Weyl's inequalities (13), we have

$$\lambda_2(L) \le p\lambda_2(L_{\text{comp}}) + ||R||_2.$$

Heuristically then, if $\lambda_2(L_{\text{comp}}) < 0$ and the error $||R||_2$ is asymptotically negligible the "two-group" solution is stable asymptotically almost surely. Using the estimates from the previous section, we show this is indeed the case provided p does not vanish too rapidly.

To realize this program, we first consider stability of the complete graph. A direct computation reveals that

$$L_{\text{comp}} := n(\kappa - 1)\text{Id} + \begin{pmatrix} -\kappa\mathbf{1} \otimes \mathbf{1} & \mathbf{1} \otimes \mathbf{1} \\ \mathbf{1} \otimes \mathbf{1} & -\kappa\mathbf{1} \otimes \mathbf{1} \end{pmatrix}. \tag{21}$$

Let $\mathbf{\Psi} = \mathbf{0}$ and $\mathbf{\Phi} = \mathbf{e}_1 - \mathbf{e}_j$, where $\mathbf{e}_j \in \mathbb{R}^n$ denotes any of the $(n-1)$ remaining standard basis vectors. Setting $\mathbf{v} = (\mathbf{\Phi}, \mathbf{\Psi})^t$ or $\mathbf{v} = (\mathbf{\Psi}, \mathbf{\Phi})^t$ and performing a straightforward computation shows that

$$L_{\text{comp}}\mathbf{v} = n(\kappa - 1)\mathbf{v},$$

so that $n(\kappa - 1)$ is an eigenvalue with multiplicity $N - 2$ by linear independence. The choice $\mathbf{v} = (\mathbf{1}, -\mathbf{1})^t$ yields an eigenvalue of $-N$. Stability of the complete graph therefore demands $\kappa < 1$, in which case

$$\lambda_1(L_{\text{comp}}) = 0, \quad \lambda_2(L_{\text{comp}}) = \cdots = \lambda_{N-1}(L_{\text{comp}}) = n(\kappa - 1), \quad \lambda_N(L_{\text{comp}}) = -N.$$

Turning now the random graph case, we wish to find how small we can take p while not losing control of the error R in (20). To this end we decompose the error as

$$R = D + \tilde{E}, \tag{22}$$

where D denotes a diagonal matrix and $\tilde{E}$ denotes a symmetric matrix. The diagonal matrix D has entries

$$\begin{aligned} d_{ii} &= \kappa \sum_{j=1}^n \tilde{a}_{ij} - \sum_{j=1}^n \tilde{b}_{ij} \\ d_{i+n,i+n} &= \kappa \sum_{j=1}^n \tilde{c}_{ij} - \sum_{j=1}^n \tilde{b}_{ji} \end{aligned} \quad \text{for } 1 \le i \le n. \tag{23}$$

The matrix D therefore has entries comprised of sums of independent random variables, although the entries of D depend on one another and depend on the entries of $\tilde{E}$ as well. The matrix $\tilde{E}$ has the form

$$\tilde{E} = \begin{pmatrix} -\kappa\tilde{A} & \tilde{B} \\ \tilde{B}^t & -\kappa\tilde{C} \end{pmatrix} \tag{24}$$

The entries of $\tilde{A}$ and $\tilde{C}$ satisfy $\tilde{a}_{ij} = \tilde{a}_{ji} =_d \tilde{X}_p$ and $\tilde{c}_{ij} = \tilde{c}_{ji} =_d \tilde{X}_p$, and are independent on the upper triangular portion. The entries $\tilde{b}_{ij} =_d \tilde{X}_p$ of B are independent across the full matrix. Estimating $||R||_2$ therefore involves estimating the operator norm of two types of matrices: a diagonal matrix with entries that are sums of independent $\tilde{X}_p$ variables and a symmetric matrix with independent $\tilde{X}_p$ variables on the upper triangle. Lemma 1.7 allows us to handle the former while lemma 1.8 suffices to handle latter.

We first estimate $||D||_2$. As D is diagonal, this simply equals the entry with maximum absolute value. We simply apply lemma 1.7 directly to the N independent random variables that constitute a given diagonal entry. A direct calculation of the relevant quantities in the statement of the lemma shows that

$$\sigma^2 := \kappa^2 \sum_{j=1}^n \mathbb{E}(\tilde{a}_{ij}^2) + \sum_{j=1}^n \mathbb{E}(\tilde{b}_{ij}^2) = n(1 + \kappa^2)p(1 - p) = n(1 + \kappa^2)pq$$

Thus whenever $\lambda^2/4 \le n(1 + \kappa^2)pq$, it follows from the lemma that

$$\mathbb{P}\left(|d_{ii}| \ge \lambda\sigma\right) \le 2e^{-\lambda^2/4}.$$

By the union bound, this implies

$$\mathbb{P}\left(||D||_2 \ge \lambda\sigma\right) \le 4ne^{-\lambda^2/4}.$$

To ensure the right hand side still tends to zero as $n \to \infty$, we take $\lambda^2/4 = (1 + \epsilon)\log n$ for some $\epsilon > 0$. In turn, this places the requirement on p that

$$pq > \frac{\log n}{4(1 + \kappa^2)n}. \tag{25}$$

Substituting these choices into the previous bound demonstrates that

$$\mathbb{P}\left(||D||_2 \geq 2\sqrt{(1+\epsilon)(1+\kappa^2)pqn\log n}\right) \leq 4n^{-\epsilon} \qquad (26)$$

provided that (25) holds.

It remains to estimate $||\tilde{E}||_2$. For this task we apply lemma 1.8 from the preliminary material. We let M denote the $N \times N$ symmetric matrix

$$M = \frac{\tilde{E}}{\sqrt{p}},$$

so that its entries m_{ij} satisfy

$$\mathbb{E}(m_{ij}) = 0 \qquad \mathbb{E}(m_{ij}^2) \leq \max\{\kappa^2, 1\} \qquad |m_{ij}|^2 \leq \max\{\kappa^2, 1\}\frac{1}{p}.$$

As $0 < \kappa < 1$, we satisfy the hypotheses of lemma 1.8 provided we place one further restriction on p, i.e. that

$$Np \geq q_0^2 \log^{3/2} N$$

where $q_0 > 0$ denotes an arbitrary, fixed constant. We then have that there exists an N-independent constant C so that the estimate

$$\mathbb{P}\left(||\tilde{E}||_2 \geq 4p_0\sqrt{Np\log^{3/2} N}\right) \leq CN^{-2} \qquad (27)$$

holds for all $p_0, q_0 > 0$ and N sufficiently large. Substituting the previous estimates (26),(27) into the bound for $||R||_2$, we find

$$\lambda_2(L) \leq Np\frac{(\kappa-1)}{2} + 4p_0\sqrt{Np}\log^{3/4} N + 2\sqrt{(1+\epsilon)(1+\kappa^2)pqn\log n}$$

holds with probability at least

$$1 - 2\max\left\{4(N/2)^{-\epsilon}, CN^{-2}\right\}.$$

Taking $\epsilon = 3$ and $p_0 = (1-\kappa)q_0\delta/8$ for some $\delta > 0$, for instance, we see that

$$\lambda_2(L) \leq -\sqrt{Np}\frac{(1-\kappa)q_0}{2}(1-\delta)\log^{3/4} N + 4\sqrt{np(1+\kappa^2)\log n}$$

with probability at least

$$1 - CN^{-2}.$$

As a consequence, we find $\lambda_2(L) < 0$ asymptotically almost surely. The following theorem encapsulates the preceeding discussion.

Theorem 2.1. *Let $0 < \kappa < 1$ and $p \in (0,1)$ satisfy*

$$p \geq q_0^2 \frac{\log^{3/2} N}{N} \qquad (28)$$

for any $q_0 > 0$. Then for any $\delta > 0$

$$\lambda_2(L) \leq -\left(\frac{(1-\kappa)(1-\delta)}{2}q_0 + O\left(\log^{-1/4} N\right)\right)\sqrt{Np}\log^{3/4} N$$

asymptotically almost surely.

To establish the converse result, we appeal to the results regarding connectivity of random graphs in the preliminary material. Suppose that

$$p \leq 2(1 - \epsilon)\frac{\log N}{N}$$

for some $\epsilon > 0$. Then by lemma 1.9, a graph $(V, E) \in \mathcal{K}(N, p)$ contains isolated vertices with probability at least $1 - \exp(-(N/2)^{\epsilon/2})$. For any such graph, let j denote the index of an isolated vertex and set

$$\mathbf{v} = c_1 \mathbf{1} + c_2 \mathbf{e}_j$$

for any choice of coefficients c_1, c_2 such that $\|\mathbf{v}\|_2 = 1$. Then $\langle \mathbf{e}_j, L_2 \mathbf{e}_j \rangle = 0$, so that

$$\langle \mathbf{v}, L\mathbf{v} \rangle = \kappa c_2^2 \langle \mathbf{e}_j, L_1 \mathbf{e}_j \rangle \geq 0$$

provided $\kappa \geq 0$. By the Courant-Fischer formula (12), this implies $\lambda_2(L) \geq 0$ for any such graph. Therefore for any choice of $\kappa_N \geq 0$ we find $\lambda_2(L) \geq 0$ asymptotically almost surely. We summarize this fact in the following theorem.

Theorem 2.2. *Let $\kappa > 0$. If for some $\epsilon \geq 0$*

$$p \leq 2(1 - \epsilon)\frac{\log N}{N},$$

then $\lambda_2(L \geq 0)$ with probability at least $1 - \exp(-(N/2)^{\epsilon/2})$. In particular, the steady-state (3) is unstable asymptotically almost surely.

Remark 2.3. *From a modelling perspective, in the system of ODEs (1) only $\kappa > 0$ makes sense. However, the preceeding arguments hold if $\kappa \leq 0$ as well. When $\kappa < 0$, the statement "$\lambda_2(L) < 0$ asymptotically almost surely" proves exactly equivalent to the connectedness of the full Erdős-Rényi random graph $\mathcal{G}(N, p)$ on N vertices. In this case it is well-known that the sharp threshold is $p = \log N / N$, so that at $\kappa = 0$ a "discontinuity" occurs in the sharp threshold for stability.*

3 Estimates for critical probability

The main goal of this section is to derive formula (9). We first present formal asymptotics and then present numerical simulations supporting these asymptotics.

3.1 Formal Asymptotics

The formal asymptotics rely on the following key lemma that may be of independent interest.

Lemma 3.1. *Let P_λ denote the Poisson distribution with parameter λ. Define*

$$X := P_\lambda; \quad Y := -\kappa P_\lambda \text{ and } Z := X + Y, \quad \text{where } 0 < \kappa < 1. \tag{29}$$

Suppose that

$$\lambda = \lambda_0 \log N \quad \text{with} \quad N \gg 1, \; \lambda_0 = O(1) \tag{30}$$

and let $Z_1, \ldots Z_N$ be N independent realizations of the random variable Z. Define S and x_0 through the equations

$$\psi(x_0, S, \lambda_0) := -2\lambda_0 + S\left(\log\frac{\lambda_0}{S} + 1\right) + \frac{S - x_0}{\kappa}\left(\log\frac{\lambda_0\kappa}{S - x_0} + 1\right) = -1, \tag{31}$$

$$\frac{\partial\psi}{\partial S}(x_0, S, \lambda_0) = \log\frac{\lambda_0}{S} + \frac{1}{\kappa}\log\frac{\kappa\lambda_0}{S - x_0} = 0. \tag{32}$$

Then in the limit $N \gg 1$, we have

$$\mathbb{E}(\min(Z_1, \ldots Z_N)) \sim x_0\log N, \quad N \gg 1. \tag{33}$$

As a consequence, $\mathbb{E}(\min(Z_1, \ldots Z_N)) \sim 0$ if and only if

$$\lambda_0 = \frac{1}{2 - \kappa^{-\kappa/(\kappa+1)}(1 + \kappa)}. \tag{34}$$

Remark 3.2. *The asymptotic statement in* (33), *that*

$$\frac{\mathbb{E}(\min(Z_1, \ldots Z_N))}{x_0\log N} \to 1 \quad as \quad N \to \infty,$$

follows as consequence of the more precise statement

$$\mathbb{P}(\min(Z_1, \ldots Z_N) > x_0\log N) \sim \exp\left\{-CN^{\psi(x_0, S, \lambda_0)+1}\right\}$$

for some positive constant $C = O(\log N)$.

We first use the lemma to derive (9), then prove lemma 3.1 on page 16. As in Section 2, we decompose L into three parts

$$L = pL_{comp} + D + \tilde{E}, \tag{35}$$

where L_{comp} is defined in (21) and D and $\tilde{E}$ are defined in (22, 23, 24). The matrix $\tilde{E}$ is a symmetric random matrix whose entries have mean zero and D is a diagonal matrix whose entries are minus the row sums $\tilde{E}$. As noted in Section 2, the matrix L_{comp} has a zero eigenvalue with algebraic multiplicity one, an eigenvalue $\lambda = -N(1 - \kappa)/2$ with algebraic multiplicity $2N - 2$ and an eigenvalue $\lambda = -N < -N(1 - \kappa)/2$ with algebraic multiplicity one. With this in mind, and in the spirit of asymptotics, we formally replace

$$L_{comp} \sim -\frac{N(1 - \kappa)}{2}\text{Id}. \tag{36}$$

To gain further insight on the remaining terms, we first consider the simpler case where $e_{ij} =_d \mathcal{N}(p, pq)$ are normally distributed with the same mean and variance as the Bernoulli case. The diagonal entires of D therefore have distribution

$$d_{ii} =_d \sqrt{1 + \kappa^2}\sqrt{pqN/2}\mathcal{N}(0, 1). \tag{37}$$

If $\mathcal{N}_1, \ldots, \mathcal{N}_N$ denote N independent unit Gaussians, an argument similar to the proof of lemma 3.1 demonstrates that $\mathbb{E}(\max(\mathcal{N}_1, \ldots, \mathcal{N}_N)) \sim \sqrt{2\log N}$. Due to the symmetry of $\tilde{E}$, the entries of D are not independent. However, if we formally assume that D has independent entries this would

imply that $\mathbb{E}\left(\max(d_{ii})\right) \sim \sqrt{(1 + \kappa^2)\, pqN \log N}$. Finally, $\tilde{E}/\sqrt{pq}$ is a symmetric random matrix whose entries have mean zero and variance bounded by one. In the case of normally distributed weighted edges e_{ij}, the entries of $\tilde{E}/\sqrt{pq} =_d \mathcal{N}(0,1)$ have uniformly bounded fourth moments. From the Bai-Yin theorem (c.f. [2]), it follows that $||\tilde{E}||_2 = O(\sqrt{pqN})$ asymptotically almost surely. Thus $||\tilde{E}||_2$ is $O(\sqrt{\log N})$ smaller than the maximum entry of D if N is large. We therefore formally discard $\tilde{E}$ in (35) to obtain

$$\lambda_2(L) \sim -\frac{Np(1 - \kappa)}{2} + \sqrt{(1 + \kappa^2)\, pqN \log N}.$$

The two terms in balance precisely when p has the critical scaling $p = O(\log N/N)$. Substituting $p = p_0 \log N/N$ and setting $\lambda_2(L) = 0$ yields a critical threshold

$$p_{0c} = \frac{4\left(\kappa^2 + 1\right)}{(1 - \kappa)^2}, \tag{38}$$

so that the eigenvalue λ_2 is negative for $p_0 > p_{0c}$ and is positive for $p_0 < p_{0c}$ asymptotically almost surely.

We now turn our attention to the original problem. From (23) we have

$$\begin{aligned} d_{ii} &= (1 - \kappa)\tfrac{N}{2}p + \sum_{j=1}^n \kappa a_{ij} - \sum_{j=1}^n b_{ij} \\ d_{i+n,i+n} &= (1 - \kappa)\tfrac{N}{2}p + \sum_{j=1}^n \kappa c_{ij} - \sum_{j=1}^n b_{ji} \end{aligned} \quad \text{for } 1 \le i \le n \tag{39}$$

As in the normal case, our first simplifying assumption is to discard $\tilde{E}$ in (35). Unlike the normal case, we cannot prove that $||\tilde{E}||_2 = O(\sqrt{pqN})$ for p as small as $p = p_0 \log N/N$ (c.f. lemma 1.8). The numerical evidence in subsection 3.2 is consistent with $||\tilde{E}/\sqrt{pqN}||_2 = o(\sqrt{\log N})$ when $p = p_0 \log N/N$, however, so that $\tilde{E}$ is of lower order in this case as well. Substituting (39) and (36) into (35) and discarding $\tilde{E}$ we obtain

$$L \sim \hat{D},$$

where $\hat{D}$ denotes the diagonal matrix with entries

$$\begin{aligned} \hat{d}_{ii} &= \sum_{j=1}^n \kappa a_{ij} - \sum_{j=1}^n b_{ij} \\ \hat{d}_{i+n,i+n} &= \sum_{j=1}^n \kappa c_{ij} - \sum_{j=1}^n b_{ji} \end{aligned} \quad \text{for } 1 \le i \le n. \tag{40}$$

This amounts to approximating L by its diagonal. As in the normal case, we continue to assume that the entries $\hat{d}_{ii}$ are independent. While this is not true, in practice this assumption introduces negligile error to the overall computation. Next, we set $p = p_0 \log N/N$ and approximate the sum of n independent Bernoulli trials $\sum_{j=1}^n a_{ij}$ by a Poisson distribution. That is, we replace $\sum_{j=1}^n a_{ij} \sim P_\lambda$ with $\lambda = pn = (p_0 \log N)/2$. Then $\hat{d}_{ii} \sim \kappa P_\lambda - P_\lambda$ is the difference of two Poisson distributions. The threshold occurs precisely when $\mathbb{E}\max\left(\hat{d}_{1,1}, \ldots, \hat{d}_{N,N}\right) = 0$. By lemma 3.1 this happens precisely when $\lambda_0 = p_{0c}/2$ satisfies

$$\frac{p_{0c}}{2} = \frac{1}{2 - \kappa^{-\kappa/(\kappa+1)}\left(1 + \kappa\right)}. \tag{41}$$

This is precisely the threshold (9).

It remains to give the derivation of Lemma 3.1. We first recall Laplace's method since it plays a central role in the derivation. We state it as follows:

15

Lemma 3.3. (Laplace's Method) *Suppose that $f(x)$ is smooth inside $[a, b]$. Suppose that $f(x)$ has a global maximum at x_0 with $a < x_0 < b$ and with $f''(x_0) < 0$. Then*

$$\int_a^b \exp\left\{\frac{1}{\varepsilon}f(x)\right\} \sim \sqrt{\frac{2\pi\varepsilon}{|f''(x_0)|}} \exp\left\{\frac{1}{\varepsilon}f(x_0)\right\}\left(1 + O(\sqrt{\varepsilon})\right), \quad \varepsilon \ll 1 \tag{42}$$

If f is increasing inside $[a, b]$ and with $f'(b) > 0$, then

$$\int_a^b \exp\left\{\frac{1}{\varepsilon}f(x)\right\} \sim \frac{\varepsilon}{f'(b)} \exp\left\{\frac{1}{\varepsilon}f(b)\right\}\left(1 + O(\varepsilon)\right), \quad \varepsilon \ll 1 \tag{43}$$

See for example [12] for explanation of Laplace's method.

Proof of lemma 3.1. Let $f(t), g(t)$ and $h(t)$ be the probability density function for X, Y and Z, respectively. We have

$$f(t) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta(t - k);$$

$$g(t) = \frac{1}{\kappa} f\left(-\frac{t}{\kappa}\right);$$

$$h(t) = \int_{-\infty}^{\infty} f(s)g(t - s)ds.$$

Using the Stirling approximation formula

$$\log(k!) \sim k \log k - k + O(\log k) \quad k \gg 1.$$

we then estimate

$$f(t) \sim C \exp\left(\phi(t, \lambda)\right)$$

where we define

$$\phi(t, \lambda) := -\lambda + t\left(\log \frac{\lambda}{t} + 1\right), \quad t > 0.$$

Here and below, C will denote a positive quantity that has order at most $O(\log N)$. We then obtain

$$h(t) \sim \int_t^{\infty} C \exp\left\{\psi(t, s, \lambda)\right\} ds \quad \text{where} \quad \psi(t, s, \lambda) := \phi(s, \lambda) + \phi\left(\frac{s - t}{\kappa}, \lambda\right). \tag{44}$$

Consider the critical scaling

$$s = S \log N; \quad t = T \log N, \ \lambda = \lambda_0 \log N; \qquad S, T, \lambda_0 = O(1).$$

We then have

$$\exp\left\{\psi(t, s, \lambda)\right\} = \exp\left\{\psi(T, S, \lambda_0) \log N\right\}$$

so that the integral (44) can be estimated asymptotically using Laplace's method (42). It then yields

$$h(t) \sim C \exp\left\{\psi(T, S(T), \lambda_0) \log N\right\}, \quad t = T/\varepsilon, \quad T, \lambda_0 = O(1) \tag{45}$$

16

where $S(T)$ satisfies $\frac{d}{dS}\psi(T,S,\lambda_0)=0$:

$$\log\frac{\lambda_0}{S}+\frac{1}{\kappa}\log\frac{\kappa\lambda_0}{S-T}=0 \tag{46}$$

Elementary calculus shows that (46) has a unique root $S=S(T)>T$ which is the global maximum of $\psi(T,S,\lambda_0)$. Next, we approximate

$$\int_{-\infty}^{x_0}\exp\left\{\psi(T,S(T),\lambda_0)\log N\right\}dT \sim C\exp\left\{\psi(x_0,S(x_0),\lambda_0)\log N\right\} \tag{47}$$

This follows from the fact that the function $\psi(x_0,S(x_0),\lambda_0)$ is increasing in x_0 and from (43). Hence we have:

$$P(Z>x_0\log N)\sim 1-C\exp\left\{\psi(x_0,S(x_0),\lambda_0)\log N\right\}$$

and

$$P(\min(Z_1,\ldots Z_N)>x_0\log N)\sim (1-C\exp\left\{\psi(x_0,S(x_0),\lambda_0)\log N\right\})^N .$$

Using the elementary estimate $(1-x/N)^N\sim\exp(-x)$ as $N\to\infty$, we therefore obtain

$$P(\min(Z_1,\ldots Z_N)>x_0\log N)\sim\exp\left[-C\exp\left\{[\psi(x_0,S(x_0),\lambda_0)+1]\log N\right\}\right].$$

Thus we have a sharp threshold: if $\psi+1<0$ then $\min(Z_1,\ldots Z_N)>x_0\log N$ with probability rapidly approaching one; in the oppose case, $\min(Z_1,\ldots Z_N)<x_0\log N$ with probability rapidly approaching one. Thus $\mathbb{E}(\min(Z_1,\ldots Z_N))=x_0\log N$ precisely when $\psi(x_0,S,\lambda_0)+1=0=\frac{d}{dS}\psi(x_0,S,\lambda_0)$. Finally to show (34), we set $x_0=0$ in (32) to obtain

$$\log\frac{\kappa\lambda_0}{S}=-\kappa\log\frac{\lambda_0}{S} \tag{48}$$

so that (31) simplifies

$$0=1-2\lambda_0+S+\frac{S}{\kappa}; \qquad S=\lambda_0\kappa^{1/(\kappa+1)} \tag{49}$$

Solving (48, 49) for λ_0 then yields (34). $\qquad\square$

3.2 Numerical computations

To test the asymptotic theory, we compare the theoretical threshold (9) with numerical estimates of the threshold. For a given probability p and given system size N define $f(p,N):=\mathbb{E}(\lambda_2(L))$, where L denotes the stability matrix (6). We estimate $f(p,N)$ by taking the average of $\lambda_2(L)$ for 1000 different random realizations. To estimate p_{0c} for a fixed value of N, we use the bisection method to solve $f(p_c,N)=0$ then set $p_{0c}=p_cN/\log N$. This yields the following table.

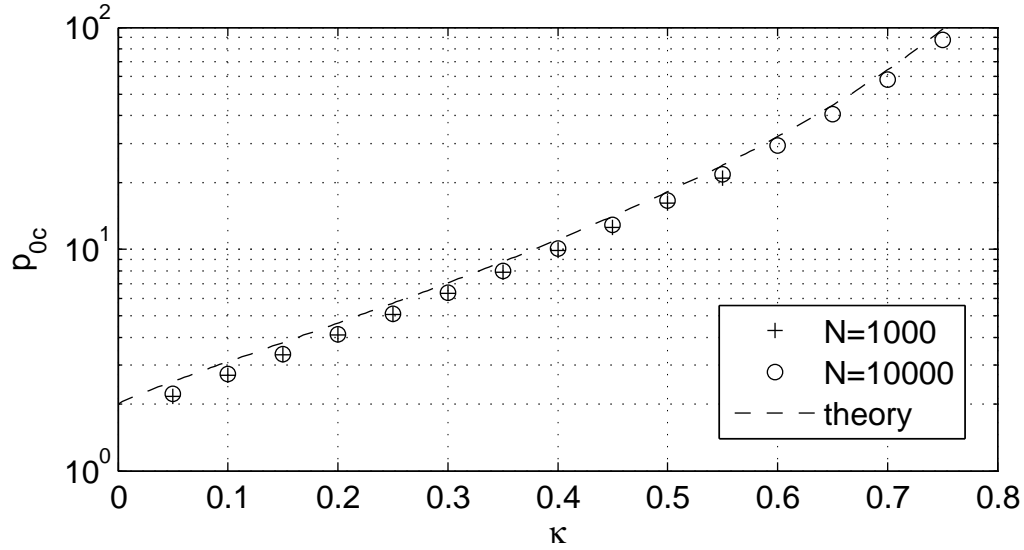Figure 2: Comparison between the numerics (see table on Page 18) and theory as given by (9) for the threshold value.

κ	Theoretical prediction for p_{0c} from (9)	p_{0c} from numerics	
		N=1000 (rel.err%)	N=10000 (rel.err%)
0.05	2.53484	2.173 (16.7%)	2.222 (14.1%)
0.1	3.10622	2.692 (15.4%)	2.737 (13.5%)
0.15	3.79413	3.331 (13.9%)	3.357 (13.0%)
0.2	4.64245	4.097 (13.3%)	4.145 (12.0%)
0.25	5.70426	5.065 (12.6%)	5.121 (11.4%)
0.3	7.05112	6.32 (11.6%)	6.364 (10.8%)
0.35	8.78366	7.886 (11.4%)	7.967 (10.2%)
0.4	11.0478	9.884 (11.8%)	10.06 (9.83%)
0.45	14.0616	12.56 (12%)	12.85 (9.44%)
0.5	18.1623	16.09 (12.9%)	16.54 (9.84%)
0.55	23.8918		21.81 (9.57%)
0.6	32.1643		29.25 (9.97%)
0.65	44.614		40.57 (9.97%)
0.7	64.3877		58.01 (11.0%)
0.75	98.1659		87.58 (12.1%)

From the asymptotics in lemma 3.1 we expect error of at least $O\left(\frac{1}{\sqrt{\log N}}\right)$. This demands very large N before we can expect to see reasonable agreement between the asymptotics and numerics. For example, as $\frac{1}{\sqrt{\log 10,000}} = 0.33$ the typical error of 10% with $N = 10,000$ in the table is still in line with expectatations. We do not include results for $N = 1000$ if $\kappa > 0.5$. This is because p_{0c} becomes too big, requiring a larger value of N than $N = 1000$ to make p sufficiently small. For

example $\kappa = 0.6$ yields $p_{0c} = 32.16$ so that $p = 32.16 \log(1000)/1000 \approx 0.2$ which is introduces an $O(p)$ error comparable with $O(1/\sqrt{\log N})$. The graph $p_{0c}(\kappa)$ is also shown in Figure 2. Good agreement is observed between the theoretical prediction and numerical computations.

As we have no rigorous proof of the simplifying assumptions in the derivation of formula (9), we attempt to verify them numerically. Specifically, we assume that the matrix $\tilde{E}$ in (35) is of lower order and that the diagonal entries in (39) are actually independent random variables. To verify the first assumption, let $Y = (X_p - p)/\sqrt{p(1-p)N}$ where X_p denotes a Bernoulli random variable. Take the critical scaling $p = p_0 \log N/N$ and consider a symmetric $N \times N$ matrix $M(N, p_0)$ with upper triangular entries $m_{ij} =_d Y$. We compute the expected operator norm $\mathbb{E}(\|M(N, p_0)\|_2)$ for fixed p_0 and N using an average of 100 independent trials. The following table shows the results as a function of p_0 and N up to $N = 10000$.

p_0	$\mathbb{E}(\|M(N,p_0)\|_2)$ (std)			
	$N = 500$	$N = 1000$	$N = 5000$	$N = 10000$
0.1	2.9063 (0.188)	2.9353 (0.169)	2.9793 (0.120)	2.9749 (0.098)
0.5	2.2394 (0.076)	2.2446 (0.057)	2.2461 (0.050)	2.2430 (0.047)
1	2.0992 (0.041)	2.1148 (0.036)	2.1087 (0.023)	2.1039 (0.028)
1.5	2.0473 (0.043)	2.0672 (0.028)	2.0648 (0.022)	2.0581 (0.029)
2	2.0218 (0.038)	2.0362 (0.033)	2.0417 (0.018)	2.0452 (0.017)
10	1.9225 (0.027)	1.9546 (0.023)	1.9865 (0.022)	1.9933 (0.017)

The value in parentheses denotes the standard deviation of the 100 trials. The results are consistent with our first assumption, that $\|M(N, p_0)\|_2 = o(\sqrt{\log N})$; in fact these numerics suggest that $\|M(N, p_0)\|_2 = O(1)$ as $N \to \infty$ and for fixed p_0. However, as $\sqrt{\log 10000} \approx 3.03$ is still rather large, a much more systematic numerical study is required to verify this conjecture with any certainty. This lies beyond the scope of the present work.

To verify the second assumption we consider the following numerical test. Let $S_i = \sum_{j=1}^{N} a_{ij}$ where $e_{ij} =_d X_p$ denote Bernoulli random variables with the dependence assumption $\xi_{ij} = \xi_{ji}$ and let $A_i = \sum_{j=1}^{N} b_{ij}$ where $b_{ij} =_d X_p$ denote fully independent Bernoulli random variables. We then compute $s := \mathbb{E}(\min(S_1, \ldots, S_N))$ and $a := \mathbb{E}(\min(A_1, \ldots, A_N))$. The second assumption essentially states that $s/a \to 1$ as $N \to \infty$. Numerically this is indeed true. In all the cases we tried, the difference between s and a was negligible. For example taking $N = 1000$, $p_0 = 1.5$ and using 2000 trials, we found that $a \approx 1.823$ and $s \approx 1.815$ with a nearly identical histogram of the sample trials.

4 Discussion

This paper presents a study of the behavior of a well-known swarming algorithm where the standard all-to-all coupling between agents is replaced by a random graph where two agents interact with some probability p. The classical 'compromise' solution is shown to lose stability when the graph is very sparse and estimates are derived on the sparseness of the graph (in terms of bounds on p) such that the clustering solution is no longer stable. While the best result we can obtain rigorously is that the compromise solution is stable when $p \geq O(\log^{3/2} N/N)$, the calculations in section 3 suggest that the critical probability scales like $p = O(\log N/N)$. Moreover, the constant is strictly larger than the threshold for connectivity of the underlying graph. Closing the gap between the rigorous result and our conjectured scaling remains a difficult open problem. The main difficulty
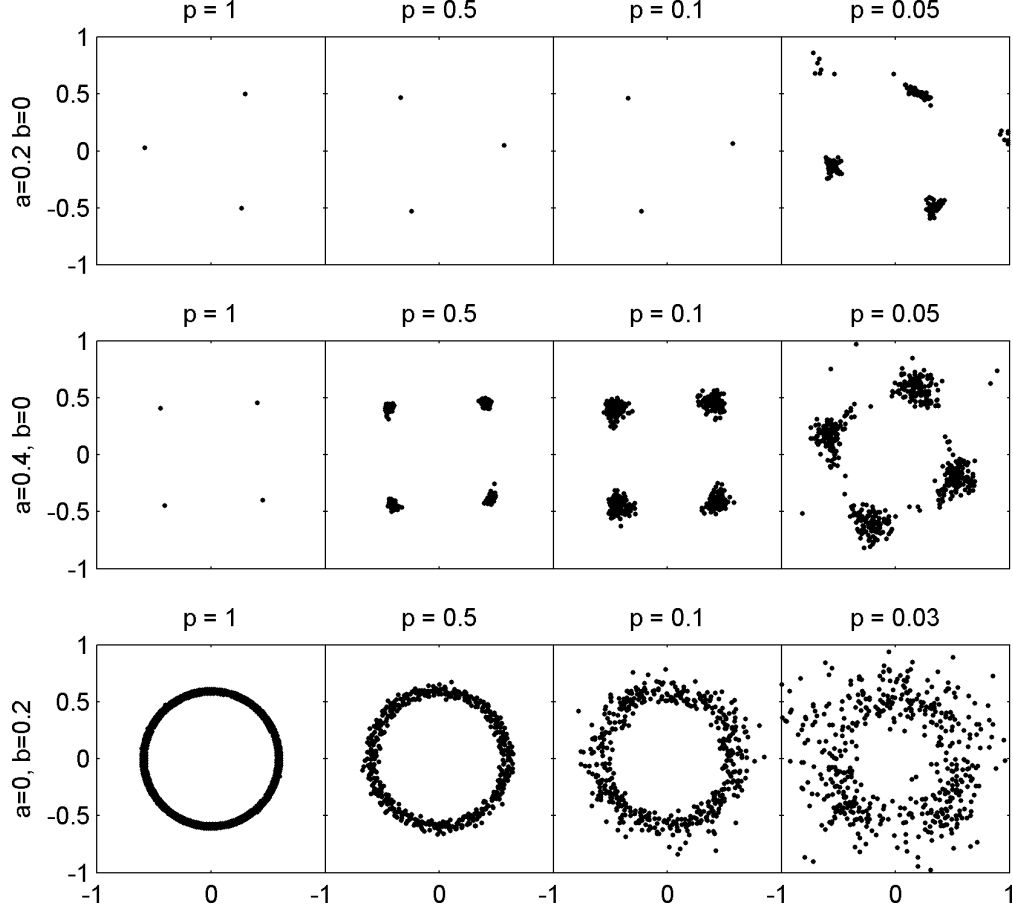
Figure 3: Examples of two-dimensional steady states for the system (1) with $F(r) = \min(ar+b, 1-r)$, e_{ij} drawn according to (4) and $N = 500$. The ODE system was simulated using the forward Euler method until the equilibrium was achieved. Note that the "shadow" of the steady state is preserved even for relatively small values of N.

lies in obtaining stronger estimates on the operator norm of Bernoulli random matrices, such as an improved version of lemma 1.8, when p scales with $\log N/N$. To the best of our knowledge such estimates do not yet exist, which demonstrates the need for better understanding of random Bernoulli matrices in the critical regime $p = O(\log N/N)$.

We note that the all-to-all coupling assumption underlying many aggregation models, while useful for analytical computations, is numerically expensive: simulating one step of an aggregation model on N fully coupled particles has a cost of $O(N^2)$. On the other hand, our analysis suggest that similar dynamics might be achieved with relatively sparse coupling, whereby each particle is coupled to only $O(\log N)$ other particles chosen at random. The cost of each step in the computation would then be of $O(N \log N)$. Figure 3 illustrates this phenomenon for some two dimensional models in which the geometry is more complex than that of the line. The sparse connectivity of the graph manifests as a 'noisy' version of the fully-connected configuration even though we use a fixed graph for the simulation and no noise is actually present. We conjecture that the expected

value over different realizations of these sparse graphs (where the connection is viewed as a random variable) would lead to an additional diffusion process. It is also an interesting open question to study how the sparsity of connections affects the confinement properties of the potential. Numerical experiments shown in figure 3 indicate that confinement is preserved even for relatively small values of p. Based on the results of this paper, it is tempting to conjecture that confinement is preserved up to $p = O(\log N/N)$.

Our current analysis readily extends to higher dimensions and to more general random graph models. In higher dimensions, 'simplex' configurations (c.f. the top row in figure 3) are the natural analogue of the 'compromise' solution that we consider in this paper. Without the random graph structure, the stability analysis for such solutions already exists [13, 23]. Extending the present analysis to these cases would therefore only requires a version of the trace method for block matrices, but in principle this extension is straightforward. Outside of this modification, the program to demonstrate stability remains the same. The assumption that all agents interact, on average, with the same number Np of other agents might be inappropriate depending on the application. The generalization of the Erdős-Rényi model $\mathcal{G}(N, p)$ due to Chung and Lu [7, 6] overcomes this difficulty by allowing for arbitrary degree sequences. Provided the minimum degree of the graph is sufficiently large, the trace method also applies when studying the Laplacian matrix for such graphs [8]. Our results should therefore extend, in a straightforward manner, to this generalized version of the Erdős-Rényi model.

Acknowledgments

References

[1] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization, 2000.

[2] Z. D. Bai and Y. Q. Yin. Necessary and sufficient conditions for almost sure convergence fo the largest eigenvalue of a wigner matrix. *Ann. Prob.*, 16(4):1729–1741, 1988.

[3] B. Bollobás. *Random Graphs (2nd ed.)*. Cambridge University Press, 2001.

[4] M. Cao, A.S. Morse, and B.D.O. Anderson. Reaching an agreement using delayed information. *Decision and Control, 2006 45th IEEE Conference on*, pages 3375–3380, Dec. 2006.

[5] Y.-L. Chuang, M. R. D'Orsogna, D. Marthaler, A. L. Bertozzi, and L. Chayes. State transitions and the continuum limit for a 2d interacting, self-propelled particle system. *Physica D*, 232:33–47, 2007.

[6] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *PNAS*, 99(2):15879–15882, 2002.

[7] F. Chung and L. Lu. Connected components in random graphs with given degree sequences. *Ann. Combin.*, 6:125–145, 2002.

[8] F. Chung, L. Lu, and V. Vu. The spectra of random graphs with given expected degrees. *Internet Mathematics*, 1(3):257–275, 2004.

[9] Maria R. D'Orsogna, Yao-Li Chuang, Andrea L. Bertozzi, and Lincoln Chayes. Self-propelled particles with soft-core interactions: patterns, stability, and collapse. *Physical Review Letters*, 96:104302, 2006.

[10] P. Erdős and A. Rényi. On random graphs i. *Publicationes Mathematicae*, 6:290–297, 1959.

[11] K. Fellner and G. Raoul. Stable stationary states of non-local interaction equations. *Math. Models Methods Appl. Sci.*, 20:22672291, 2010.

[12] M.H. Holmes. *Introduction to Perturbation Methods*. Springer, 1995.

[13] T. Kolokolnikov, Y. Huang, and M. Pavlovski. Singular patterns for an aggregation model with a confining potential, 2012. To appear, Physica D.

[14] Theodore Kolokolnikov, Hui Sun, David Uminsky, and Andrea L. Bertozzi. Stability of ring patterns arising from 2d particle interactions. *Phys. Rev. E*, 84(1), 2011. Rapid Communications.

[15] L. Moreau. Stability of continuous-time distributed consensus algorithms. *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, 4:3998–4003 Vol.4, Dec. 2004.

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

[17] R. Olfati-Saber. Flocking for multi-agent dynamic systems: algorithms and theory. *Automatic Control, IEEE Transactions on*, 51(3):401–420, March 2006.

[18] R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *Automatic Control, IEEE Transactions on*, 49(9):1520–1533, Sept. 2004.

[19] W. Ren and E. M. Atkin. Distributed multi-vehicle coordinated control via local information exchange. *International Journal of Robust and Nonlinear Control*, 17(10-11):1002–1033, 2007.

[20] W. Ren and R. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control:Theory and Applications*. Communications and Control Engineering. Springer, London, 2008.

[21] W. Ren, K. Moore, and Y. Chen. High-order and model reference consensus algorithms in cooperative control of multi-vehicle systems. *ASME J. Dyn. Sys. Meas. Control*, 129(5):678–688, 2007.

[22] Wei Ren and R.W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *Automatic Control, IEEE Transactions on*, 50(5):655–661, May 2005.

[23] Hui Sun, David Uminsky, and Andrea L. Bertozzi. Stability and clustering of self-similar solutions of aggregation equations, 2012. accepted in J. Math. Phys.

[24] T. Tao. *Topics in Random Matrix Theory*. American Mathematical Society, 2012.

[25] J. H. von Brecht and D. Uminsky. On soccer balls and linearized inverse statistical mechanics, 2012. J. Nonlin. Sci.

[26] J. H. von Brecht, D. Uminsky, T. Kolokolnikov, and A.L. Bertozzi. Predicting pattern formation in particle interactions. *Math. Models and Meth. in Appl. Sci.*, Suppl. 4, 2012.

[27] Van H. Vu. Spectral norm of random matrices. *Combinatorica*, 27(6):721–736, 2007.

[28] E. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. Math.*, 67:325–327, 1958.