Effects of social interactions in collective motion

Jack Thomas Parley

supervised by
Dr. Romualdo Pastor

June 21, 2016
Abstract

The Vicsek model of collective motion is known to undergo a noise-induced phase transition from a disordered state with random velocities, to a collectively ordered state where all the particles tend to move in the same spontaneously selected direction. In this work, we have studied the outcome of this model on various social network topologies, in order to investigate how social interactions affect the transfer of directional information. The original results for off-lattice particles are first reproduced. The model is then implemented on random and small-world graphs, where the effects of increasing connectivity, in the first case, and the long-range interactions brought about by the rewiring of edges, in the second case, are studied. On the other hand, scale-free topologies are found to have a higher critical noise $\eta_c$ than homogeneous topologies with the same connectivity, with the high-degree nodes playing a crucial role. Weighted interactions are also studied, for which there is found to exist an optimal weighting exponent which gives the highest possible value of the order parameter at equilibrium.
Contents

1 Introduction 1

2 Social Networks 3
  2.1 Network definition and topological properties ............... 3
  2.2 Topological properties of social networks ................. 5
  2.3 Models of social networks .................................. 6
    2.3.1 Erdos Renyi ........................................ 6
    2.3.2 Small world (Watts-Strogatz) model .......... 10
    2.3.3 Barabási-Albert model ........................... 13
    2.3.4 Uncorrelated configuration model (UCM) ......... 17

3 2D Vicsek model 19
  3.1 Model definition ........................................... 20
    3.1.1 Algorithmic implementation ......................... 23
  3.2 Nature of the phase transition ............................ 24
  3.3 Networks as mean-field representation of interactions .... 26

4 Vicsek model on networks 27
  4.1 Alternative expression of the order parameter ............ 27
  4.2 Random networks ........................................... 29
    4.2.1 Phase transition .................................... 29
    4.2.2 Effect of connectivity on phase transition ....... 30
    4.2.3 Ordering time ....................................... 31
  4.3 Small-world networks ...................................... 34
    4.3.1 Small-world effect .................................. 34
    4.3.2 Ordering time ....................................... 35
  4.4 Scale-free networks ....................................... 37
    4.4.1 Effect of scale-free topology on $n_e$ ............ 37
    4.4.2 Finite-size scaling .................................. 38
    4.4.3 Hierarchy of the order parameter .................. 41
    4.4.4 Ordering time ....................................... 41
    4.4.5 Hysteresis for extrinsic noise ...................... 43
  4.5 Weighted networks ........................................ 45
    4.5.1 Dynamics without noise .............................. 45
    4.5.2 Optimal exponent .................................... 46
5 Conclusions
1 Introduction

Collective phenomena, in which repeated interactions among many individuals parts produce patterns on a scale larger than themselves, are within us and all around us. Some of the best examples of these complex systems can be found in the collective motion of animals, in which there are two clearly defined levels of organization to be linked together: the animal and the group (in other collective phenomena, such as protein interactions or ecological webs, it is more difficult to establish exactly on which level to study the system) [1]. They also provide some of the most spectacular and fascinating sights in the natural world: flocks of birds turning in unison or migrating in well ordered formation (Figure 1.1); fish shoals splitting and reforming as they outmaneuver a predator...

Indeed, the challenge of understanding how hundreds or thousands of organisms can self-organize and move together in the same direction without the presence of any gradient field or any apparent leader has attracted many scientists for a long time. The main tool used in describing the dynamics of flocking are self-propelled particle (SPP) models. In SPP models “particles” move with constant absolute velocity and each one has a local interaction zone within which it responds to other particles. This interaction may include repulsion, attraction and alignment. For example, Couzin et al. (2002) [2] proposed a model in which individual animals have three interaction radii of increasing size, so that individuals are attracted to neighbours over a larger range than they align, but also possess a strong repulsion zone at short range. This simple model can produce complex patterns: as the alignment radius is increased (keeping the repulsion and attraction radii constant), individuals go from a loosely packed stationary swarm, to a torus where individuals circle around their center of mass, and finally to a parallel group moving in a common direction.

The model we will study in this work is the original SPP model introduced by Vicsek et al. (1995) [3], where only alignment, to which noise is added, is taken into account. Despite its simplicity, Vicsek’s model in 2D and 3D generates spectacular dynamical patterns that are highly reminiscent of the movement of flocks. In addition, it leads to a phase transition, which can be studied borrowing methods from statistical mechanics. This phase transition is not only interesting in its own right, but is also observed in, for example, groups of moving locusts, whose collective motion undergoes a similar transition when the density reaches a critical value [4].
Of course, the simplicity of Vicsek’s interaction rule leaves out many important aspects. Real interactions are much more complicated. For example, one of the most celebrated mechanisms for transferring information is the waggle dance of the honeybee. Waggle dances are performed by honeybee foragers that have successfully found nectar or pollen. The dance is a figure eight pattern, the direction and duration of which are correlated with the direction and the distance from the bee hive to the food source. Uninformed bees in the hive follow a dance and then fly in the direction of and for the distance encoded by the dance, after which they may repeat the dance themselves and produce a positive feedback mechanism through which information is transferred.

It is clear that making movement decisions often depends on social interactions among group members. Couzin et al. (2005) [5] proposed a simple model through which they show how a group can make consensus decisions with only a very small proportion of individuals. In this model the informed individuals balance the influence of their preferred direction and their social interactions with a weighting term (in this work we will also use weighted interactions).

Also worth mentioning, on the other hand, is the work by Rosenthal et al. (2015) [6], in which they reveal the hidden network of interaction in schooling fish, by analyzing the sensory input of each individual.

In this work we will investigate how directional information is transferred on different models of social networks, which are presented in Section 2. In Section 3 we reproduce the original Vicsek results in 2D, and discuss the theoretical framework for studying the results in Section 4.

Figure 1.1: Flocking behaviour
2 Social Networks

2.1 Network definition and topological properties

We can define a network as a collection of points joined together by lines. In this work we will follow the customary jargon and refer to points as nodes or vertices and the lines as edges. In essence they are an abstract way of representing a set of components or individual parts which are linked together in some way. Examples of systems which can be modeled as networks include the Internet, social networks, citation networks, and biological networks such as food webs or neural networks. The network representation of systems is useful and widespread since it can be used to obtain information regarding the topology of the system, and how this topology affects the dynamics such as ordering or flocking.

In the case of social networks, the nodes represent individuals and edges the presence of social relations between pairs of individuals. Our working hypothesis is that when swarming, animals may not look only at the behaviour of their closest neighbors in euclidean distance, but may also be affected by the behaviour of other animals to whom they are socially related, even if they are at a larger euclidean distance.

Although scientists in a wide variety of fields have, over the years, developed an extensive set of tools for analyzing and modeling networks, here we focus on three key ways of characterizing a network [7].

Degree distribution The degree of a node $k$ is the number of other nodes to which it is connected. One of the most fundamental of network properties is the frequency distribution of node degrees. Each network model will have an associated $P(k)$, which will denote the probability that a randomly chosen node has degree $k$.

Clustering coefficient Clustering refers to the property of certain networks to show transitivity in their connections, and can be quantified by the clustering coefficient $c_i$ which represents the average probability that a pair of $i$'s connections are also connected to one another. That is, in the case of social networks, it measures the increased likelihood that two of your friends are also friends of each other. Numerically it can be calculated as $c_i = \frac{e_i}{k_i(k_i-1)/2}$, where $e_i$ is the number of edges between the nearest neighbours of $i$, and $\frac{k_i(k_i-1)}{2}$ is the total number
of edges that could exist between these neighbours. The average clustering coefficient of the network is found from each local coefficient as
\[ \langle c \rangle = \frac{\sum c_i}{N}. \]

**Average path length** The average path length is defined as the average number of steps along the shortest paths for all possible pairs of network nodes. The standard algorithm (and the one used in this work) is the breadth-first search. A single run of the breadth-first search algorithm finds the shortest distance from a single source node \( s \) to every other vertex in the same component of the network as \( s \). Although we may only want the shortest distance between a single pair of nodes \( s, t \), there is no procedure known for calculating such a distance that is faster in the worst case than calculating the distances from \( s \) to every other node using breadth-first search and then throwing away all of the results except for the one we want.

![Figure 2.1: Basic metrics characterizing a vertex \( i \) in the network. A, The degree \( k \) quantifies the vertex connectivity. B, The shortest path length identifies the minimum connecting path (dashed line) between two different vertices. C, The clustering coefficient provides a measure of the interconnectivity in the vertex’s neighborhood. As an example, the central vertex in the figure has a clustering coefficient \( C = 1 \) if all its neighbors are connected and \( C = 0 \) if no interconnections are present.](image)
2.2 Topological properties of social networks

The main properties of social networks, as discovered from empirical results for real networks and databases, are related to the three main topological measures just presented.

1. Long-tailed degree distributions
   For a large number of networks, including the World Wide Web, the Internet and different social networks, it has been discovered that the degree-distribution is long-tailed, with a power-law decay. That is, there are "hubs", or individuals with many more connections than the average. The power law exponent has been measured in social networks such as sexual contacts, movie actors or citations, yielding values between 2 and 3 [12].

2. Small-world
   The small-world concept describes the fact that despite their often huge size, in most social networks there is a relatively short path between any two nodes. The most popular manifestation of small worlds is the six degrees of separation concept, uncovered by the social psychologist Stanley Milgram (1967), who concluded that there was a path of acquaintances with a typical length of about six between most pairs of people in the United States.

3. Clustering
   A common property of social networks is that cliques form, representing circles of friends or acquaintances in which every member knows every other member. Most empirical results show the clustering coefficient in social networks is much higher than what random interactions would predict.
2.3 Models of social networks

In order to investigate the effects of the different properties of social networks on swarming and flocking dynamics, we will consider different theoretical network models that reproduce some of the properties of these social networks. These models have been implemented on Fortran using jagged array structures, which allow us to save the adjacency list (a list with each node and its neighbours) without occupying too much memory.

2.3.1 Erdos Renyi

The random graph is the most basic of network models. Although there are two closely related models for generating random graphs, in this work I will focus on the Erdos-Renyi model, and it is this model I will use in the simulations. It is named after Paul Erdos and Alfred Renyi, who published a celebrated series of papers about the model in the late 1950s and early 1960s [8]. The other model was introduced contemporaneously and independently by Edgar Gilbert. Gilbert’s model is slightly different in that each edge has a fixed probability of being present or absent, independently of the other edges; while in the Erdos-Renyi model all graphs on a fixed vertex set with a fixed number of edges are equally likely.

General properties The model is defined by two fixed parameters: $n$, the number of vertices, and $m$, the number of edges. To generate the network we choose $m$ pairs of vertices uniformly at random and connect them with an edge. For the graph to be simple, that is without self-edges or multiedges, one must also impose that the pairs chosen are distinct and not already connected.

Due to the nature of its generative process, the random graph model is not defined in terms of a single randomly generated network, but as an ensemble of networks. The ensemble of all possible simple graphs with $n$ vertices and $m$ edges is represented by $G(n,m)$, where each graph has a uniform probability $P(G) = \frac{1}{\Omega}$.

Although some properties of the random graph are straightforward, such as the average degree $\langle k \rangle = \frac{2m}{n}$, most mathematical work has to be carried out on a slightly different model defined by $G(n,p)$. The parameter fixed here is not the number of edges, but the independent
probability of an edge existing between each distinct pair of vertices. Thus $P(G)$ can now be written as

$$P(G) = p^m (1 - p)^{\binom{n}{2} - m} \quad (2.1)$$

Once $G(n, m)$ has been replaced by $G(n, p)$, the number of edges is no longer fixed. However as the number of edges $m$ is the outcome of creating an edge with probability $p$ at each possible distinct pair of vertices $\binom{n}{2}$, it is straightforward to see we have the standard binomial distribution:

$$P(m) = \binom{\binom{n}{2}}{m} p^m (1 - p)^{\binom{n}{2} - m} \quad (2.2)$$

the mean value of which is given by:

$$\langle m \rangle = \sum_{m=0}^{\binom{n}{2}} m P(m) = \binom{n}{2} p \quad (2.3)$$

and therefore

$$\langle k \rangle = \sum_{m=0}^{\binom{n}{2}} \frac{2m}{n} P(m) = \frac{2}{n} \binom{n}{2} p = (n - 1)p \quad (2.4)$$

A crucial property of a network is its degree distribution. Finding the degree distribution for $G(n, p)$ is also straightforward. A given vertex in the graph is connected with independent probability $p$ to each of the $n - 1$ other vertices, and there are $\binom{n-1}{k}$ ways to pick $k$ vertices, so:

$$p_k = \binom{n-1}{k} p^k (1 - p)^{n-1-k} \quad (2.5)$$

It can be shown that in the limit of large $n$ equation 2.5 becomes a Poisson distribution:

$$P(k) = e^{-c} \frac{c^k}{k!} \quad (2.6)$$
Where $c$ is the mean degree $\langle k \rangle$. This is the reason why the Erdos-Renyi graph is also called the Poisson random graph.

**Giant component** For the purpose of this work, and for many other applications, it is important that the network be fully connected, i.e. possess a giant component which fills the whole network. The giant component is defined as a component whose size grows in proportion to $n$, and is therefore extensive. One of the most interesting predictions of the Erdos-Renyi model is that as $p$ is gradually increased from 0 (no edges), to 1 (all nodes connected), the size of the large component does not increase gradually. There is, in fact, a *phase transition*: there is a value of $p$ where there suddenly appears a giant component.

Why is this the case? We denote by $u$ the average fraction of vertices in the graph which do not belong to the giant component. $u$ can also be seen as the probability that a randomly chosen vertex does not belong to the giant component. For a vertex $i$ not to be connected to the giant component through another vertex $j$, either there is no edge between them or vertex $j$ itself is not part of the giant component. This gives:

$$u = (1 - p + pu)^{n-1} = \left[1 - \frac{c}{n-1}(1-u)\right]^{n-1} \quad (2.7)$$

For the limit of large $n$, the following expression holds for $u$:

$$u = e^{-c(1-u)} \quad (2.8)$$

But if $u$ is the fraction of vertices not in the giant component, then the fraction of vertices that are in the giant component is $S = 1-u$. Eliminating $u$ in favor of $S$ then gives us

$$S = 1 - e^{-cS} \quad (2.9)$$

This last equation predicts a transition between two distinct regimes: for equation 2.9 to have a non-zero solution, the following condition must be fulfilled.

$$\frac{d}{dS}(1 - e^{-cS}) = 1 \text{ or } ce^{-cS} = 1 \quad (2.10)$$
Imposing $S = 0$, we obtain $c = 1$ or $\langle k \rangle = 1$ as the critical point of the transition, which corresponds to a probability of $p_c = \frac{1}{n}$ (for large $n$). In other words, this is the threshold required to observe a connected network.

As a side note: besides its theoretical interest, this phase transition may have profound biological implications. In fact, it has been argued that as the diversity of catalysts in a network of chemical reactions increases there may be an analoguous transition to collectively autocatalytic sets, thus providing an explanation for the origin of the simplest life forms [9].

**Limitations** The Poisson random graph, despite being one of the best studied, has some severe shortcomings as a network model. Unlike most real-world social networks, it shows essentially no transitivity or clustering. The expression for the clustering in a random graph is $C = \frac{\langle k \rangle}{n-1}$, which tends to 0 as we increase $n$. Moreover, real networks typically have a right-skewed degree distribution, with most vertices having low degree but with a small number of high-degree “hubs” in the tail of the distribution (Figure 2.2). Different models (Watts-Strogatz, Albert-Barabasi) are therefore needed to better model real-world networks.

![Figure 2.2](image)

Figure 2.2: Comparison of degree distribution $P(k)$ in double logarithmic scale between Erdos-Renyi graph (short tail) and scale-free graph (long tail); both $N = 10^5$ and average degree $\langle k \rangle = 20$. 

9
2.3.2 Small world (Watts-Strogatz) model

General properties The small-world model, first introduced in a paper by Duncan J. Watts and Steven H. Strogatz [10], is an attempt to interpolate between regular and random networks. At one extreme, we have a perfectly regular graph (a circle) with high transitivity (high clustering coefficient) but a very large average path length. At the other extreme, a totally random network shows the small average path length characteristic to many real-world networks, but it has a vanishing clustering coefficient. Thus this model provides a connection topology which is neither completely random nor regular; like many biological, technological and social networks, which lie somewhere between the two extremes.

In order to perform the interpolation, the following random rewiring procedure is considered (Figure 2.3). Starting from a ring lattice with \( n \) vertices and \( k \) edges per vertex, we rewire each edge at random with probability \( p \). This construction allows us to tune the graph between regularity (\( p = 0 \)) and disorder (\( p = 1 \)), and thereby to probe the intermediate region \( 0 < p < 1 \).

The degree distribution for any value of \( p \) is straightforward to calculate [7]. In the initial ring every vertex has the same degree \( c \): the network is a regular graph. Once we add shortcuts to the circle to make the small-world model, the degree of a vertex is \( c \) plus the number of shortcut edges attached to it. As \( cp \) shortcuts on average end at any particular vertex, and the number \( s \) of shortcuts is Poisson distributed, \( ps \) will take the form:

\[
P(s) = e^{-cp} \frac{(cp)^s}{s!}
\]

We can find the degree distribution \( P(k) \) by putting \( s = k - c \) into this last equation:

\[
P(k) = e^{-cp} \frac{(cp)^{k-c}}{(k-c)!}
\]

\( k \geq c \)
Therefore the degree distribution has a peaked shape with a lower cut-off, and so in this respect it does not mimic real-world networks.

**Average path length and clustering** However, the importance of the small-world model lies in two other structural properties: their characteristic path length and their clustering coefficient. Their analytical expressions are more complicated; indeed no exact expression for the mean distance has yet been found. The limiting cases are clear though. In the regime where \( n \gg k \gg \ln n \gg 1 \) (which guarantees the connectedness of the random graph), we find that \( L \sim \frac{n}{2k} \) and \( C \sim \frac{3}{4} \) as \( p \to 0 \), while \( L \approx L_{\text{random}} \sim \frac{\ln(n)}{\ln(k)} \) and \( C \approx C_{\text{random}} \sim \frac{k}{n} \ll 1 \) as \( p \to 1 \). These limiting cases might lead one to suspect that large \( C \) is always associated with large \( L \), and small \( C \) with small \( L \).

![Watts-Strogatz algorithm and effects](image)

**Figure 2.3:** Watts-Strogatz algorithm and effects, reproduced from [10]

On the contrary, Figure 2.3 shows that there is a broad interval of \( p \) over which \( L(p) \) is almost as small as \( L_{\text{random}} \), yet \( C(p) \gg C_{\text{random}} \). This is due to the fact that for small \( p \), each shortcut has a highly non-linear effect on \( L \), contracting the distance not just between the pair of vertices that it connects, but between their immediate neighbourhoods, neighbourhoods of neighbourhoods and so on. By contrast, an edge removed from a clustered neighbourhood to make a shortcut has, at most, a linear effect on \( C \); hence \( C(p) \) remains practically unchanged for small \( p \) even though \( L(p) \) drops rapidly.
Small-world connectivity has important consequences for networks of coupled dynamic systems. For example, infectious diseases are predicted to spread much more easily and quickly in a small world; and small-world networks of coupled phase oscillators synchronize almost as readily as in the mean-field model, despite having orders of magnitude fewer edges. In the realm of collective motion, small-world connectivity will also have important consequences.
2.3.3 Barabási-Albert model

As already mentioned, the empirical study of real-world networks has shown that many of them have a scale-free degree distribution, that is, \( P(k) \) follows a power law for large \( k \). We have already seen that neither the random graph nor the Watts-Strogatz model are capable of explaining this topology. Accounting for this scale-freeness will require a shift from modeling network topology to modeling assembly and evolution. This is the idea behind the Barabási-Albert model.

Algorithm Barabási and Albert (1999) [11] were the first to argue that the scale-free nature of real networks is rooted in two generic mechanisms shared by many real networks: growth and preferential attachment. On the one hand, most real-world networks are open systems that grow by the continuous addition of new nodes. On the other, the likelihood of connecting to a node depends on the node’s degree; for example a web page will be more likely to include hyperlinks to popular documents with already high degree. These two ingredients constitute the basis of the Barabási-Albert algorithm:

1. *Growth*
   - Starting with a small number \( (m_0) \) of nodes, at every time step, we add a new node with edges \( m \leq m_0 \) that link the new node to \( m \) different nodes already present in the system.

2. *Preferential attachment*
   - When choosing the nodes to which the new node connects, we assume that the probability \( \Pi \) that a new node will be connected to node \( i \) depends on the degree \( k_i \) of node \( i \), such that

     \[
     \Pi(k_i) = \frac{k_i}{\sum_j k_j}
     \]  

(2.12)

After \( t \) time steps this procedure results in a network with \( N = t + m_0 \) nodes and \( mt \) edges.

In my simulations, for simplicity, I have chosen the initial nucleus to be a clique: \( m_0 \) nodes where each node is connected to every other node. \( m_0 \) is taken as \( m_0 = m + 1 \), where \( m \), as explained earlier, is the number of new edges created at each step.
Degree distribution As seen by performing numerical simulations of this algorithm (Figure 2.4), the procedure just described produces a scale-free degree distribution, with average degree \(2m\). Various analytical approaches have been taken to study the distribution. Either way, it can be shown that the resultant degree distribution is independent of any of the parameters \((m, m_0)\) and follows a power law with exponent \(\gamma = 3\) \((k^{-3})\).

The main relevance of a scale-free degree distribution lies in the different values of its moments, \(<k^n>\):

\[
<k^n> = \sum_k k^n P(k) \simeq \int_m^{\infty} k^n P(k) dk \sim \int_m^{\infty} k^{n-\gamma} dk
\]

For the average degree to be well defined, we require \(\gamma > 2\), and in the range of interest \((2 < \gamma < 3)\), the second moment diverges \((<k^2> \to \infty)\).

In reality, no vertex has infinite degree, and so the upper limit of the previous integral cannot be \(\infty\). Rather, there is an upper cut-off \(k_c(N)\), which is a growing function of network size. Replacing \(\infty\) by \(k_c(N)\), we have:

Figure 2.4: Degree distribution for BA network, size \(N = 10^6\), generated on FORTRAN, fitted to \(P(k) \sim k^{-3}\)
\[ \langle k^n \rangle \sim \int_m^{k_c(N)} k^{2-\gamma} dk \sim \frac{k_c(N)^{3-\gamma} - m^{3-\gamma}}{3-\gamma} \] (2.14)

Since \( k_c(N) \gg m \), the following scaling relation holds:

\[ \langle k^n \rangle \sim k_c(N)^{3-\gamma} \] (2.15)

It is usually observed in scale-free networks that \( k_c(N) \sim N^{\beta} \), so:

\[ \langle k^2 \rangle \sim N^{\beta(3-\gamma)} \] (2.16)

Therefore the variance of the degree distribution diverges with increasing network size, and this will have important effects on the dynamical processes taking place on the network.

**Average path length and clustering coefficient** As discussed earlier, a characteristic feature of real networks is the coexistence of clustering and short path lengths. Thus we need to investigate whether the network generated by the model has a small-world character. While the Barabasi-Albert model captures the power-law tail of the degree distribution, these other properties may not agree with empirical results on real networks.

As regards the clustering coefficient, there is yet no analytical prediction for the Barabasi-Albert model. However, numerical simulations show that the clustering coefficient of the Barabasi-Albert model decreases with the network size, following approximately a power law \( C \sim N^{-0.75} \), which, while a slower decay than the \( \frac{k}{N} \) decay observed for random graphs, is still different from the behavior of the small-world models, where \( C \) is independent of \( N \).

Finally, numerical simulations indicate that the average path length is smaller in the Barabasi-Albert network than in a random graph for any \( N \), indicating that the heterogeneous scale-free topology is more efficient in bringing the nodes close than is the homogeneous topology of random graphs. It is found that the average path length of the Barabasi-Albert network increases approximately logarithmically with \( N \), the best fit following a generalized logarithmic form [12]
\[ l = A \ln (N - B) + C \] (2.17)

Analytical results indicate that there is a double logarithmic correction to the logarithmic N dependence, i.e., \( l \sim \frac{\ln(N)}{\lnh(N)} \) [13].
2.3.4 Uncorrelated configuration model (UCM)

The last model used in my simulations is the uncorrelated random scale-free network. This type of network is a useful null model to check the accuracy of analytical solutions of dynamical processes defined on complex networks. In order to generate the network a variant of the configuration model is used.

The configuration model (CM) [7] allows one to construct random networks with any prescribed degree distribution \( P(k) \). To construct a network with the original definition of this algorithm, we start assigning to each vertex \( i \), in a set of \( N \) vertices, a random degree \( k_i \) drawn from the probability distribution \( P(k) \), with \( m \leq k_i \leq N \) (no vertex can have a degree larger than \( N \)), and imposing the constraint that the sum \( \sum_i k_i \) must be even. The network is completed by randomly connecting the vertices with \( \sum_i k_i / 2 \) edges, respecting the preassigned degrees. The result of this construction is a random network whose degrees are, by definition, distributed according to \( P(k) \) and in which, in principle, there are no degree correlations, given the random nature of the edge assignment.

While this prescription works well for bounded degree distributions, in which \( \langle k^2 \rangle \) is finite, one has to be more careful when dealing with networks with a scale-free distribution, which, for \( 2 < \gamma \leq 3 \), yield diverging fluctuations, \( \langle k^2 \rangle \to \infty \), in the infinite network size limit. In fact, it is easy to see that, if the second moment of the degree distribution diverges, a completely random assignment of edges leads to the construction of an uncorrelated network, but in which a non-negligible fraction of self-connections (a vertex joined to itself) and multiple connections (two vertices connected by more than one edge) are present. Self-connections and multiple connections make perfect mathematical sense; however they are not desired for the simulation purposes of this work.

If one simply forbids multiple connections and self-connections during the generating process, correlations of a mixed disassortative nature appear [14]. These degree-degree correlations are characterized by means of the average degree of the nearest neighbours of the vertices of degree \( k \):

\[
\bar{k}_{nn} = \sum_{k'} k' P(k'|k)
\]  

(2.18)

When \( \bar{k}_{nn}(k) \) is an increasing function of \( k \), the corresponding network is said to exhibit assortative mixing by degree, i.e. highly connected vertices are preferentially connected to highly connected vertices and vice-versa, while
a decreasing $\bar{k}_{nn}(k)$ function is typical of disassortative mixing, highly connected vertices being more probably connected to poorly connected ones.

The appearance of the correlations is due the effects of the cut-off (or maximum expected degree) $k_c(N)$. To avoid them, Romualdo et al. [14] propose the uncorrelated configuration model (UCM), utilising the fact that in order to have no correlations the cut-off can scale at most as $k_s(N) \sim N^{1\over 2}$ (the so-called structural cut-off). Essentially UCM consists in following the same scheme as the configuration model but subjecting the pre-assigned degrees to the constraints $m \leq k_i \leq N^{1\over 2}$ as well as $\sum_i k_i$ even. The constraint on the maximum possible degree of the vertices ensures that $k_c(N) \sim N^{1\over 2}$, allowing for the possibility to construct uncorrelated networks.
3 2D Vicsek model

As stated in the introduction, the Vicsek model, first introduced in 1995 by Vicsek et al. [3], is perhaps the simplest model displaying a transition to collective motion; in the study of active matter it plays a prototypical role, similar to the one played by the Ising model for equilibrium ferromagnetism. The simplicity of the model responds to the ubiquity of collective motion phenomena at all scales, from groups of large vertebrates to subcellular collective dynamics. This ubiquity strongly hints at the existence of some universal features, possibly shared among the many different situations, regardless of many individual-level details. We can approach this universality by constructing minimal models of collective motion, that is models stripped of as many details as possible and only equipped with the basic features that we believe characterize the problem, typically its fundamental symmetries and conservation laws [15].

The only rule of the model is at each time step a given particle driven with a constant absolute velocity assumes the average direction of motion of the particles in its neighbourhood of radius $r$ with some random perturbation added. Using simulations it can be shown that, despite its simplicity, this model results in a rich, realistic dynamics (like the flocking behaviour shown in Figure 3.2), including a kinetic phase transition from no transport to finite net transport through spontaneous symmetry breaking of the rotational symmetry.

(a) Disordered state  
(b) Ordered state

Figure 3.1: The system undergoes a phase transition from no transport to finite net transport, reproduced from [3]
3.1 Model definition

The model describes the overdamped dynamics of a collection of $N$ self-propelled particles (SPP) moving in a two-dimensional plane, characterized by their position $r_n(t)$ and direction of motion given by the angle $\theta_n(t)$, such that $v_n(t) = v_0 e^{i \theta_n(t)}$. All particles move with the same constant speed $v_0$, according to the time-discrete dynamics

$$r_{n}^{t+\Delta t} = r_{n}^{t} + \Delta tv_{0}e^{i \theta_{n}(t+\Delta t)}$$  \hfill (3.1)

Particles tend to align their direction of motion with the one of their local neighbours, and $\theta_{n}^{t}$ depends on the average direction of all particles ($n$ included) in the spherical neighborhood $U_n$ of radius $R_0$ centered on $n$. This alignment is hampered by a white noise term, which can be introduced in two different ways giving very different results.

Intrinsic (scalar) noise This was the original algorithm proposed by Vicsek et al. The direction of motion is defined by a single angle $\theta_n^t$, with $v_n^t = v_0(\cos \theta_n^t, \sin \theta_n^t)$, and the orientation dynamics may simply be written as

$$\theta_{n}^{t+\Delta t} = \text{Angle}[V_n] + \eta \xi_{n}^{t}$$  \hfill (3.3)
where $\xi^t_n$ is a zero average, delta-correlated (white) scalar noise

$$\langle \xi^t_n \rangle = 0, \quad \langle \xi^t_n \xi^k_j \rangle \sim \delta_{ik} \delta_{nj}$$

(3.4)

uniformly distributed in $[-\pi, \pi]$. Thus $\eta = 1$ is the largest meaningful noise, and results in a collection of independent random walkers. The dynamics are synchronous, meaning that all particles’ positions and headings are adjusted at the same time.

In equation 3.2 the function ”Angle” returns the angle defining the orientation of the average vector $V_n(t)$, and $k_n(t)$ is the number of particles within the spherical neighbourhood (including the n-th particle itself.

In the intrinsic noise case each particle receives the signal sent by the neighbors perfectly, but then it may decide to do something else and move in a different direction. Thus, the intrinsic noise comes from the ”free will” of the particles, so to speak; namely, from the uncertainty in the particle’s decision mechanism.

Extrinsic (vectorial) noise The second interaction rule, proposed by Grégoire and Chaté in [16], is given by

$$\theta^{t+\Delta t}_n = \text{Angle}[V_n + \eta \xi^t_n]$$

(3.5)

where $V_n$ is again the average velocity including the n-th particle, and $\xi$ is a random unit vector, delta-correlated in time and in the particle index.

The extrinsic noise consists in that the signal received by the particle is blurred (because, say, the environment is not completely transparent and the particle cannot see its neighbors very well). The consequence is the same as in the intrinsic case, namely that the particle may move in a different direction to the one dictated by the neighbors. However, in contrast to the intrinsic noise which came from the ”free will” of the swarming animals, extrinsic noise can be thought of as produced by a blurry environment.

A key observation is that while the intrinsic (scalar noise) is independent of the degree of local order, the vectorial noise becomes weaker as
Table 1: Parameters in the Vicsek model

<table>
<thead>
<tr>
<th>Raw parameters</th>
<th>Effective parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$ Number of particles</td>
<td>$\rho = \frac{N}{L^2}$ Density</td>
</tr>
<tr>
<td>$L$ Size of the box</td>
<td>$L' = \frac{L}{r_0}$ Size of the box relative to the interaction radius</td>
</tr>
<tr>
<td>$v_0$ Particle speed</td>
<td>$K = \rho(\pi r_0^2)$ Average number of interactions per particle</td>
</tr>
<tr>
<td>$r_0$ Radius of interaction vivinity</td>
<td>$l = \frac{(v_0 \Delta t)}{r_0}$ Step size relative to interaction radius</td>
</tr>
<tr>
<td>$\Delta t$ Integration time-step</td>
<td>$\eta$ Noise intensity</td>
</tr>
<tr>
<td>$\eta$ Noise intensity</td>
<td>$\eta$ Noise intensity</td>
</tr>
</tbody>
</table>

The local order is increased. If the vectors involved in the summation of neighbouring velocities (numerator of Eq. 3.5) all point in the same direction, the unitary noise vector will be much less effective than if they were pointing in different directions.

The parameters involved in simulating the Vicsek model are summarized in table 1. Although one can reduce the number of effective parameters to five, a numerical exploration of this parameter space is still very difficult, and what is valid for one region may not be valid for another region. This is the main reason why the nature of the transition is still not entirely clear (as will be further explained in section 2.3), as in some regions (low speeds and low particle speeds) it appears continuous, while in others it may appear discontinuous [17]. Of particular importance is $l = \frac{(v_0 \Delta t)}{r_0}$, the step size relative to the interaction radius. When $l$ is small ($l < 0.5$) we will say that the system is in the low-velocity regime. Otherwise, the system will be in the high-velocity regime.

In order to study the spontaneous symmetry breaking, we measure the amount of order in the system at any moment using the following polar order parameter:

$$\psi_\eta(t) = \frac{1}{Nv_0} \left| \sum_{n=1}^{N} v_n(t) \right|$$

(3.6)

where the subscript $\eta$ indicates the dependence on the noise intensity. The system shows a transient time, after which it reaches a steady state. In the
stationary state, we can average $\psi_\eta(t)$ over noise realizations, or equivalently, over time. The stationary order parameter $\psi(\eta)$ is thus defined as

$$\psi(\eta) = \lim_{t \to \infty} \langle \psi_\eta(t) \rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T \psi_\eta(t) dt$$

(3.7)

### 3.1.1 Algorithmic implementation

Although the update algorithm is fairly simple, a naive implementation would entail testing all $(i,j)$ pairs of particles, an operation which would scale as the square of the number of particles $O(N^2)$. This approach would quickly become unmanageable for system sizes of more than a few thousand particles, which is a major problem if we are interested in asymptotic (i.e. long time and large $N$) properties.

The way around this problem can be found using techniques originally developed in molecular dynamics, namely the linked-list cell MD algorithm. The idea is rather simple, even if its algorithmic interpretation may not be so straightforward. We divide the full box $L^2$ in boxes of linear size $r_0$, and assign at each timestep each particle to a given box. Once this is done, it is clear that we only have to compute the interactions with the particles belonging to the neighbouring 9 boxes (Figure 3.3). The particles belonging to each cell are organized using linked lists, using two data structures: the 'head', where head $[c]$ holds the index of the first atom in the c-th cell, and a second array in which each element $i$ points to the position of the next particle in the cell. At any fixed total density, the mean number of particles contained in these boxes does not grow with $N$, so that the number of operations needed to find all the interacting couples grows only linearly with $N$. Since assigning particles to boxes is also an order $N$ operation, it is immediate to conclude that the computational time is of order $N$ rather than $N^2$ like the naive algorithm.

![Figure 3.3: Linked-list cell algorithm](image)
3.2 Nature of the phase transition

The appearance of a long range ordered phase (collective motion) is at first surprising, given that the Mermin-Wagner theorem states that no system with short-range interactions can break a continuous symmetry in 2D and achieve long-range order \[15\]. A classical example is the XY model (in \(d=2\)), where only *quasi long range order* can be achieved, as the order parameter decays algebraically but with a very small exponent. Of course, the Mermin-Wagner theorem applies to equilibrium systems, whereas the Vicsek model is out-of-equilibrium. In fact, it can be shown with simplified arguments \[15\], that it is because the orientation fluctuations are coupled to motion (orientation information can spread both through diffusion and particle motion), that effective long-range interactions appear and there can be a true phase transition.

As regards the nature of the transition, for the extrinsic noise case it is indisputably discontinuous. However for the intrinsic noise there has been an ongoing dispute as to the nature of the transition. Figure 3.5 shows the results of simulations I have carried out for intrinsic noise, in systems of increasing size for a fixed density. As noted in \[16\], phase transitions occur in the limit \(N \to \infty, L \to \infty, \rho = \text{const}\).

![Figure 3.4: Constant density \(\rho = 4.0\), phase transition for increasing system size](image)

Figure 3.4: Constant density \(\rho = 4.0\), phase transition for increasing system size
It is clear from this figure that the transition remains continuous for increasing size, at least in the low velocity regime (here \( l = \frac{v_0 \Delta t}{v_n} = 0.03 \)). However, this behaviour cannot be generalized immediately for the whole parameter space. In \([18]\) it was shown that in the high-velocity regime (\( l = 0.5 \)), for very large system sizes the order parameter undergoes an apparently discontinuous transition. Such a discontinuity is confirmed by the Binder cumulant \( G(\eta) \) \([16]\):

\[
G(\eta) = 1 - \frac{\langle \psi_\eta^4 \rangle}{3\langle \psi_\eta^2 \rangle^2}
\]  

(3.8)

The Binder cumulant can intuitively be interpreted as a measure of how much a probability distribution function (PDF) differs from a Gaussian. The idea behind this measure is that for a first order phase transition the PDF \( P(\psi_\eta) \) should be bimodal, and therefore the Binder cumulant should have a sharp minimum at the transition. This minimum is indeed observed, which constitutes a strong argument for the discontinuity of the transition. However, in a recent paper \([19]\), Nagy, Daruka and Vicsek pointed out the formation of travelling bands in the high velocity regime. These bands always wrap around tightly in a periodic direction, which suggests they are an artifact of the toroidal topology induced by the periodic boundary conditions. In \([17]\), they have performed numerical simulations where they “shift” the boundary conditions to break the periodicity, and found that the discontinuity disappeared. Moreover, Albano et al. \([20]\) have also performed extensive finite-size scaling analysis and dynamic scaling analysis yielding results such as a hyperscaling relationship which are consistent with what is known for second order phase transitions in equilibrium systems. During this work, therefore, I will assume the intrinsic noise version of the Vicsek model gives a continuous transition analogous to second order phase transitions in equilibrium.

The question as to whether the phase transition is first order or second order is not just of academic interest. If the transition were indeed of second order, the individual fish forming a school could adjust their parameters to put the whole group close to the phase transition. Near the critical point the correlation length diverges, thus making it possible for the entire school to respond collectively to an attack. However, if the phase transition were first order, this would not be possible as the correlation length typically remains finite.
3.3 Networks as mean-field representation of interactions

As mentioned in the previous section, the XY model (where neighbouring spins tend to align their orientations) does not have a long-range order-disorder transition. It is the motion of the particles in the Vicsek model, which generates long-range interactions between these, that allows the system to undergo a phase transition. In [21] Aldana et al. propose a network model in an attempt to capture this behaviour. It is a 2D version of the original Watts-Strogatz small-world model, where the nodes start out on a lattice connected to \( K \) neighbors. These connections are then rewired randomly in order to lose all spatial correlations. As in the Watts-Strogatz model, this rewiring is controlled by the rewiring probability \( p \).

The network model can be considered analogous to the mean-field version of Vicsek’s model, and equivalent when \( p = 1 \) (where all spatial correlations are lost). This situation also corresponds to the limiting case of very large particle speeds (\( l >> 1 \)), because the noise in the direction of each particle destroys all spatial correlations in one time step.\(^1\)

\[ \text{Figure 3.5: (a) Schematic representation of long-range spatial correlations created due to motion; (b) Model used by Aldana et al. to capture these long-range interactions (figures from [17])} \]

\(^1\)The mean-field network approach neglects the coupling between local density and global order, as well as the generation of spatial structures such as clusters of particles, and has been criticised on this basis. Indeed, the mean-field approach assumes the system is spatially homogeneous
4 Vicsek model on networks

Instead of the 2D Vicsek model with local interactions in space, from here on we will study the model considering each particle to be situated on a node in a network of social interactions. To define the orientation dynamics, we will use the adjacency matrix. The adjacency matrix is a square \((N \times N)\) matrix such that each element \(A_{ij}\) is one when there is an edge from vertex \(i\) to vertex \(j\), and zero when there is no edge. The diagonal elements of the matrix are all zero, since we only consider simple graphs, which do not allow edges from a vertex to itself (loops).

Given a network characterized by its corresponding matrix \(A_{ij}\), the intrinsic noise scheme becomes (taking \(v_0 = 1\), that is, unitary velocity vectors):

\[
\theta_i(t + \Delta t) = \text{Angle} \left[ \sum_j A_{ij} v_j + v_i \right] + \eta \xi_i^t \tag{4.1}
\]

\[
v_i(t + \Delta t) = (\cos(\theta_i(t + \Delta t)), \sin(\theta_i(t + \Delta t))) \tag{4.2}
\]

The extrinsic noise implementation, on the other hand, will take the form

\[
v_i(t + \Delta t) = \frac{\sum_j A_{ij} v_j + v_i + \eta \xi_i^t}{\left\| \sum_j A_{ij} v_j + v_i + \eta \xi_i^t \right\|} \tag{4.3}
\]

4.1 Alternative expression of the order parameter

For the purposes of this work it will be useful to express the order parameter \(\psi = \frac{1}{\sqrt{v_0}} \left| \sum_{i=1}^{N} v_i(t) \right|\) in a different manner. Instead of measuring order in terms of the absolute value of the average velocity, we would like to find an expression in terms of the orientational angles of the individual particles. Inspired by the expression for the interaction Hamiltonian in the XY model \((H \sim \cos(\theta_i - \theta_j))\), we proposed the following order parameter

\[
\phi = \frac{\sum_{i=1}^{N} \cos(\theta_i - \bar{\theta})}{N} \tag{4.4}
\]

where \(\bar{\theta}\) is the average orientation. Interestingly, \(\phi\) and \(\psi\) were found to be entirely equivalent. This is straightforward to see. If we take \(v_0 = 1\),
the velocity vectors $v_i$ can be expressed as $(\cos(\theta_i), \sin(\theta_i))$, and so $\psi$ can be written in terms of the angles $\theta_i$ as

$$\psi = \left| \left( \sum_i \cos(\theta_i), \sum_i \sin(\theta_i) \right) \right|$$ (4.5)

On the other hand, expression 4.4 can be expanded using the cosine angle difference identity:

$$\phi = \frac{\sum_i^N \cos(\theta_i - \bar{\theta})}{N} = \frac{1}{N} \left( \sum_i \cos \theta_i \cos \bar{\theta} + \sum_i \sin \theta_i \sin \bar{\theta} \right)$$ (4.6)

The next step is to find $\cos(\bar{\theta})$ and $\sin(\bar{\theta})$ in terms of the individual angles $\theta_i$. Given that $\bar{\theta}$ is defined as the angle corresponding to the average velocity

$$\bar{\theta} = \text{Angle}[V] = \text{Angle} \left[ \frac{1}{N} \left( \sum_i^N \cos(\theta_i), \sum_i^N \sin(\theta_i) \right) \right]$$ (4.7)

From here we can find $\cos(\bar{\theta})$ and $\sin(\bar{\theta})$ as the first and second component respectively of the normalized vector $\frac{V}{\|V\|}$. Finally, introducing these into 4.6 it is possible to reach the same expression as in 4.5, that is

$$\phi = \left| \left( \sum_i \cos(\theta_i), \sum_i \sin(\theta_i) \right) \right| = \psi$$ (4.8)

As will be seen in the following sections, although $\phi$ is equivalent to the original Vicsek order parameter, it will be very useful in some situations.
4.2 Random networks

4.2.1 Phase transition

For random graphs the most important parameter in generating the network is its connectivity \( \langle k \rangle \). Of course, if we want the graph to be fully connected, \( \langle k \rangle \) has to be larger than a certain threshold. Once we have correctly generated the random graphs, we can relate \( \langle k \rangle \) to the number of interactions per particle as defined in table 1 for the 2D Vicsek model in space. That is, a certain \( \langle k \rangle \) will correspond to a density \( \rho \) through the relation \( K = \rho (\pi r_0^2) \). In fact, this is the idea behind using network models as a mean-field vision of the system, only taking into account random interactions and forgetting about the spatial correlations.

Figure 4.1 compares the phase transition with intrinsic noise for \( \rho = 4.0 \) and the approximately corresponding (i.e. same number of interactions) random graph with \( \langle k \rangle = 12 \). For the random graph the finite-effects are much less important, and we observe a smooth continuous transition for \( \langle \psi \rangle \).

Figure 4.1: Phase transition for systems with \( \rho = 4.0 \), and for random graph with \( \langle k \rangle = 12 \)

\(^2\)The symbol \( \langle ... \rangle \) denotes averages over noise realizations, and \( [... \) over network realizations. One has to bear in mind that a set of parameters defines an ensemble of networks with similar statistical properties.
4.2.2 Effect of connectivity on phase transition

Using the alternative order parameter $\phi$ introduced in section 3.1, it is possible to find the exact expression for $\psi(\eta)/\phi(\eta)$ in the limit of a fully connected graph, that is, in a graph where everyone is connected to everyone else. In this situation, each agent receives as a signal the mean orientation of the whole system, and the group becomes a collection of random fluctuations around the same common value:

$$\phi = \frac{1}{N} \sum_{i=1}^{N} \cos(\theta_i - \bar{\theta}) = \frac{1}{N} \sum_{i=1}^{N} \cos(\eta \xi_i) \quad (4.9)$$

As $\xi_i$ is delta-correlated in time and in the particle index, we can treat it as a continuous random variable in the range $[-\pi, \pi]$, and therefore we have:

$$\left\langle \psi \right\rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(\eta \xi) d\xi = \frac{\sin(\eta \pi)}{\eta \pi} \quad (4.10)$$

This is illustrated in Figure 4.2. As the average degree $\langle k \rangle$ is increased, the curve $\left\langle \psi \right\rangle$ resembles more and more solution 4.10, which is drawn in red. It is worth noting that in the limiting case there is in fact no phase transition. Thus a larger connectivity induces a larger noise threshold, and in the limiting case of full connectivity the threshold tends to $\eta = 1$ (infinite noise).

![Approach to maximum connectivity in random graphs of size $N = 10^4$](image)

Figure 4.2: Approach to maximum connectivity in random graphs of size $N = 10^4$
4.2.3 Ordering time

Besides the noise-induced phase transition, we have also studied the ordering dynamics without noise. Starting from random initial conditions for the orientations, we can study how long it takes the system to reach the fully ordered state. This can be done by defining an upper threshold for the order parameter.

In the disordered state what we have is a collection of $N$ randomly oriented velocities. $\psi$ in this state is the modulus of the sum of $N$ independent random vectors, thus by the central limit theorem $\psi \sim \frac{1}{\sqrt{N}}$. $\psi$ will rise up from the disordered state ($\psi \sim \frac{1}{\sqrt{N}}$) and reach the threshold after a given number of steps.

Figure 4.3 shows averages carried out over different network realizations and different repetitions on each network (starting from differing initial random conditions).

![Figure 4.3: Ordering time on ER N=1000, from $\langle k \rangle = 5$ to 20, 20 networks and 20 repetitions on each network](image)

The ordering time, as one would expect, falls when increasing $\langle k \rangle$. In the
limit $\langle k \rangle \to N$ the system should only need one step to reach the uniformly oriented state. Intuitively one could think that the ordering time falls with $\langle k \rangle$ in a similar fashion to the average path length. If the nodes are more closely packed together, this will enhance the transfer of information. To test this, we use the breadth-first search algorithm (described in Section 2.4) to measure the average path lengths in the same range as in Figure 4.3. Although the average path length also falls with $\langle k \rangle$, when we plot the ordering time as a function of this path length the fit isn’t very good (Figure 4.4a). We obtain a much better fit in Figure 4.4b, using the spectral gap of a linear version of the model. The idea behind the linear model is to define a process on the network similar to the Vicsek dynamics (equation 4.1) but without imposing normalization. That is, we can define the process

$$v_i(t + \Delta t) = \sum_j A_{ij} v_j + \frac{v_i}{k_i + 1}$$ (4.11)

Where $v_i$ are now scalar and at each time step they are updated to the average of themselves with their neighbours. We can write this process as a linear operator acting on $\{v_i\}$, in the following manner

$$v(t + \Delta t) - v(t) = \frac{d}{dt} = A_{ij} - \frac{k_i \delta_{ij}}{k_i + 1} v(t) = L v$$ (4.12)

Given that the dynamics is linear, one can study the relaxation time, i.e. the time it takes the system to average out the differences and reach a uniform state. It is straightforward to see that 0 is always an eigenvalue of $L$, and so the relaxation time will be governed by the smallest non-zero eigenvalue (also termed the spectral gap). Specifically, one can predict the ordering time from the inverse of the absolute value of this first eigenvalue

$$\tau \sim \frac{1}{|\lambda_s|}$$ (4.13)

To test whether the theory indeed works, we evaluate the spectral gap for ER graphs with the same connectivities as in figure 4.3. To do the spectral decomposition one must work with adjacency matrices in sparse form, where only the non-zero elements are saved (otherwise the calculations would be too lengthy).

As we can see in Figure 4.4b, when the prediction $\frac{1}{|\lambda_s|}$ is plotted against the ordering time we have a much better fit than, for example, with the path
length (figure 4.4a) \(^3\).

\[3\]

Although in this case the linear model works, in the heterogeneous networks it breaks down and doesn’t give good results.

Figure 4.4: Comparison between path length and inverse eigenvalue as predictors of the ordering time
4.3 Small-world networks

4.3.1 Small-world effect

As explained in section 1.2.2, the Watts-Strogatz or small-world algorithm consists in interpolating between a regular ring-type lattice and a random graph. The initial graph, where each node is connected to its \( k \) nearest neighbours, shows large spatial correlations between the nodes, whereas in the \( p \to 1 \) limit all spatial correlations are lost. This was used by Aldana et al. [21] to mimic the appearance of long-range interactions in the system. Although they used as starting point a 2D lattice instead of a ring, their results were similar to Figure 4.5, where I have plotted the phase transition for different values of \( p \). As expected, the appearance of long-range interactions (\( p \to 1 \)) favours an ordered system, increasing \( \eta_{\text{crit}} \). Another way of seeing the effect is by plotting the order parameter as a function of \( p \) for different values of the noise amplitude. In every case there is a threshold in the rewiring probability above which the system can sustain order. This threshold becomes larger as the noise is increased (Figure 4.6).

![Figure 4.5: Phase transition for WS graphs with \( N = 10^3 \), \( p = 0.15, 0.24, 0.5, 1.0 \)](image-url)
4.3.2 Ordering time

As in the random graph case, we have studied the ordering time without noise for networks with growing probability of rewiring $p$. As seen in Figure 4.7a, it falls sharply with $p$, as the number of shortcuts is increased. We repeat the same process as before, and calculate the average inverse eigenvalue for $L$ in the same range of $p$, and again plot it against the ordering time obtaining a good linear fit (Figure 4.7b).
(a) Average ordering time on WS graphs, N=2500, ⟨k⟩ = 8

Figure 4.7: Ordering time as a function of p, and the corresponding eigenvalue prediction

(b) Plot of average inverse of smallest eigenvalue against ordering time
4.4 Scale-free networks

4.4.1 Effect of scale-free topology on $\eta_c$

The divergence of the fluctuations $\langle k^2 \rangle$ for degree distributions of the form $k^{-\gamma}$ ($-2 < \gamma \leq 3$) has been shown to have important effects on the dynamical processes taking place on these networks. For example models in epidemiology predict that this kind of structure leads to a vanishing threshold of infection [22].

Figure 4.8 compares the phase transition on a Barabási-Albert network with homogeneous graphs possessing the same average degree. We can see that the BA network has a considerably higher value of $\eta_c$ for the same connectivity.

Figure 4.8: $\langle \psi(\eta) \rangle$, each point averaged over 20 different networks after 500 equilibration steps, all networks $N = 10^4$

Because the disordered state doesn’t have a vanishing order parameter (systems are finite, and $\psi \sim 1/\sqrt{N}$) it is not clear where the transition actually takes place. Therefore we need a method for determining $\eta_c$. Because the transition is second-order, we can characterize the critical point by diverging fluctuations (in the $N \to \infty$ limit). In this case we will use the variability of the order parameter at equilibrium on the various networks

$$\Delta = \frac{\sqrt{\langle \psi^2 \rangle - \langle \psi \rangle^2}}{\langle \psi \rangle}$$  \hspace{1cm} (4.14)
This quantity is akin to a susceptibility, so we expect it to show a peak in the vicinity of the critical point.

Figure 4.9 shows a plot of the variability $\Delta$ for AB networks of size $10^4$ and $10^5$. We can see both peak at the same point, but the peak is much higher for $N = 10^5$ (due to the fact that in 2nd order phase transitions it diverges as $N \to \infty$). Given that $\eta_c$ increased when switching to a heterogeneous network, one could expect that maybe $\eta_c$ is an increasing function of the degree variance $\langle k^2 \rangle$. However in BA graphs it would be hard to discern an increase in $\eta_c$ when increasing the network size $N$, as $\langle k^2 \rangle \sim \log(N)$ (only grows logarithmically with network size). To see the effect of increasing $N$ we will need scale-free graphs with $2 < \gamma < 3$, which we will generate using the UCM algorithm.

![Figure 4.9: variability $\Delta$ of the order parameter at different values of the noise, $N = 10^4$ and $N = 10^5$ BA graphs](image)

### 4.4.2 Finite-size scaling

For scale-free networks with an exponent $2 < \gamma < 3$, there will be important finite-size scaling effects, given that

$$\langle k^2 \rangle \sim N^{\beta(3-\gamma)}$$

(4.15)

where in our case $\beta$ will be $\frac{1}{2}$. We would expect $\eta_c$ to grow for networks of increasing size $N$, and converge to a $\eta_c(\infty)$, the critical noise in the ther-
modynamic limit which will depend on the exponent $\gamma$. This is precisely what was observed. For three different values of $\gamma$ (2.01, 2.25 and 2.5), the variability $\Delta$ was studied for networks of increasing size. As $N$ is increased $\eta_c$ is displaced towards higher values of noise amplitude (Figure 4.10).

![Graph (a) $\gamma = 2.5$](image_a)

![Graph (b) $\gamma = 2.01$](image_b)

Figure 4.10: the peak of the variability $\Delta$ moves towards higher noise

To figure out the dependence $\eta_c(N)$, we take $\eta_c(N_i)$ as the maximum of $\Delta(\eta)$ for each size. As is common in the analysis of equilibrium systems when applying finite-size scaling theory [23], the dependence $\eta_c(N)$ is then fitted to a function of the form

$$\eta_c(N) = \eta_c(\infty) - aN^{-b} \quad (4.16)$$

where $\eta_c(\infty)$, $a$ and $b$ all depend on $\gamma$. The results of the fits for $\eta_c(\infty)$ were: 0.8053 ($\gamma = 2.01$), 0.9616 ($\gamma = 2.25$), 1.028 ($\gamma = 2.01$) and are plotted in Figure 4.11 (linear scale) and 4.12 (double logarithmic scale). We can see that as the network becomes more and more scale-free ($\gamma \to 2$) it becomes more resilient to noise. For the $\gamma = 2$ case, in the $N \to \infty$ (thermodynamic) limit $\eta_c \to 1$, meaning the transition occurs at 'infinite' noise.
Figure 4.11: Critical noise $\eta_c$ as a function of $N$ in linear scale

Figure 4.12: $\eta_c(\infty) - \eta_c(N)$ on a double logarithmic scale
4.4.3 Hierarchy of the order parameter

To investigate the role played by the high-degree nodes in a scale-free distribution, we have studied how the order in the system is distributed among the differently connected nodes. At first it seemed reasonable to study this by defining

$$\psi_k = \frac{\sum_{i \in S_k} v_i}{N_k}$$  \hspace{1cm} (4.17)

where $S_k$ is the subset of nodes with degree $k$. However, this measure of order doesn’t give us much information because it is impossible to distinguish genuine order from the effect of having fewer nodes of a given degree. In fact, Equation 4.17 is equivalent to

$$\sum_{i \in S_k} \cos(\theta_i - \bar{\theta}_k)$$

where $\bar{\theta}_k$ is the average orientation of nodes of degree $k$. This is not the measure we want, as we would like to compare the individual orientations to the average of the whole group. In other words, we will use

$$\phi_k = \frac{\sum_{i \in S_k} \cos(\theta_i - \bar{\theta})}{N_k}$$  \hspace{1cm} (4.18)

which can be interpreted as the average 'error' the particles make relative to the group orientation. The results of this analysis are shown in figure 4.13. As the level of noise is increased, the low-degree nodes fall into disorder, and it is the high-degree nodes who sustain the order in the system. They are less susceptible to the noise thanks to the signal they receive, the better quality of which we can relate to the 'many wrongs' concept [1]. It is the result of integrating many errors, because it is a sum over the velocities of many more particles than their low-degree counterparts. High-degree nodes effectively act as leaders and allow the system to sustain more noise.

4.4.4 Ordering time

Using the same method as in Sections 3.2.3 and 3.3.2, the effect of increasing the network size $N$ on the ordering time without noise has been studied. In figure 4.14, $\tau_{BA}(N)$ is compared to $\tau_{ER}(N)$, the equivalent measure but for random graphs. Both grow logarithmically with network size, but for the BA networks the growth is slower and the ordering time is always lower. This is reminiscent of the average path length on both types of networks, which
grows logarithmically but with a double logarithmic correction ($\sim \frac{\ln(N)}{\ln\ln(N)}$) in the BA case, making it an even more compact structure than the random graph.

These results are without noise, but one could also ask what effect the noise has on the ordering dynamics. Figure 4.15 shows the relative time it takes for nodes of degree $k$ to reach a given threshold (in this case $\phi = 0.9$).
Whether in the presence of noise or not, low-degree nodes take longer to reach the ordered state, as one would expect. Interestingly, however, when the noise is introduced it affects differently the low-degree nodes and the high-degree nodes. The low-degree nodes take longer when there is noise, but there is an opposite effect on their high-degree counterparts. For them the presence of fluctuations speeds up the process.

Figure 4.15: Dependence of ordering time on degree, done for BA networks $m = 3, N = 10^5$

4.4.5 Hysteresis for extrinsic noise

Although in this work we have mostly focused on intrinsic noise, which gives a continuous transition, we have also studied some properties of the extrinsic version of the algorithm. As explained previously, extrinsic or vectorial noise can be interpreted as being the result of a blurry environment (as opposed to an uncertainty in the decision mechanism), and gives a discontinuous transition. A common property in discontinuous or first-order transitions is the existence of hysteresis loops, where the output of a system depends on its past history. This is also observed in the Vicsek model with extrinsic noise, as shown in figure 4.16, where we have run the model on a BA network of size $N = 10^4$. To study hysteresis, we increase $\eta$ by $\Delta \eta$ every $N_{\text{steps}}$ time steps. The way to ascertain the presence of hysteresis is by varying a parameter we may define as a ‘velocity’: $v = \frac{N_{\text{steps}}}{\Delta \eta}$. If the number of steps is too low we may simply be not allowing the system to reach equilibrium, but if we increase $N_{\text{steps}}$ up to a high enough value we can make sure that this is not
the case. In figure 4.16 the first loop is drawn with only 200 equilibration steps at each noise, but then $N_{\text{steps}}$ is increased to $10^3$ and the loop is still observed.

What this means is that the critical noise is higher if we start from an initially ordered state. In other words, it is harder to disorganize an already ordered state, which is probably related to the observation (already made in the introduction) that the vectorial noise is more or less effective depending on the local order.

Is anything like this observed in real collective motion? In nature, groups are likely to move between collective states as conditions change, and as a consequence of this the previous history of individual orientations and positions can have an influence on the collective behavior as behavioral parameters change. That is, the transition between behavioral states depends on the previous history (structure) of the group, even though the individuals have no explicit knowledge of what that history is (thus there is a ‘collective memory’). Couzin et al. [24] argue that this may be a generic property of transitions between collective behaviors.

Figure 4.16: Hysteresis loops for increasing $N_{\text{steps}}$ on BA network $m = 3, N = 10^5$. 
4.5 Weighted networks

In all the network models we have used up to this point we have considered all edges to be identical. This is an over-simplification, as in a real-world network different connections may have different strengths, the intensity of the social tie depending on factors such as kinship.

A way of modeling this is by defining weights such that $w_{ij} = w_0 (k_i k_j)^\alpha$ (where $\alpha$ represents the strength of the correlation). This is the weight-degree correlation observed in several networks such as the WAN (World-wide Airport Network). [25]

We can introduce this specific form of $w_{ij}$ into the general Vicsek dynamics and reach the following expression for the intrinsic noise implementation:

$$\theta_i(t + \Delta t) = \text{Angle} \left[ \sum_j A_{ij} k_j^\alpha v_j + k_i^\alpha v_i \right] + \eta \xi_i^t \quad (4.19)$$

$$v_i(t + \Delta t) = (\cos(\theta_i(t + \Delta t)), \sin(\theta_i(t + \Delta t))) \quad (4.20)$$

And the extrinsic noise algorithm becomes:

$$v_i(t + \Delta t) = \frac{\sum_j A_{ij} k_j^\alpha v_j + k_i^\alpha v_i + (k_i^\alpha + A_{ij} k_j^\alpha) \eta \xi_i^t}{\left\| \sum_j A_{ij} k_j^\alpha v_j + k_i^\alpha v_i + (k_i^\alpha + A_{ij} k_j^\alpha) \eta \xi_i^t \right\|} \quad (4.22)$$

Where we have simplified the $k_i^\alpha$ term, the weight of each node’s own velocity is given by $w_{ii}$, and $w_0 = 1$ in order to recover the unweighted dynamics for $\alpha = 0$. The idea behind the weighting is that each particie will give more or less importance to the information coming from its neighbours depending on this neighbour’s degree.

The simulations in this section will be carried out on BA scale-free graphs, as we will need degree heterogeneity to see the effects of the weighting.

4.5.1 Dynamics without noise

The effect of the exponent $\alpha$ is clear in the two limit cases.

When $\alpha \to \infty$, low-degree nodes copy their most connected neighbour, while the high-degree nodes simply copy themselves. Depending on the particular structure of the network the system may or may not be able to reach
the uniformly oriented state. If the large-degree nodes aren’t connected be-
tween them, there may be ‘jamming’ and the information won’t be able to
spread.

In the $\alpha \to -\infty$ limit, on the other hand, it is even clearer that regardless
of the particular graph it will end up disintegrating as the lower-degree nodes
(the majority) simply copy themselves and do not change over time.

Figure 4.17 corroborates this picture. The time required to reach order
increases for both very high (positive) values of $\alpha$ and for negative values. We
can also see that for very positive values there is a high variance, indicating
the dependence on network structure mentioned above.

Figure 4.17: Figure obtained for BA graphs of size $N = 10^3$ $m = 4$, with
weighting exponents ranging from $\alpha = -1$ to $\alpha = 6.0$. Each point obtained
from averaging over 150 repetitions on 150 networks

4.5.2 Optimal exponent

Figure 4.18 shows the order parameter at equilibrium for different values of
the weighting exponent, for BA graphs of size $N = 10^4$ in the presence of a
noise amplitude $\eta = 0.1$. We can observe that it decreases faster for negative
$\alpha$’s than for positive $\alpha$’s.
Figure 4.18: Order parameter at equilibrium for noise $\eta = 0.1$ and $\alpha$ ranging from -3 to 6. Figure obtained for BA graphs size $N = 10^4$, averaging over 8 networks and 3 repetitions on each network for each point.

This can be visualized drawing the network using the Python package Networkx. In figure 4.19 a BA graph with $N = 100$ nodes and $m = 3$ is represented. From this unweighted network, we can increase or decrease the exponent $\alpha$, thus varying the strength of the connections according to $\sim (k_i k_j)^{\alpha}$. Some of these connections will become weaker and weaker until they play no role in the interaction dynamics. Having defined a threshold, we can go through all nodes $i$ and eliminate neighbours $j$ from their adjacency list if $k_j^{\alpha}$ is below the threshold.

Figure 4.19: Initial BA network before weighting, $N = 100$, $m = 3$

Figure 4.20 shows the result of this process for $\alpha = 2.0$ and $\alpha = -2.0$. In
The $\alpha = 2.0$ case, the connections that are left all point towards high-degree nodes. For $\alpha = -2.0$ the network is more sparsely connected, thus negative weighting has a more harmful effect on the ordering. This can explain the asymmetry found in the ordering times (Figure 4.17).

![Diagram](image)

(a) $\alpha = 2.0$  
(b) $\alpha = -2.0$

Figure 4.20: Effect on network for weighting exponents $\alpha = 2.0$ and for $\alpha = -2.0$, deleting connections below a strength threshold of 0.1

What perhaps is more interesting here, however, is that there seems to be a maximum in $\psi(\alpha)$. When the simulations were repeated for higher values of the noise in a narrower interval $[-1, 1]$ (Figure 4.21), we found that indeed there is always an optimal weighting exponent, approximately $\alpha \sim 0.2$. The existence of an $\alpha_{\text{max}} > 0$ seems reasonable, as a positive weighting exponent lends more importance in the interactions to the higher-degree nodes, whom we have shown in Section 3.4.3 to be more 'informed' or aligned.

In fact, the best way of seeing the optimal parameter is by repeating the plot of $\phi_k$ from Section 3.4.3 but now with a weight $\alpha_{\text{max}} = 0.2$ attached to the dynamics. As seen in figure 4.22, the effect of the weight is to change the shape of the curve $\phi_k$ by lifting the low- and intermediate-degree nodes, transferring information to them from the high-degree nodes.

$\phi_k$ is also plotted for an even higher $\alpha = 1.0$, but now the weighting no longer has a positive effect. The reason for this is that if $\alpha$ is too big, the system is basically subject to the random fluctuations of the few high-degree nodes. The increased variance in the signal $\psi(t)$ (figure 4.23) for high $\alpha$ corroborates this idea.
Figure 4.21: Order parameter at equilibrium for noises $\eta = 0.2, 0.5, 0.6, 0.7$ and $\alpha$ ranging from -1 to 1. Figure obtained for BA graphs size $N = 10^4$, averaging over 8 networks and 3 repetitions on each network for each point.

Figure 4.22: Relative order of nodes of degree $k$ on weighted BA networks of size $N = 10^5$, noise level $\eta = 0.65$
Figure 4.23: Increase of variance in $\psi(t)$. From top to bottom: $\alpha = 0.0$ (unweighted), $\alpha = 1.0$ and $\alpha = 3.0$. Noise level $\eta = 0.4$ in all cases.
5 Conclusions

By studying Vicsek’s model on various network topologies, we have explored how social interactions affect the transfer of directional information in collective motion. As in the original articles, we show Vicsek’s model with intrinsic noise leads to a phase transition both for off-lattice particles and on graphs, the second-order (continuous) nature of which we have provided further evidence for. In order to study the phase transition on networks, we have introduced a different expression for the order parameter, which is shown to be equivalent to the original.

On random graphs, which can be seen as a mean-field representation of the off-lattice interactions, as the connectivity is increased the order parameter $\psi(\eta)$ is shown to converge in the fully connected limit to a solution we find analytically. Because the connection topology in social networks is usually between totally random and regular, we use the small-world model to interpolate between these two extremes. We find the long-range interactions produced by the rewiring of edges is essential for the system to be able to sustain order, and the critical value of the probability of rewiring $p$ beyond which order is possible is shown to increase with the noise. In both random and small-world networks, the ordering time without noise is found to decrease with $\langle k \rangle$ and $p$ respectively. We propose an analogous linear operator containing the adjacency matrix, whose eigenvalues give a prediction for this ordering time.

On the other hand, many social networks have a scale-free degree distribution. For the same average degree, we have found a scale-free topology has a higher critical noise $\eta_c$ than homogeneous topologies, i.e. the system is more resilient. Moreover, when the exponent $\gamma$ approaches 2, the transition is found to occur at infinite noise ($\eta = 1$) in the thermodynamic limit ($N \to \infty$). This higher resilience is thanks to the role played by the high-degree nodes, which in the presence of noise can integrate many errors and sustain the long-range order of the system. Finally, weighted interactions (correlated to degree) on scale-free graphs show the existence of an optimal weighting exponent which gives the highest possible order in the system at any given noise.

Regarding future developments to this research, further understanding of collective motion will require greater consideration of the rules adopted by individuals, of how individuals interact with the environment and of between-individual differences. Experimental work is also needed in order to carry
out detailed statistical comparisons between real flocks and SPP models.
References


