Imperial College London

Department of Computing

The Dynamical Complexity of Swarms, Flocks and Shoals

by

Filipe Peliz Pinto-Teixeira

Submitted in partial fulfilment of the requirements for the MSc Degree in Computing Science/Artificial Intelligence of Imperial College London

September 2011

Abstract

In this project the natural flocking and swarming behaviour demonstrated by a variety of animal aggregations is studied. A correlation between dynamical complexity and the flocking patterns created by such animal groups is closely analysed.

For this study to be conducted various flocking and swarming models were carefully analysed in order to determine the best way to create a biologically accurate, but not overly complex flocking model.

Having developed such a model which allowed for complex behavioural control using only one parameter, the dynamical complexity of the resulting flocking patterns was measured. In order to perform such measurements various dynamical complexity measures including Interaction Complexity, Information Integration and Causal Density were analysed with the objective to determine which of these would be appropriate for establishing the correlation between flocking and dynamical complexity.

However before proceeding with the complexity analysis, a technique for measuring the components or variables of a flock had to be developed. Various measurement techniques were designed, and the results achieved from each of these measures were carefully and strenuously analysed to determine their accuracy. Having eliminated the inaccurate measurement techniques, a true method for simply measuring a flock is discovered thus reaching a conclusion about the dynamical complexity exhibited by these flocks.

List of Figures

1	Box Three Dimensional World
2	An Object's Local Rotation axis
3	Possible Bird Perceptions
4	Bird Neighbourhoods
5	Effects of Window Size
6	Small Scale: Flocking under a Leader
7	Large Flock: Flocking under a Leader
8	Causal Relations as Flock Size Increases
9	Small Scale: Average Velocity Measured
10	Small Scale: Average Position Measured
11	Small Scale: Velocity Difference Measured
12	Small Scale: Position Difference Measured
13	Small Scale: Speed Difference Measured
14	Small Scale: Difference from Home Measured
15	Small Scale: Position Difference from Main Flock Measured
16	Large Scale: Average Velocity Measured
17	Large Scale: Average Position Measured
18	Large Scale: Difference of Velocity Measured
19	Large Scale: Difference of Position Measured
20	Large Scale: Speed Difference Measured
21	Large Scale: Position Difference from Home Measured
22	Large Scale: Position Difference from Flock Measured
23	Example Flocking at Stagnation Point
24	Illustrative Example of Leader Driven Flocks
25	Leader Analysis with Windowing technique
26	The Measurement Method Weaknesses
A.1	The Welcome Screen
A.2	Simulation Window
A.3	Main Simulation Controls
A.4	Camera Properties
A.5	Flock Drives
A.6	Neighbourhood Definitions
A.7	Controls Window
A 8	About Window 66
R 1	Bandom Flocking: Flocking Factor $= 0$ 70
B 2	Natural Flocking: Flocking Factor $= 1$ 71
B.2 B.3	Stagnated Flocking: Flocking Factor $= 10$ 79
1.0	$\sim 1000000000000000000000000000000000000$

List of Tables

1	Main Experimentation Settings
2	Small Scale: Norm of Residuals
3	Small Scale: Norm of Residuals
4	Causal Density for a Simple Stagnated Flock
B.1	Flock Settings For Primary Patterns

Contents

1	Introduction	1
2	Related Work	3
	2.1 Flocking, Swarming and Shoaling Behaviour	. 3
	2.1.1 Natural Flocks and Swarms	. 3
	2.1.2 Particle Swarm Optimization	4
	21.3 The Boid Model	5
	2.1.9 The bold worder	. 0
	2.1.4 Fuzzy Flocking	. 1 0
	2.1.5 Extending Dous	. 3
	2.1.0 Other Models	11 10
	2.2 Dynamical Complexity	. 12 19
	2.2.1 Interaction Complexity \dots Δ	12 12
	2.2.2 Information Integration Ψ	. 13
	2.2.3 Causal Density	. 14
	2.2.4 Quantifying Flocking and Swarming	. 16
3	Methodology	18
	3.1 The Flocking Model	. 18
	3.1.1 The Social Forces	20
	3.1.2 Other Model Forces	22
	3.1.3 The Final Model and The Model Parameters	. <u>2</u> 2
	2.2 The Analysis Procedure	. 20 94
	3.2 The Analysis Procedure	24 95
	3.2.1 FIOIII FIOCKING TO Data	. 20 96
	3.2.2 From Data 10 Causal Density	. 20
	3.2.3 Experiments Conducted	. 27
4	Results	29
	4.1 Detecting The Leader	29
	4.2 Causal Density	33
	4.2 Causa Density	. 00 34
	4.2.1 Junan Scale Results	20
	4.2.2 Large Scale Results	. 39 14
	4.2.5 Combining The Results	. 44
5	Discussion	45
6	Limitations	50
0		00
7	Future Work	53
8	Conclusion and Summary	55
A	knowledgements	56
R	ferences	57
A	opendix	59
А	Simulation User Guide	59
	A 1 Getting Started	50
	A 2 The User Interface	. 55 60
	A 2.1 The Simulated World	. 00
	A.2.1 The Simulated Wohld	. UU
	A.2.2 The Simulation Manager	
	A.2.9 Uamera Uontrois \ldots	. 03

	A.3 A.4 A.5	A.2.4 Drive Forces	64 65 66 66 67 68
в	The B.1 B.2 B.3	3 Primary Flocking Patterns Random Flocking Pattern Natural Flocking Pattern Stagnated Flocking Pattern	69 70 71 72
\mathbf{C}	Prir	nary Code Snippet	73

1 Introduction

The aerial display of large flocks of birds is a stunning example of collective behaviour in animal aggregations [16]. Collective animal behaviours are not only demonstrated in avians but also in swarms of bees [17] as well as schools of fish [15]. This behaviour exhibited in animal groups is a fascinating subject, arising not as a result of leadership but rather through the group dynamics.

Collective behaviour in animals suggests a measure of intelligence within these groups. Collective animal behaviour is not only interesting to observe but also provides an insight into the academic methodology used in various applications to measure the underlying individual behaviour within a group, namely that each individual performs a simple task resulting in a complex collective behaviour [8].

An unavoidable question is to which animals or subsets thereof demonstrate complex group dynamics. For example a group of dogs dominated by an Alpha male is not relevant for the purpose of this discussion. The behaviour which forms the underlying premises of this study is in the so called "Intelligent Group", that is to say behavioural patterns created by the social dynamics within the group as a collective. According to [18] [21] a swarm or group needs to satisfy five principles in order to demonstrate intelligence:

- 1. The Proximity Principle A swarm should be able to do simple space and time computations. This is relatively simple as each member of a swarm must be aware of its own position at any given time.
- 2. The Quality Principle A swarm should be able to respond to it's environment based on local quality observations.
- 3. The Principle of Diverse Response The swarm should not allocate all of its resources into the same effort. Meaning that a swarm must be able to perform some sort of task distribution.
- 4. **The Principal of Stability** The swarm should not immediately or instantaneously change its behaviour.
- 5. The Principal of Adaptability The swarm must be able to change its behaviour when the change is computationally better.

Schools of fish, bee swarms and certain flocks of birds satisfy the above principles, hence why it is acceptable to focus on these animal groups.

The study of the emergent behaviour demonstrated by these groups is already being applied successfully. Examples of such applications include ant algorithms [6], particle swarm optimizations [19] as well as various simulation models [18] [25] [26].

It is known that these groups do demonstrate some coordinated social behaviour. This is obvious by simple observation.

For example when observing the flocking behaviour demonstrated by birds, the birds manage to fly quite rapidly around each other without any collisions occurring. Another example is when ants locate a food source they will all eventually take the same path to the food source, all this is achieved not due to a leader but rather the social dynamics occurring within these groups.

It is this social aspect which results in the dynamic behaviour of these groups. The purpose of this study is to better understand this dynamic behaviour by measuring the dynamical complexity demonstrated. A system is said to exhibit dynamical complexity if both integrated and segregated activity occur within the system [30]. At the same time the system must demonstrate a good balance between said activities in order to exhibit a high dynamical complexity. If a system is too integrated then it is unable to provide a large range of responses, i.e. it is inflexible. On the other hand if a system is too segregated then it is unable to gather its resources in order to provide a strong collective response. A highly unified or segregated system will have a low dynamical complexity.

In terms of the flocking, swarming or shoaling behaviour exhibited, a flock for example will be overly integrated when each member has the same velocity and moves in the same direction. On the other hand a flock will be overly segregated when each member acts independently and possibly even more randomly.

An assumption could be made that the degree of the social aspect of these swarms may have a direct correlation with the dynamical complexity exhibited. The question which this study attempts to answer is simply what is the measured dynamical complexity of the various behaviours demonstrated by these groups.

In order to answer this question three major objectives have to be met:

- 1. A model simulating the desired behaviour has to be developed.
- 2. A single parameter which allowed for complex control over the flocking behaviour has to be integrated into the model. The purpose of this parameter is to sweep the flocking patterns from random patterns to natural flocking and finally to stagnated movement.
- 3. A method of generating a set of time series from the flocking behaviour has to be designed. These time series are needed in order to conduct the dynamical complexity analysis. Most importantly is that each time series has to uniquely quantify the behaviour of each component (flock member).
- 4. Each of the methods developed has to be carefully tested in order to ensure that the time series generated accurately reflects the necessary data. This means that a series of tests and measures have to be conducted in order to ensure the validity of the final results returned.
- 5. Finally the dynamical complexity of the flocking behaviour under different conditions (set by the sweeping parameter) has to be measured.

In analysing the correlation between dynamical complexity and flocking patterns, the hope is to better understand and demonstrate how dynamical complexity is an accurate measurement of the fragile balance between segregated and integrated activity in dynamical systems.

Please note that for the rest of this study the terms flock member, swarm member and group member will be used interchangeably, even though there is a significant difference between a flock of birds and a swarm of bees, this study is primarily concerned with the collective behaviour in these animal aggregations.

2 Related Work

The following section will a describe series similar studies which are of relevance to this study. The study performed was divided into two sections, the first being creating an accurate simulation of the behaviour described in Section 1 and the second being the measurement of the dynamical complexity. As such, the study of relevant pieces of work will be divided into those 2 sections.

2.1 Flocking, Swarming and Shoaling Behaviour

2.1.1 Natural Flocks and Swarms

It is known that birds flock, bees swarm and fish swim together as a result of each member wanting to stay close to its group but at the same time wanting to avoid collisions. According to various authors such as [2], [26] and [3] this behaviour is a result of the evolutionary pressure from several factors, such as the desire to be protected from predators (strength in numbers) or the drive to optimise search patterns (e.g. bees looking for flowers to pollinate).

Interestingly enough there is no evidence which suggests that this flocking behaviour is bounded in any way. Regardless of the size of these groups the underlying behaviour remains constant. This is due to the fact that each member of the flock does not perceive the entire flock but rather a small subset, usually 2-8 other members depending on the size of the flock as well as the species [3].

The idea of each flock member only perceiving relatively few other flock members is important because it means the amount of "thinking" each member has to do is independent of the size of the group [26]. This means that the complexity of the behaviour should not increase as the flock increases. If this were not the case then it would entail that we would notice a natural maximum flock capacity (which is not the case).

Whilst in natural flocks there is no upper bound, there is in simulated flocks. Meaning that in simulated flocks there exists a maximum number of birds which can be modelled. This is due to the complexity of the flocking algorithm described being basically $O(N^2)$. That is, the work required to run the algorithm grows as the square of the flock's population [26]. This is due to a variety of factors, some of which are:

- The need to compute for each member of the group the x closest members.
- Computing the field of view as well as adjusting and updating the field of view as the bird rotates and translates.
- The need to compute the social forces each bird receives as well as the social forces each bird imposes.
- Computing the velocity which depending on the model can include computing social drives, gravitational forces, wind and aerodynamic forces.

Due to these computational constraints the model developed had to be limited and simplified, not only in terms of the group sizes but also in terms of other factors which will be discussed in Section 6.

Another significant question we have to ask is, when does flocking or swarming occur?

There is no unified reason for all species. However each animal group usually demonstrates some flocking ability under certain conditions.

For example flocks of starling birds tend to produce the most interesting and beautiful aerial displays shortly before sunset when they return to their roosting area [2]. The flocking behaviour exhibited by starlings is not only beautiful to watch but also quite complex when carefully analysed [2]. The flock demonstrates strong cohesion but at the same time, a strong group alignment meaning that they move together and also quite closely together, while avoiding collisions. This is impressive considering the speeds at which these birds move. Another example of when significant swarming occurs is when honey bees look for, and move to new nests [4]. The procedure itself can be analysed in terms of the emergent group behaviour First the nest will send out "scout" bees in order to locate acceptable nests. When an acceptable nest has been located by a scout, it will inform some of its members of the location of the nest via a dance. If the nest is truly desirable then more and more bees will go to the location. The true swarming behaviour occurs when there is a large move to the new nest. Most of the bees within the swarm will not be aware of the location of the new nest yet they are still able to move to that location quite efficiently. [20] hypothesized that once airborne, the swarm is directed toward the chosen nest site by scouts who streak rapidly through the airborne swarm cloud in the appropriate direction.

Interestingly enough it seems that the most complex flocking or swarming behaviour is usually goal driven, whether it is to locate a new home or to roost for the night. This idea of providing the flock or swarm with a biological goal was quite important in developing the model described in this project.

2.1.2 Particle Swarm Optimization

As mentioned before, a popular application of this collective intelligence is in Particle Swarm Optimizations [35] [8]. The purpose of a Particle Swarm Optimization (PSO) is to find the optimum solution to a complex problem where other approaches such as a linear search is not computationally efficient.

Essentially a PSO algorithm is a simplified application of swarming theory where each particle or member of the swarm attempts to find the optimum solution. The interesting part is that these particles adhere to a set of social rules which essentially drive and guide the search procedure.

In the basic PSO algorithm each particle is represented by an n-dimensional position; the number of dimensions vary from one optimization problem to the next. The most common PSO algorithms work as follows [35]:

- 1. Begin with an initial set of particles, typically randomly distributed throughout the problem space.
- 2. Calculate the velocity for each particle in the swarm.
- 3. Update the position of each particle using the updated velocity and the previous position.
- 4. Go back to Step 2 and repeat the procedure until the swarm converges or until some other stopping condition is met.

Each position update is driven by:

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t)$$
(2.1)

Where $x_{ij}(t)$ is the position of particle *i* in the *jth* dimension and $v_{ij}t(t)$ is the velocity of the particle.

As can be seen above the primary drive force is the velocity which is computed. PSOs are of interest to this study due to the rules which are used to compute the new velocity. These rules are primarily social rules, driven by the whole swarm, and as such reflect a form of the "collective intelligence" demonstrated by fish or birds. The velocity is computed by [8]:

$$v_{ij}(t+1) = v_{ij}(t) + c_1 r_{1j}(t) [y_{ij}(t) - x_{ij}(t)] + c_2 r_{2j}(t) [\hat{y}_j(t) - x_{ij}(t)]$$
(2.2)

 c_1 and c_2 refer to acceleration constants which are used to scale the contributions of the cognitive and social components respectively.

 $r_{1j}(t), r_{1j}(t) \sim U(0, 1)$ are used to add a stochastic element to the swarm. $y_{ij}(t)$ refers to the most desirable position for particle *i*. $\hat{y}_j(t)$ refers to the best position found by the swarm.

It is important to note that the desirability of a position is determined by a problem specific evaluation function, which for the purpose of this study is not relevant.

For the velocity calculation above there are three components which drive the swarm movement, namely:

- 1. **The Previous Velocity** Also known as the momentum. This simply refers to the particle's memory of its previous flight.
- 2. The Cognitive Component $c_1r_{1j}(t)[y_{ij}(t) x_{ij}(t)]$ Which in this case refers to the particle's own desire to stay close to the best position it has located within the swarm. An example of what could be classified as "best position" in terms of biological flocking could simply be a predefined location or "home" position.
- 3. The Social Component $c_2r_{2j}(t)[\hat{y}_j(t) x_{ij}(t)]$ Which in this case refers to the particle's desire to move close to the best position found by the swarm. Again an example of what could be seen as "best position" in terms of flocking could be the desire for a bird to move towards the mean of the flock, for the sake of safety from predators.

Notice how the social components defined above is a primary driver for the swarm and as such leads to a collective intelligence from the particles.

PSOs are a very simplified real world application of swarming theory, however PSOs are driven by the desire to solve an optimization problem and as such, while the ideas behind a PSO reflect that of the desired behaviour, the actual application of them are too problem specific and are not ideal for creating a simulation of natural swarming behaviour. Whilst it might be possible to modify the basic PSO algorithm to result in better swarming, such modifications would not result in a acceptable simulation.

2.1.3 The Boid Model

The most common or traditional approach for flocking is the Boid Model developed by Reynolds [26]. The Boid algorithm to flocking is based on the work done in particle systems [24] and as such uses some techniques and methods implemented by those systems. Particle systems are collections of a large number of dynamic "fuzzy objects", where each individual has its own behaviour based on certain properties such as colour, opacity or location.

The Boid model uses a similar approach where it also uses a large number of particles, but the primary difference is that in the Boid model each particle is replaced by a full geometric model with a geometric position and orientation. Meaning that the Boid model allows for complex geometrical objects to exist within a predefined "world" or "universe". These geometrical objects are not only more complex in terms of appearance but also in terms of behaviour.

The behaviour of these objects allows for a more natural flocking to occur due to geometric flight being incorporated into the model. The incorporation of geometric flight gives each object the ability to move and transform along and tangent to a three dimensional curve. Thus giving each object the ability to not only change position within its world but also change its orientation and shape.

Geometric flight is based on applying incremental changes and updates along the object's forward direction (positive Z axis). Whilst these translations allow for the object to move forward and backward, steering is achieved by rotating about the local X and Y axis (pitch and yaw) which in turn realign the local Z axis. Real world geometric flight is continuous which is why incremental changes are used. The application of incremental changes is a discrete approximation to continuous real world flight [26].

Each object in the model also has a velocity which allows the object to move within the world. It is important to note that in geometric flight, momentum plays a significant role. A maximum velocity must also be introduced into the model in order to ensure that the objects' flight remains within realistic speeds. The introduction of a maximum velocity as well as incorporating momentum is to constrain movement into a more realistic approximation. In essence this will ensure that each boid(object) in the model does not perform impossible movements, such as performing an immediate 180 degree turn.

Taking into account everything above, each boid or flock member in a 3D world is represented by 2 vectors. The first [x, y, z] representing its position within the global co-ordinate system (i.e. its position within the world) and the second [X, Y, Z] representing its current rotation within its local axis. A graphical representation of both the world's global co-ordinate system and object's local rotation axis is given in Figure 1 and Figure 2 respectively.





Figure 1: A simplified example of a box 3D world where the dot represents the position of the object in the world. For example a valid position for the object above could be (3, 3, -5)

Figure 2: A visual representation of an object's local rotational axis. This local axis is fixed to the object and moves with the object throughout the world's coordinate system

Geometric flight relates to translation, pitch and yaw, but does not constrain roll which refers to rotation about the object's local Z axis. This is usually used for modelling the banking of the object. Banking, which is often omitted from various models for the sake of simplicity is based on the tangential component of acceleration. With correct banking (what pilots call a coordinated turn) the object's local space remains aligned with the "perceptual" or "accelerational" coordinate system [26].

By representing a geometrical object as defined above, it is possible to create a highly accurate simulation, however such a representation can be complex and many models may choose to omit certain aspects such as the 3rd dimension, which greatly reduces complexity, but also reduces model accuracy. Omitting the roll of the object is another method of simplifying the model, and in many cases omitting the roll is acceptable. For example omitting the roll only affects an object's field of view in a minor way (i.e. the roll affects the rotation of the field of view whilst the pitch and yaw affect the direction of the field of view). In essence meaning that the roll is unlikely to affect what can be seen by the object.

The primary drive behind the Boid flocking model is the application of social rules to the flock. While the model does make use of basic gravitational forces it is these social rules that are the primary forces which make up the velocity calculation. The social rules are:

• Separation - The primary drive for the birds to fly away from each other in order to avoid

colliding with each other.

- Alignment Each bird tries to match the velocity of its neighbouring birds.
- **Cohesion** Each bird tries to stay close to its neighbours and the flock as a whole by attempting to move towards the center of the flock.

Separation being a natural desire to avoid collisions together with alignment allows birds to move freely within the flock and it is because of these two rules that one bird's movement can cause a chain of effects which impact the whole flock. Separation is based entirely on each birds position and if the distance between two birds falls too low then the separative rule will push birds away from each other. Conversely, alignment is purely based on velocity as each bird tries to match the velocity (i.e. the speed and direction) of all of its neighbours.

Cohesion is based on the position of all the birds of the flock and is indeed the force which keeps the birds together. Since the mean of the flock depends on all the birds' positions, the mean moves with the flock and as such the birds attempt to follow the mean which results in birds following each other but not in an explicit manner.

The question which remains regarding the Boid model is how boids perceive each other. Since the social rules are the primary drive force, the perceptive abilities of the boids has a large impact on the model. The Boid model does not attempt to simulate the actual perceptive abilities of each boid i.e. it does not simulate vision or hearing, but rather the model attempts to provide the same information that is available to an animal as the end result of its perceptual and cognitive processes. The Boid model restricts perception by limiting the number of neighbours a boid can perceive to only those which fall within a certain range. While this approach does limit a boid's perception it is not ideal, as each boid is still able to perceive more than a real bird. The problem is that any boid which falls within the range is perceived perfectly, which is unrealistic.

Limiting the perception of the birds is not only important for the sake of a realistic simulation, but it was determined through various experiments that without perceptual limitations a natural flocking could not be achieved [26] [14] [11]. The model defined in this paper agrees with those findings. Good natural flocking is difficult to achieve without limiting the perceptual range of the flock members.

The basic Boid model allows for a certain degree of natural flocking to occur, however stagnation can quickly occur. The model stagnates when all the flock members match each others velocity. In this case the flock tends to move in the same direction at the same speed. Some modifications such as giving the flock a random target to follow or restraining the flock within an area can produce more significant flocking, even adding obstacles to the environment as done by Reynolds [26] often achieves good flocking. However true natural flocking such as that demonstrated by starlings occurs in an unbounded environment with few or no obstacles, thus these modifications are counter intuitive to the behaviour of a natural flock. Later in this paper modifications are discussed which allow for more natural flocking to occur, without the need for random targets or obstacles.

2.1.4 Fuzzy Flocking

While most models such as [26] [14] [11] [25] attempt to create an accurate mathematical model, [1] on the other hand suggests an approach based on fuzzy logic. [1] reasons that a mathematical model requires a transition from a linguistic description to mathematical formula which is seldom straightforward, thus the desirability of a fuzzy approach.

In many ways this approach is an extension of the original work by Reynolds. Each flock member is represented by a position and orientation within the world. However this approach extends the flock member model by including an internal state (e.g. fear). The internal state of each flock member affects the drives which control the movement of that flock member. For example if a bird is experiencing "fear" (state) then likely the bird will flee (action). The idea of having a state affecting the action selection does allow for a wider range of behaviours to occur.

Another difference in comparison to the Boid model is that the world defined by [1] is modelled as a collection of animates (i.e. birds or in Reynold's case boids) meaning that the world lacks obstacles for the animates to avoid. With regards to perception, the fuzzy approach models a visual perception by allowing each bird to perceive only what it can see. This visual perception is further limited by adding a maximum range to which a bird can see. This means that a flock member sees fewer of its neighbours and as such is more biologically accurate.

In modelling the drives which control the flock, [1] extends and modifies the original Boid model. It extends on it by including additional forces defined by Hepner[13], these additional forces are:

- **Homing** Which simulates a birds desire to fly towards the roosting site which is a predefined location within the world. This external non social force usually results in some interesting flocking behaviour and prevents stagnation from occurring.
- Velocity Regulation Each bird attempts to maintain its velocity at a desirable level. This is a more biologically accurate and preferable approach when in comparison to the approach of implementing a maximum velocity.
- Interaction Which combines the separation and cohesion drives defined in the Boid model (Section 2.1.3). If two birds are too close they repel each other, if they are too far they do not affect each other, and if they are anywhere in-between they are attracted to each other. This approach is more plausible than the cohesive force due to the fact that each bird is unlikely to know the position of the centre of the flock and as such a global cohesion is unrealistic. It was Pliny [23] [1] [12] who also noted that "It is a peculiarity of the starling kind that they fly in flocks and wheel round in a sort of circular ball, all making towards the centre of the flock".

One of the issues with the attraction rules (repulsion and alignment) is that a bird flying away is a consequence of a neighbouring bird being "too close" just as a bird flying towards the main flock is a consequence of the neighbours being "too far away". The problem noted by [1] is that there are no precise values such as 0 or 5 with regards to "too close" and "too far", indeed their meaning is fuzzy. True we may set the actual values ourselves but the approach used by [1] is to model each value (e.g. "too close") as a fuzzy set which means that a fuzzy value is uniquely defined by its membership function. Due to the interpretation of each fuzzy value being subjective the membership function is not unique; meaning it is left to the modeller to decide.

What must also be noted is that it is possible for a transition of fuzzy values to occur. For example if the current fuzzy value a bird's neighbour is "too close" then as this bird flies further away its fuzzy value will change, the rate of change as well as the extent to which the change occurs is again up to the modeller to decide. All this simply means that it is possible to model the actions of the birds as a set of if-then rules such as the following defined by [1]:

a1: if (distance is close enough) then (flight direction is keep direction)a2: if (distance is too far) then (flight direction is keep direction)

Similarly the repulsive forces can also be modelled:

r1: if (distance is too close) then (flight direction is keep direction) r2: if (distance is too close) and (position is in front or left) then (flight direction is turn right)

The rest of the drives defined by Hepner[13] and Reynolds[26] are represented in a similar fashion.

This approach resulted in less collisions occurring between birds, which in essence makes it a better model in terms of biological accuracy. Unfortunately in order for a thorough study of the dynamical complexity to be conducted, one of the requirements is the ability to sweep the behaviour of the flock by using one parameter to dictate whether the flock is moving around randomly, flocking in a beautiful and natural fashion or completely stagnated. Incorporating a parameter which would allow for such control in this fuzzy model would be difficult and might not even be possible, thus this approach while intuitive is not ideal for this study.

2.1.5 Extending Boids

While the Boid model may achieve some basic flocking it has many limitations. This is why there have been many extensions to the model. One of the most interesting extensions by [14], aims to not only compensate and correct the deficiencies, but also add realistic flight forces to the simulation. The model developed by [14] is aimed at simulating an aerial display of thousands of starlings and is an extension of their earlier work [11] which aimed to simulate the shoaling behaviour in fish.

The "boids" in this model have the same basic principle in that they have a position and an orientation (Shown in Figures 1 and 2 on 6). However these boids are extended in that they also have a mass. This does not only allow the flock to be more heterogeneous but also allows for the incorporation of the speed control formula originally used in [11]. This formula shown in Equation 2.3 is one approach to applying the velocity regulation defined by Hepner[13].

$$F_{\tau_i} = \frac{m}{\tau} (v_0 - v_i) \cdot e_{x_i} \tag{2.3}$$

Where m is the mass of bird i, τ is the relaxation time and v_0 is the velocity it tends towards. The purpose of this speed control component is to encourage birds to tend towards a certain forward velocity, this means that there is no maximum velocity cap; which as a consequence allows birds to catch up to each other.

The drives of the birds in the model proposed by [14] are a modified version of the social drives in the Boid model. One of the modifications made to each of the drives (cohesion, separation and alignment) is the way in which birds perceive their neighbours. [14] incorporates a range which birds must fall within in order to be interpreted as neighbours, but the range is not static and adapts itself based on the number of birds currently in view. As previously stated each group member only perceives a certain number of its neighbours [3]. This adaptive perception aims to simulate this behaviour in a computationally feasible manner. The formula shown in Equation 3.9 on page 20 is used in the model defined in this paper and as such will be discussed in detail later. The perception of neighbours (for certain neighbourhoods) is also further limited by the visual range of each bird. Meaning a bird is only a neighbour of another if that bird can actually be seen and if it falls within the adaptive perception range.

The alignment force remains unchanged in that each bird tries to match the forward velocity of its neighbouring birds by simply computing the difference of forward velocities. The Separative force on the other hand is changed based on the introduction of a safety radius r_h . If the distance between 2 birds drops below this safety radius then both birds experience a strong repulsive force and no cohesive force. If on the other hand the distance between two birds does not fall below the safety radius, but is still below the standard separation radius r_sep then both birds experience a repulsive force following a halved Gaussian, with σ the standard deviation being set such that at the border of the separation zone the force is almost zero (Equation 2.5). The total separative force each bird experiences is then the weighted average of Equation 2.5 taken from all other birds (Shown in Equation 2.4).

$$F_{s_i} = -\frac{w_s}{|N_i|} \sum_{j \in N_i} g(d_{ij}) d_{ij}$$
(2.4)

$$g(x) = \begin{cases} 1 & \text{if } x \le r_h \\ exp\left(-\frac{(x-r_h)^2}{\sigma^2}\right) & \text{if } r_h < x \le r_{sep} \end{cases}$$
(2.5)

Where d_{ij} is the distance between bird *i* and *j*, w_s is the weight of the separative force and N_i is the neighbourhood of bird *i*.

Another modification made is to the cohesive force. As noted by [2] the outer borders of a flock are denser than the interior of the flock, therefore the model by [14] forces birds to be drawn to the center more strongly when the birds are at the border of the flock; which is done by multiplying by the degree to which a bird is away from the centre.

This is achieved by measuring a bird's degree of centrality, which is calculated as the length of the mean vector of direction towards a bird's currently perceived neighbours (Equation 2.7). The centrality C_i is close to 0 when bird *i* is close to the interior or centre of the flock and has values 0.5-0.75 when close to the borders of the flock. Once the centrality has been computed using a larger neighbourhood N_G , the total cohesive force can be computed as shown in Equation 2.6. The purpose of using a larger neighbourhood for cohesive forces is to more strongly encourage birds to fly together rather than fly away from each other; using a larger neighbourhood for cohesion promotes these desires.

$$F_{c_i} = C_i \frac{w_c}{|N_i|} \sum_{j \in N_i} \delta_{ij} d_{ij}; \quad \delta_{ij} = \begin{cases} 0 & d_{ij} \le r_h \\ 1 & d_{ij} > r_h \end{cases}$$
(2.6)

$$C_i = \frac{1}{|N_G|} \left\| \sum_{j \in N_G} d_{ij} \right\|$$

$$(2.7)$$

Another force defined by Hepner[13], the homing force, is incorporated in this model by means of including a roosting area. [14] introduced this behaviour in order to limit the flocking to a specific area without explicitly restricting the flocking area. Each bird experiences an attractive force to the roosting f_{Roost_i} which consists of a horizontal f_{RoostH_i} and vertical f_{RoostV_i} attraction. The horizontal force depends on the degree to which a bird heads outward from the roosting area, i.e. the bird will experience a strong attractive force to the roost if it is heading away from the roost and a weak attractive force if it is heading towards the roost.

The vertical force however f_{RoostV} is proportional to the distance from the preferred height above the roost z. If a bird is too high or to low from z then it will experience a strong vertical attractive force.

All these forces defined by [14] are modelled as follows:

$$f_{Roost_i} = f_{RoostH_i} + f_{RoostV_i} \tag{2.8}$$

$$f_{RoostH_i} = \pm w_{RoostH} \left(\frac{1}{2} + \frac{1}{2} (e_{x_i} \cdot n) \right) \cdot e_{y_i};$$

$$(2.9)$$

$$f_{RoostV_i} = -w_{RoostV}(vertical\ distance) \cdot z; \qquad z = (0, 0, 1)^T$$
(2.10)

Where w_{RoostH} and w_{RoostV} are the weighting factors for the horizontal and vertical attractive roost forces. e_{x_i} and e_{y_i} are the are the x and y coordinates of the bird with regards to the world coordinate system. The final "steering" velocity can then be computed as:

$$F_{Steering_i} = F_{Social_i} + f_{\tau_i} + f_{Roost_i} + f_{\zeta_i} \tag{2.11}$$

Where f_{ζ_i} is a random unit vector sampled from a uniform distribution used as a method of modelling imperfect flight and F_{Social_i} is the weighted sum of the social forces.

In order to truly create an accurate simulation of bird flocking, flight dynamics must be taken into account thus the model by [14] incorporates fixed wing aerodynamics as defined by [36]. During horizontal flight with a constant cruising speed v_0 , the lift L_0 balances the weight mg of the bird and the bird generates thrust T_0 that balances the drag D_0 [14]. The lift drag and thrust created by a bird can be modelled by the following equations [36]:

$$L_0 = \frac{1}{2}\rho S v_0^2 C_L = mg; \quad D_0 = \frac{1}{2}\rho S v_0^2 C_D = T_0$$
(2.12)

$$L_{i} = \frac{v_{i}^{2}}{v_{0}^{2}} L_{0} = \frac{v_{i}^{2}}{v_{0}^{2}} mg; \quad D_{i} = \frac{C_{D}}{C_{L}} L$$
(2.13)

$$F_{Flight_i} = (L_i + D_i + T_0 + mg); \quad L_i = L_i \cdot e_{z_i}; \quad D_i = -D_i \cdot e_{x_i}; \quad T_0 = T_0 \cdot e_{x_i}$$
(2.14)

For the above equations g is a vector reflecting standard gravity, v is the speed, ρ is the air density, S is the wing area, C_L and C_D are the lift and drag coefficients.

Once the social, roosting and aerodynamic forces have been determined the velocity can be computed [14]:

$$v_i(t + \Delta t) = v_i(t) + \frac{1}{m} (F_{Steering_i}(t) + F_{Flight_i}(t)) \Delta t$$
(2.15)

$$r_i(t + \Delta t) = r_i(t) + v_i(t + \Delta t) \cdot \Delta t$$
(2.16)

Where $v_i(t)$ and $r_i(t)$ is the velocity and position of bird i and Δt is the time steps of the model.

This model creates a very accurate simulation of the flocking behaviour of starlings. The accuracy of this model is shown by comparative study conducted by [14]. In this study [14] compares the behaviour of their simulation against that of the behaviour of real starling flocks. Various aspects are analysed such as the turning radius, the volume and the borders of the flock. In most of the studies the simulated flocks performed relatively close to real flocks.

This may seem like the ideal model but for the purpose of this study it is too complex. With such an accurately defined model it might be difficult to achieve full control of the flock using only one parameter. Even if such a parameter could be designed, it too would be complex as it would likely have to effect many components of the model. The study of [14] did however provide a large amount of guidance with regards to not only improving the "realness" of the flock but also with regards to achieving natural flocking behaviour without having to resort to obstacles, random targets, predators or other components which would reduce the realism of the flock.

2.1.6 Other Models

It is worth noting that there exist many different types of models, all promoting and simulating different aspects of this animal aggregation. Covering all these models in this paper would not be feasible so this section of the report is dedicated to a brief summary of other models which were studied.

As previously stated some flocking simulations may stagnate (i.e. all members of the flock behave in the same manner). A common work around for this problem is the introduction of a 2nd swarm or flock, known as the predator swarm. This swarm is entirely goal driven to chase the 1st swarm known as the prey swarm. The resulting model is known as the predator-prey model.

The idea of introducing a 2nd chasing swarm introduces a new force which the members of the prey swarm experience, namely a "fleeing" force. Each member will move away from members of the prey swarm. Usually this "fleeing" force will take precedence over all other rules. The introduction of predators is not only used to create more interesting flocking behaviour but is also used to simulate natural predator-prey behaviour patterns [37]. This approach has also been used in Particle Swarm Optimizations to prevent all particles from converging on the same suboptimal solutions [8].

The predator-prey model was initially considered for this study. Reason being that the predatorprey model provides an inherent control parameter which can be used to sweep the behaviour of the flock, namely the aggressiveness of predators. Predators which are overly aggressive will create a more chaotic flocking behaviour due to chasing prey more rapidly, while predators with very little aggression will not give chase to the prey and hence allow them to stagnate more readily. Discovering a mid-aggressive behaviour for predators would likely produce the desired "natural" flocking. Predator aggressiveness could simply be defined as the maximum velocity a predator is allowed to move. In the end this model was not chosen due to the ultimate goal of the study being the analysis of natural flocking without external influencing factors.

As noted previously by [1] and [11] one approach to producing interesting flocking is the introduction of heterogeneous swarms. Allowing flock members to posses different internal states and properties can result in a more interesting dynamic behaviour, and indeed this also reduces the chances of stagnation occurring.

Taking this idea even further is the introduction of mixed species flocking [22] which is similar to the predator-prey model in that it introduces a variety of swarms, all with different behaviours and drives. These swarms can be in competition with each other for a food source or even just for space to move; regardless of the conditions in which these swarms operate, observing their behaviour while they try to remain with their own species does produce interesting flocking patterns. This is different to the predator-prey model in that the swarms are not necessarily enemies, but rather just attempting to co-habitat in the same world.

A final alternative approach to flocking and swarming worth noting is that of modelling the swarm based on agent architectures [32] [31]. The agent based approach is of interest due to most of the social interactions between flock members occurring as a result of agent communication policies. In this case agents (flock members) communicate with each other explicitly via passing messages to each other. This is an interesting contrast in comparison to the traditional social drives which are simply the result of flock members being able to perceive each other.

While explicit communication between flock members results in some flocking behaviour, such an approach could not be used for this study for the same reason the model proposed by [1] could not be used, namely the lack of being able to introduce a sweeping control parameter. Since each agent operates according to an internal policy and each agent's policy might be different; the only way to control the entire flock with one parameter would be to modify the global communication policy agents' share. Modifying the communication topology might also produce the desired flocking. However achieving such significant modifications with a single parameter could prove challenging.

2.2 Dynamical Complexity

Dynamical Complexity is a measurement of the extent to which a system balances integrated and segregated activity. There are many methods of determining the dynamical complexity of a system, the following sections will not only investigate some of these methods but will also investigate the steps needed before one of these measurement methods can be used; the most important of which is the creation of a time series of data which uniquely measures the different components of a system.

2.2.1 Interaction Complexity

Neural complexity expresses the extent to which a system is both dynamically segregated, so that small subsets of the system tend to behave independently, and dynamically integrated, so that large subsets of the system tend to behave coherently [29]. Interaction Complexity also called Neural Complexity is a technique often used to measure the dynamical complexity of matters regarding the brain and mind. An example of such work by Shanahan [30] measures the dynamical complexity of small world networks of spiking neurons. In this case each component of the system (component in this context referring to a neighbourhood of neurons) is measured uniquely by measuring its mean firing rate (i.e. the rate at which a clusters of neurons spike).

Other works which make use of Neural Complexity are [29] and [34] with the latter being an examination of how such a statistical measures (in this case Neural Complexity) can be used to evaluate the integration of incoming stimuli with ongoing neural interactions.

In order to calculate neural complexity the entropy of a system must be determined. The entropy of a system refers to uncertainty or variability. Essentially if a variable S significantly differs from one time step to the next then this indicates that it carries a maximum amount of information, thus giving S a high entropy. Conversely however if S remains relatively the same, then this indicates that the variable carries a minimal amount of information, thus giving it a low entropy. The entropy of a system is calculated as follows [34]:

$$H(X) = -\sum_{m=1}^{M} p_m log_2(p_m)$$
(2.17)

Where H(X) refers to the entropy of a variable X in a system, M is the number of discrete states a system can take and p_m is the probability of each of these states occurring such that the sum of all probabilities p_m is equal to 1. H(X) is high if the system has many equally likely states (i.e. the system has an equal likelihood of generating each of the states) and H(X) is low if and only if the system attains a single state with p = 1 [34].

The next component needed in order to calculate the neural complexity is the Mutual Information. Mutual Information refers to how much information one variable carries about another. This is denoted as $MI(X_j; S - X_j)$ which determines the mutual information between a variable X_j and the rest system of variables S excluding X_j :

$$MI(X_{i}; S - X_{i}) = H(X_{i}) + H(S - X_{i}) - H(X)$$
(2.18)

The final step before calculating the neural complexity is to calculate the integration I(X), which according to Tononi [34] measures the total entropy loss which is the difference between the sum of the entropy of each component and the entropy of the whole system:

$$I(S) = \sum_{i=1}^{N} H(X_j) - H(S)$$
(2.19)

Where H(S) is the entropy of the whole system and $H(X_j)$ is the entropy of the *j*th component.

Finally the interaction complexity can be calculated as follows [34] [30]:

$$C(S) = \sum_{i=1}^{N} MI(X_i; S - X_i) - I(S)$$
(2.20)

Even though Neural Complexity is suited for measuring aspects about the brain and consciousness it could easily be applied to a flocking simulation. This is due to the only thing being required is a collection time series of data representing the different components(members) of the system(flock).

2.2.2 Information Integration Φ

According to [33] and [29] Information Integration is a method of measuring the total amount of information a system can integrate. [33] defined Φ as the "effective information" across the informational "weakest link" which is called the "minimum information bipartition" [29]. The question is, what is effective information, how can it be computed with regards to the system being measured, and in what way can it be used to locate the minimum information bipartition.

Effective information is computed as the mutual information (Equation 2.18) across a partition (a partition referring to a portion of components of the system) except that all the measurements from one of the subsets or components of the system is made to have maximum entropy. The minimum information bipartition is the component of the system with the lowest effective information.

As computing the effective information depends on one of the components of the system having maximum entropy, it is important be able to replace the outputs all possible subsets of a system with uncorrelated white noise [29]. As can be imagined in a system where each component is capable of reaching a large number of different states it can be difficult or computationally expensive to force a subset of the system to demonstrate maximum entropy. It is important to note that according to [29] effective information across the majority of the components of the system is useful only if it helps determine which partition is the minimum information bipartition.

According to [29], Φ reflects causal interactions because it is based on effective information, which is a directed version of mutual information. [29] also states that Φ cannot be used to measure non-trivial systems because:

- As stated above replacing the outputs of a system with uncorrelated noise such that a component exhibits maximum entropy is infeasible simply because most real world systems are capable of reaching a large amount of unique states, some of which may even be unknown. The states a system might be able to generate could even be continuous.
- Computing Φ requires the calculation of effective information across every bipartition of a system. Meaning that there is a factorial growth in the number of partitions which must be analysed as the grows system larger. This makes the computation of Φ computationally infeasible for large systems. Considering some flocks or swarms can have thousands or even tens of thousands of members, using Φ to determine the dynamical complexity could prove difficult.

In order to understand effective information, [29] gives an example of a Hopfield-type network with (k, n-k) bipartition of the network, A|B with A having k components of the system and B having n-k, i.e. the rest of the components of the system. The effective information can be computed by [29]:

$$EI(A \to B) = H(A) + H(B) - H(AB)$$

$$(2.21)$$

Where the outputs or measurements (in the context of flocking) of A are replaced by uncorrelated white noise. According to [29]: $EI(A \rightarrow B) = k$ bits. Meaning that $EI(B \rightarrow A) = n$ -k bits which in turn, means:

$$EI(A \leftrightarrow B) = EI(A \to B) + EI(B \to A) = n \text{ bits}$$

$$(2.22)$$

Since Equation 2.22 does not depend on k, the effective information is the same for all bi-partitions, therefore the information integration is given by:

$$\Phi = n \ bits \tag{2.23}$$

Information Integration, while being a seemingly accurate measurement is not the best choice for this study. This is due to flock behaviour not being finite, i.e. a flock usually does not occur in a finite number of unique states meaning that the data retrieved from the flocking simulation is continuous. As such calculating the effective information will be difficult or inaccurate. This together with the computational complexity of Φ rising with the size of the system (flock) makes it non-ideal measurement.

2.2.3 Causal Density

The methods studied in Sections 2.2.1 and 2.2.2 have been mostly applied to the area of brains and measuring brain related aspects such as consciousness. An alternative method proposed by Granger [10] not only been applied to brain analysis but also to various other topics with reasonable success [27][9]. This measure called *Granger-causality* [10] is based on measuring the causal influences between the components of a system. In essence it measures the degree to which one variable influences or causes another. With regards to a swarm this simply refers to how much one member of a swarm will attract or repel another member.

[30] provides an example in order to aid in understanding Granger Causality:

By looking at 3 time series $X_1(t)$, $X_2(t)$ and $X_3(t)$, where each time series refers to the outputs generated by a specific component or variable of the system. $X_1(t)$ can be described by the following autoregressive model:

$$X_1(t) = \sum_{j=1}^m A_j X_1(t-j) + B_j X_2(t-j) + C_j X_3(t-j) + \varepsilon_{ABC}(t)$$
(2.24)

Where *m* is maximum number of lagged observations in the model (m < TheSizeOfEachTimeSeries). A, B and C are vectors which contain the coefficients of the model which are indexed by the model order and ε_{ABC} is the prediction error for each time series. If we then consider the following model for $X_1(t)$ where $X_2(t)$ is omitted:

$$X_1(t) = \sum_{j=1}^m A_j X_1(t-j) + C_j X_3(t-j) + \varepsilon_{AC}(t)$$
(2.25)

Where ε_{AC} is the prediction error of the time series with $X_2(t)$ omitted. It can then said that X_2 Granger-causes X_1 if the variance of ε_{ABC} is significantly less than the variance of ε_{AC} . In other words X_2 Granger-causes X_1 if including X_2 in the model aids in predicting X_1 . It is also important to note that X_2 Granger-causes X_1 only if X_2 influences X_1 greater than the other components of the system (X_3) . To summarise, Granger Casualty measures the degree to which X_2 causes X_1 over and above the rest of the system.

With this it is possible to determine which components of the system have the most pull over the system. Which in terms of a swarm means that it makes possible the detection of which swarm members are the primary "drivers" of the swarm.

From this, the *causal density* of a system S can be computed as follows [28]:

$$cd(X) = \frac{1}{n(n-1)} \sum_{i \neq j} F_{x_i} \to x_j | x_{ij}$$
 (2.26)

Where x_{ij} denotes a subsystem of X with the components X_i and X_j omitted. Unit causal density, which measures the proportion of all causal relations which are statistically significant can be measured as follows [30]:

$$cd(X) = \frac{\alpha}{n(n-1)} \tag{2.27}$$

Where α is the number of pairs of variables X_i and X_j , such that X_i Granger Causes X_j .

It is important to note that the time series $X_1(t)$, $X_2(t)$ and $X_3(t)$ must be covariance stationary. Meaning that neither the mean or the auto-covariances depend on the time or spatial index. Formally covariance stationary can be defined [7]: A stochastic process y_t is covariance stationary if there exists a mean m, a constant variance s^2 and a series of auto-covariances g_s such that:

1. $E(y_t) = m$ for all integers t

- 2. $E((y_t m)^2) = s^2$ for all integers t and
- 3. $E((y_t m)(y_{t+j} m)) = E((y_s m)(y_{s+j} m))$ for all integers s,t and j.

Where E is the mean or expectations operator.

Achieving a covariance stationary time series is important as it allows an F-Test to be performed in order to determine the significance of the Granger causality between two variables.

Using causal density would allow the evaluation of the degree to which different flock members drive the whole flock in different ways. Essentially if a flock moves about randomly (meaning that each member of the flock does not influence any other members) the result would be a low causal density. On the other hand if a flock moves as one unit, then that would mean that all birds influence each other equally which in turn results again in a low causal density. This is simply a theory of what the results of a causal density analysis could be. The actual results of such an analysis are discussed later.

2.2.4 Quantifying Flocking and Swarming

A recurring issue appears in Sections 2.2.1, 2.2.2 and 2.2.3 in that they all require the system to produce "outputs", i.e. the system must produce various time series $X_1(t), X_2(t), X_3(t)$. $X_N(t)$ etc which correspond to the N components or variables of the system. This introduces one of the major issues of the study, how can flocking behaviour be quantified in such a manner to produce these time series?

This depends on the type of system being analysed. For example in the analysis of spiking neural networks the mean firing rate can be used [30]. However in analysing a flock this introduces a problem due to there being a lack of a truly unique way to quantify the behaviour of flock members (system variables). The study conducted compensated for this problem by using a variety of imperfect measurements (Discussed later in Section 3.2.1) and then analysing how good each measurement method was in quantifying the flock. The details of each measurement method and its analysis is discussed in Section 3.2 while the issues and fall backs of each measurement method is discussed in Section 6.

In attempting to develop a suitable quantification method or "time series generating" method, a study conducted by Seth [27] which aimed to measure *G*-emergence outlined one method of measuring a swarm such that a Granger Causality analysis could then be performed. In [27], Seth calculates the heading and speed of a bird as follows:

$$\alpha_i = \alpha_i + a_1 \theta_1 + a_2 (\pi + \theta_2) + a_3 \theta_3 + r_1 \tag{2.28}$$

$$s_i = s_i + a_4 \overline{ds} + r_2 \tag{2.29}$$

Where α_i is the heading, s_i is the speed, θ_1 is the bearing to the center of mass of the swarm, θ_2 is the bearing to the closest bird, θ_3 is the bearing to the mean heading of all the other birds, \overline{ds} is the difference of speed between bird *i* and all other birds and r_1, r_2 are random numbers in the range [-0.01, 0.01]. The vector *a* is used to scale the relative contribution of each factor.

The above equations are not the measurements used by Seth but they do highlight a possible measurement, namely \overline{ds} . The difference of speed not only takes into account the speed of the bird itself but also relative to all other birds, thus identifying it apart from the other birds.

This may not be a perfect way of uniquely measuring a bird relative to the flock but it was used as one of the measurements methods conducted in the study described in this paper. Another study by Cavagna [5] also highlights the difference of speed between bird i and all the other birds. In [5] birds are measured by studying their behavioural fluctuations around the mean flock's velocity and calculating said fluctuation as follows:

$$u_i = v_i - \frac{1}{N} \sum_{k=1}^N v_k \tag{2.30}$$

The above equation is essentially used to calculate the difference of velocities (difference between bird i's velocity and the mean flock velocity) from which the difference of speeds can easily follow.

Looking back at [27], the model used by Seth is a simple 2 dimensional Boid Model and the actual method used to perform the measurements is to take the distance between a bird's current position $[x_i, y_i]$ and the center of the environment. This approach can easily be extended to the third dimension and thus a more realistic flocking model.

Even the measurement methods outlined by [5] and [27] are not perfect but they both were sufficient in generating the necessary data for the purposes of those studies. A final important note is to remember, as previously stated in Section 2.2.3 the generated time series must be covariance stationary. For some of the measurement methods defined in Section 3.2 this is already the case, but for others some preprocessing is required, such as calculating the differences of the original time series.

Regardless of the measurement approach taken, it is important to validate that these measurements return the data in a correct and desirable format needed for measuring the dynamical complexity of the flock.

3 Methodology

As mentioned the goal of this study is to measure the dynamical complexity of the flocking or swarming behaviour demonstrated by animals. Remember that one of the requirements for the study to be conducted is the definition of a singular control parameter which can be swept between a certain range such that at one extreme, e.g the minimum; the flocking pattern is random (As shown in Appendix B.1), the other extreme, e.g the maximum; the flocking pattern has stagnated (As shown in Appendix B.3) and some setting in between the minimum and maximum results in the desired "natural" flocking (As Shown in Appendix B.2). None of the existing models studied provided such complex control with only one parameter thus a new model had to be developed. The details of this model are discussed in the following section.

The analysis itself has to be conducted in stages, this is not only due to the weaknesses of the model but also due to the open question of how best to measure the flocking behaviour. Since the validity of the measurement methods is not confirmed and the results of the dynamical complexity analysis were unknown, everything had to be questioned and tested. The procedure by which the various measurements methods are possibly validated or invalidated (depending on the method) is discussed later and became one of the bigger issues of the study.

3.1 The Flocking Model

After careful consideration and study of the various flocking patterns exhibited by different animal groups, a specific species, namely starling flocks showed the most promise. This is not only due to the fact that various studies have been performed regarding starling flocks [14] [5] [2] but also because the flocking behaviour demonstrated by starlings is some of the most beautiful flocking patterns observed.

The model defined in this paper is an attempt at simulating those flocking patterns in a biologically real manner. The primary focus of the model proposed is the proper definition and control of the social attributes. The reason for the focus on the social attributes is because it is through these social attributes that a "sweeping" parameter can be found which controls the behaviour of the flock.

The flocking model is primarily based on the original Boid Model[26] with some influences from [14] and [13]. Each bird in the model has: **1**. A position within the 3 dimensional simulated world represented by a co-ordinate vector $p_i(t) = (x, y, z)$. **2**. A rotation vector $r_i(t) = (X, Y, Z)$ which represents its current rotation with respect to the bird's local axis. Note that while the rotation of the bird with respect to its local axis varies, the axis itself remains fixed and moves with the bird. Similar examples of (**1**) and (**2**) are given during the discussion of the Boid Model and are shown in Figures 1 and 2 on page 6. **3**. A velocity $v_i(t) = (\overline{vx}, \overline{vy}, \overline{vz})$ which represents the bird's speed in each axis and is the means by which the bird moves around the world. Each bird also has a mass $m_i = \pm 80$ which is not only to make the flock more heterogeneous but also to implement the speed control proposed by [14].

This model does not include banking, thus the roll of a bird is omitted. It is acceptable to omit the roll in this model because the primary reason for maintaining the rotation of the bird to begin with (other than being biologically real) is to allow for the modelling of the bird's visual perspective, and roll does not affect what a bird can see, but only the rotation of what a bird can see. Which for the purpose of this study is not important. It is indeed the pitch and the yaw which is important, due to having a direct impact on the direction of what a bird is looking at and thus affecting what the bird can see.

In order to model vision of a bird, the angle of rotation of a bird must be updated as a bird flies around the simulated world. At the end of each time step t the rotation of a bird i is updated

as follows:

1. The new pitch $\gamma_i(t)$ and yaw $\delta_i(t)$ of bird *i* is calculated based on the current velocity of the bird:

$$||v_i(t)|| = \sqrt{\overline{vx}_i^2(t) + \overline{vy}_i^2(t) + \overline{vz}_i^2(t)} \qquad Velocity \ Magnitude \tag{3.1}$$

$$\gamma_i(t) = atan2(-\overline{v}\overline{z}_i(t), \overline{v}\overline{x}_i(t)); \quad \delta_i(t) = \arcsin\overline{v}\overline{y}_i(t)/||v_i(t)||$$
(3.2)

where \overline{vx} , \overline{vy} and \overline{vz} are the individual components of the velocity vector v_i .

2. The following rotational matrices are created which are used to model pitch and yaw rotation transformations:

$$R_Y(\gamma_i(t)) = \begin{pmatrix} \cos \gamma_i(t) & 0 & \sin \gamma_i(t) \\ 0 & 1 & 0 \\ -\sin \gamma_i(t) & 0 & \cos \gamma_i(t) \end{pmatrix}$$
(3.3)

$$R_Z(\delta_i(t)) = \begin{pmatrix} \cos \delta(t)_i & -\sin \delta_i(t) & 0\\ \sin \delta_i(t) & \cos \delta_i(t) & 0\\ 0 & 0 & 1 \end{pmatrix}$$
(3.4)

3. Using the new three dimensional rotation matrices above the new rotation of bird i with respect to its local axis can be updated by:

$$r_i(t) = R_Y(\gamma_i(t))R_Z(\delta_i(t))r_i(t-1)$$
(3.5)

By maintaining the rotation of bird i as it moves throughout the world it is then possible to check if one bird is in view of another and as such model basic visual perception for birds. The procedure for checking if a bird j is in view of another bird i is as follows:

$$fov = \cos\theta \tag{3.6}$$

$$\lambda_{ij}(t) = ||r_i(t)|| \cdot ||p_j(t) - p_i(t)||$$
(3.7)

$$\psi_{ij}(t) = \begin{cases} 1 & \text{if } \lambda_{ij}(t) \ge fov \\ 0 & \text{if } \lambda_{ij}(t) < fov \end{cases}$$
(3.8)

Where θ indicates in radians the amount a bird is able to see from it's central viewpoint. For example $\theta = radians(180)$ would mean the bird is capable of seeing all around it. A bird j is in view of another i bird if $\lambda_{ij}(t) \ge fov$ otherwise it is not in view.



Figure 3: The viewing radius of a bird. The blue portion indicates the central bird's field of view. The red line indicates it's visual range.

Note that the vision in Equation 3.8 has no maximum range, meaning that a bird i can see another bird j fully and perfectly as long as j falls within the field of view of i. The maximum visual range (i.e. how far a bird can see) is taken into account before checking if one bird j is in view of another bird i. The reason this is not incorporated as part of Equation 3.8 is because a different maximum range is defined for each social rule, this is explained in detail next.

3.1.1 The Social Forces

1

Before defining the actual social forces each bird experiences, it is important to define the local neighbourhoods of each bird. As previously stated each bird only perceives its 6 or 7 closest neighbours [1] [14], thus defining neighbourhoods in terms of a metric distance would result in birds seeing too few or too many neighbours, so using the topological distance will result in a better model. However with regards to the simulation model, determining a bird's closest neighbours at each time step is computationally expensive therefore a compromise is made.

The compromise made is to use an adaptive perception range as proposed by Hildenbrandt [14]. This means that each bird's perceptual range will increase if it sees too few birds and it will decrease if it sees too many birds. The goal being that each bird attempts to keep in view a certain target number of birds. Each bird will adapt its perceptual range as follows:

$$Per_i(t) = (1-s)Per_i(t-1) + s\left(Per_{max} - Per_{max}\frac{|N_{gen_i}(t)|}{n_c}\right)$$
(3.9)

Where $0 \le s \le 1$ is the interpolation factor which controls the rate at which the perceptual range $Per_i(t)$ of bird *i* contracts and expands. N_{gen_i} is the currently perceived neighbourhood of bird *i* and n_c is the number of birds which should be perceived.

In a very large flock with 14,000 - 15,000 birds perceiving 6-7 neighbours is not a problem, however modelling such a large number of birds is computationally expensive thus the number of birds which were modelled were around 1000-3000. As a consequence n_c could not be explicitly set to 6-7 due to their being too few birds, thus n_c was actually set to a percentage of the number of birds in the simulated flock, specifically 10%.

Taking more inspiration from [14], this model also defines multiple neighbourhoods. This not only allows more control of the flocking patterns but was also a valuable element in creating the sweeping parameter which will be referred to as the flocking factor ξ . The Neighbourhoods are defined as follows where N_i is the perceived flock of bird *i* (i.e. all birds of the flock except *i*):

• The general neighbourhood which is used for modifying the perceptive range of bird i. A bird j is in the neighbourhood of bird i if the distance d_{ij} between those birds is less than or equivalent to the current perceptive range of the bird $Per_i(t)$.

$$N_{gen_i}(t) = \{ j \in N_i : 0 \le d_{ij}(t) \le Per_i(t) \} \qquad General \ Neighbourhood \tag{3.10}$$

• The separative neighbourhood which refers to which birds impose a repulsive force on a bird *i*. In this case a bird is only considered a neighbour if the distance to that bird is within a safety radius R_s . The safety radius simply refers to the private flight area of a bird, and should any other bird enter that area then there is a risk of a collision occurring which is why repulsive forces are experienced at such a point. Vision ψ_{ij} at this stage is not incorporated, this is because birds should be able to avoid each other based on their other senses, not only sight:

$$N_{sep_i}(t) = \{ j \in N_i : 0 \le d_{ij}(t) \le R_s \} \qquad Separative \ Neighbourhood \tag{3.11}$$

• The cohesive neighbourhood which is a local group of birds which a bird *i* will try to move towards. In the following case the perceptive range is doubled; this is to allow for birds to move closer together into larger sub-flocks. Furthermore this neighbourhood incorporates

the visual senses ψ_{ij} of the bird. Incorporating vision is because birds should only move towards a position which is known, and a position is only known if it can be seen:

 $N_{coh_i}(t) = \{ j \in N_i : 0 \le d_{ij}(t) \le 2Per_i(t) \text{ and } \psi_{ij}(t) = 1 \}$ Local Cohesive Neighbourhood (3.12)

• The alignment neighbourhood refers to which birds, a bird *i* will try to match in terms of velocity. In this case vision is incorporated because birds should only be able to match velocities if they know the velocity of other birds, and a bird will only know the velocity of another bird if it can see that bird. The perception is altered by an adjustment value κ . The purpose of this is so that the flocking factor (sweeping parameter) ξ can have the desired affect on the behaviour of the flock. By increasing the size of the alignment neighbourhood according to ξ a stagnated flock is more likely to occur, just as decreasing ξ and indeed the neighbourhood will likely result in a more random flocking pattern.

$$N_{ali_i}(t) = \{ j \in N_i : 0 \le d_{ij}(t) \le \kappa Per_i(t) \text{ and } \psi_{ij}(t) = 1 \}; \qquad \kappa = \begin{cases} \xi & \text{if } \xi \ge 1\\ 1 & \text{otherwise} \end{cases}$$
(3.13)

An example of how the different neighbourhoods are interpreted by a bird is shown in Figure 4.



Figure 4: The Neighbourhoods of a Bird: The blue shaded area represents the central bird's perception. The alignment neighbourhood $N_{ali_i}(t)$ is not included due to it taking various sizes depending on the value of $(1 \le \kappa \le 10)$, however the alignment neighbourhood is at minimum as small as the general neighbourhood and at maximum as large as 10 times the bird's current perception range, which often encompass the whole flock.

The social forces of the model are made up of 4 components which essentially are modified versions of the original social rules proposed by Reynolds [26]:

$$F_{social_i}(t) = af_{\alpha_i}(t) + bf_{\beta_i}(t) + cf_{\eta_i}(t) + \frac{\xi^2}{5}f_{\tau_i}(t)$$
(3.14)

Where f_{α_i} is the separative force, f_{β_i} is the local cohesive force, f_{η} is the alignment or matching force and f_{τ_i} is the global cohesive force. The contribution of each of the primary forces are weighted by a, b and c. The global cohesive force is weighted $\frac{\xi^2}{5}$ which was an empirically determined weighting measure which allowed the flocking factor ξ to exhibit the desired control. The

reason being that for high values of ξ the flocking should stagnate and in order for this to occur, the flock should move as one, which is why the flocking factor must scale the contribution of the global cohesive force.

As previously mentioned, the separative force f_{α_i} , is used to model the desire for birds to avoid collisions. Limiting the repulsive force means that bird *i* will only experience the following repulsion:

$$f_{\alpha_i}(t) = \sum_{j \in N_{sep_i}(t)} \frac{||p_i(t) - p_j(t)||}{d_{ij}(t)}$$
(3.15)

The cohesive force is used to model a bird's desire to stay close to the flock. In this model a bird i will experience 2 cohesive forces. The first being a local cohesive force which is restricted to the birds in the cohesive neighbourhood (Defined in Equation 3.12). Which is simply the difference between a bird i's current position and the mean position of the neighbourhood N_{coh_i} :

$$f_{\beta_i}(t) = p_i(t) - \frac{1}{|N_{coh_i}(t)|} \sum_{j \in N_{coh_i}(t)} p_j(t)$$
(3.16)

The second being a global cohesive force which is applicable to the entire flock. It is agreed by many authors [26] [25] [14] [5] that global perception may not be realistic but achieving a stagnated flocking pattern is difficult without using a large perceptive range.

This global cohesive force is only in the model as to allow for the introduction of the flocking factor. True natural flocking occurs without birds' knowing the state of all the other birds thus low values of a global cohesive force should result in a better flocking and indeed this is the case for this model (As shown in Appendix B.2). A low flocking factor makes the global cohesive force negligible. The global cohesive force is computed as the difference between a bird *i*'s position and the mean position of the perceived flock N_i ;

$$f_{\tau_i}(t) = p_i(t) - \frac{1}{|N_i|} \sum_{j \in N_i(t)} p_j(t)$$
(3.17)

The final and arguably the most important social force is that of a bird's desire to match the velocity of its neighbouring birds. This can simply be computed as the difference between a bird *i*'s velocity and the mean velocity of the alignment neighbourhood:

$$f_{\eta_i}(t) = v_i(t) - \frac{1}{|N_{ali_i}(t)|} \sum_{j \in N_{ali_i}(t)} v_j(t)$$
(3.18)

The social forces defined in this section are enough to achieve a basic flocking, however these rules are not enough to create natural flocking. This is because once the members of each sub-flock (or even the entire flock) match their velocities, there would be no motivation or reason for them to change direction. Meaning that each sub flock would fly continuously in the same direction at the same speed. This is not only boring to watch but is not even even remotely close to natural flocking. The next section covers how this problem can be solved in a simple manner.

3.1.2 Other Model Forces

A few methods have been discussed in Section 2.1 with respect on how to achieve a more natural flocking pattern. These methods have ranged from giving the flock a random target to follow, to constricting the flock within a certain region and even introducing obstacles for the flock to avoid.

For this model the method chosen was the introduction of a roosting site, thus adding another force to the flock representing flock members' desire to roost. It is also important that not all birds should fly home all the time, otherwise a different kind of stagnation occurs where the flock would simply circle around its roosting site.

Rather each flock member must make the decision to fly home based on its locally perceived environment. Each flock member experiences a constant weakened force to fly home, this could simply signify the desire to not stray too far away from a known safe location. In addition to this constant force, a flock member will also make the decision to fly home if it is currently alone or perceives too few other birds in its local neighbourhood. The idea behind this is that if a flock member is secluded then it is essentially under greater risk of being attacked by a predator, which is why it will likely decide to fly back to its home location to look for other birds.

On the other hand if a flock member perceives too many neighbours and feels "crowded" that bird will fly away from the home location in order to be in area with less risks of a collision occurring. This means that each flock member has the internal goal of roosting at the most ideal position but at the same time not feeling too crowded. This behaviour is modelled as follows:

$$f_{\sigma_i}(t) = \begin{cases} d\frac{\xi}{10}h_i(t)^2 & \text{if } t \mod \overline{TL} = 0 \text{ and } \frac{N_{ali_i}(t)}{n_c} < P_r \\ \frac{\xi}{10}h_i(t) & \text{otherwise} \end{cases}$$
(3.19)

$$h_i(t) = p_i(t) - H (3.20)$$

Where f_{σ_i} is the final force to go home, h_i is the calculated force to fly home, H = (0, 0, 0) is the home location and d is a scaling factor for when the strengthened force is used. Thus a bird *i* will experience a stronger drive to go home if:

- 1. A sufficient amount of time has passed since a bird has made the decision to fly home, (t mod $\overline{TL} = 0$) where \overline{TL} is the amount time that passes before the bird makes the decision once more. Larger values means a bird will stay on its current course for a longer period of time.
- 2. The number of birds it currently perceives is less than a predefined value $\left(\frac{N_{ali_i}(t)}{n_c}\right) < P_r$ where $P_r = 0.1$ is the likelihood of a bird choosing to flying home.

The reason for incorporating the flocking factor ξ , with the roosting force is to allow birds to change direction once they have unified and stagnated. Without an increased roosting force, a stagnated flock will not change direction, thus the need to incorporate the flocking factor with the roosting force.

The final force added to the model is a random force f_{ω_i} drawn from a standard Gaussian. This is used to not only add a stochastic element to the model but also to model imperfect flight. This makes the final non-social forces of the model:

$$F_{other_i}(t) = f_{\sigma_i}(t) + e f_{\omega_i} \tag{3.21}$$

Where e is a scaling factor used to weigh the contribution of the random force.

3.1.3 The Final Model and The Model Parameters

Using the forces defined in the previous sections, the velocity of each bird is updated as follows:

$$v_i^*(t) = F_{social_i}(t) + F_{other_i}(t) + v_t(t-1)$$
(3.22)

In order to make the flight of the birds more realistic a velocity cap is introduced to the model $v_{max} = 10$. If a bird's velocity v_i^* is ever greater than the maximum velocity v_{max} then that bird's velocity is slowly tapered down back to the maximum allowable velocity using the speed control proposed by [14] (As shown in Equation 3.23 where Δ is the relaxation time). This speed control is important as it allows birds to slightly go over the maximum velocity, which means that slow birds

will have an opportunity to catch up to the main flock, thus reducing the likely hood of stragglers occurring during the simulation.

$$v_{i}(t) = \begin{cases} v_{i}^{*}(t)((v_{max} - v_{i}^{*}(t))(m_{i}/\Delta)) & \text{if } v_{i}^{*}(t) > v_{max} \\ v_{i}^{*}(t) & \text{otherwise} \end{cases}$$
(3.23)

Once the velocity is regulated to realistic speeds the position of the bird can be updated:

$$p_i t = p_i (t - 1) + v_i(t) \tag{3.24}$$

The Field of View, θ	90 radians
Minimum Perception, Permin	0
Maximum Perception, Permax	1000
Percentage of Birds kept in view, n_c	10
Interpolation Factor, s	0.1
Safety Radius, R_s	5-
Separative Weight, a	2
Cohesive Weight, b	2
Alignment Weight, c	1
Roosting Weight , d	5
Alignment Weight, e	1
Time Lapse, \overline{TL}	1
Return Home Probability, P_r	0.1
Maximum Velocity, v_{max}	10

Table 1: Settings used for the experiments conducted. Note: the flocking factor ξ is not included above as it changes for each simulation run.

As can be easily observed the model incorporates a large number of different factors and parameters, each of which with many possible values. In order to simplify the experimentation an application was developed to interact with the model in an easy and real-time fashion, the user guide for this application is shown in Appendix A. Furthermore that in Appendix C a code snippet from the simulation environment shows the calculation of the social and roosting forces defined in the previous section.

For the dynamical complexity analysis, Table 1 shows the parameter values which were used. These values were chosen because they resulted in the desired range of behaviours depending solely on the value of the flocking factor ξ . An example of the 3 extreme flocking patterns exhibited is given in Appendix B.

3.2 The Analysis Procedure

The method chosen for analysing the simulated flocking behaviour is that of Causal Density using *Granger Causality* [10] (Discussed in Section 2.2.3). Before being able to determine the causal density, the simulation environment must produce a collection of time series which quantifies the different components of the system.

In the context of the flocking behaviour the "components of the system" refers to a small number n of special indicator birds which are selected at random. The purpose of having these indicator birds is that each bird i will produce a time series X_i which quantifies its behaviour. Ideally since every bird has the potential to act independently, every bird should produce a time series. However in a simulation which can include up to 3000 birds, the amount of data being produced would be too large, thus the existence of "indicator" birds.

3.2.1 From Flocking To Data

A problem briefly discussed in Section 2.2.4 must now be addressed. Each "indicator" bird must produce a time series X_i , the question is how can such a time series be produced ?

This turns out to be one of the more key issues of the study because the time series X_i needs to be an accurate indication of the behaviour of that bird and each element of the time series must be a singular number. This is very difficult in a 3 dimensional world where a bird is usually identified via its position $p_i(t) = (x, y, z)$. The work around for this problem is to introduce several different measures in the hope that at least one (or more) uniquely quantifies the behaviour of a bird. It should be noted that the reason only the indicator bird itself is measured, is because taking the neighbourhood of the indicator bird into account would often result in overlapping neighbourhoods between indicator birds which in turn produced strange and inaccurate results.

The reader should note that each of these measures are not perfect, meaning that there can be some ambiguity when comparing the time series produced by one bird and the time series produced by another bird. The fall-backs of each of these measures will be addressed later in Section 6. Each indicator bird is measured at each time step of the simulation, thus no data or behaviour is omitted. The measures used are:

- 1. The Average Velocity Each bird is quantified according to its current average velocity. So at each time step t, indicator bird i will be measured as $X_i(t) = \overline{v_i(t)}$ (Where $\overline{v_i(t)}$ is the average calculated by summing the velocity components and dividing by three). A bird's velocity is always changing mostly according to the behaviour of its flock mates (as a consequence of the social rules), thus this measure could possibly identify how a bird moves with regards to the forces it receives from its flock mates.
- 2. The Average Position Each bird is quantified according to its current average position. Meaning that at each time step t, indicator bird i will be measured as $X_i(t) = \overline{p_i(t)}$. Since a bird is likely to follow its locally perceived sub-flock, its position should be close to the mean position of that sub-flock. Such a measure might produce the desired time series.
- 3. The Position Difference In this case each bird is measured according to the distance it has moved from one time step to the next. Since the measurements are made every millisecond this essentially measures the bird's speed at each point of the simulation $X_i(t) = \sqrt{(x_i(t) - x_i(t-1))^2 + (y_i(t) - y_i(t-1))^2 + (z_i(t) - z_i(t-1))^2}$. Where x, y and z are the components of the position p_i . This simply tracks by how much a bird changes its position as it moves through the world. This has the advantage that the time series produced is usually co-variance stationary as required.
- 4. The Average Velocity Difference This simply measures by how much a bird's velocity changes from one time step to the next. This is calculated as the absolute value of the average of the difference between the velocity at one time step t and the velocity of the previous time step (t-1): $X_i(t) = \overline{|v_i(t) v_i(t-1)|}$.
- 5. The Difference from Home A bird is measured according to how far it currently is from the roosting site H = (0, 0, 0). This is an interesting approach as it is similar to that used by Seth [27] and unlike the previous measures it makes use of a point of reference $X_i(t) = \sqrt{(x_i(t))^2 + (y_i(t))^2 + (z_i(t))^2}$. This addresses several issues with regards to having to use the average of the position or velocity
- 6. The Position Difference from The Flock Which will also be referred to as the Difference From Mean. Each bird is measured according to how far it is from the mean position of the perceived flock:

 $X_i(t) = \sqrt{(x_i(t) - P_{ix}(t))^2 + (y_i(t) - P_{iy}(t))^2 + (z_i(t) - P_{iz}(t))^2}$ where the mean position of the perceived flock for bird *i* is $P_i(t) = (P_{ix}(t), P_{iy}(t), P_{iz}(t))$ and is calculated by Equation 3.25. This takes some inspiration from [27] except that is uses a moving point of reference

which is dependent on the behaviour of **all** the flock members. Theoretically this means that all changes within the flock should be explicitly measured unlike the previous measures where the behaviour of the whole flock is implicitly measured as a consequence of the social rules. In essence this should reflect the behaviour of the bird relative to the whole flock more accurately.

$$P_i(t) = \frac{1}{|N_i|} \sum_{j \in N_i} p_j(t)$$
(3.25)

7. The Speed Difference from The Flock - Briefly mentioned by [27] and [5] this also quantifies the behaviour of a bird relative to a moving reference point. A bird is measured according to the difference of speed between itself and the average speed of the perceived flock $(X_i(t) = |v_i(t)| - S_i(t))$ where $v_i(t)|$ is the magnitude of the velocity of a bird *i*, which in this case reflects the speed since the measurements are being made per millisecond and $S_i(t)$ is the average speed of the whole flock calculated by Equation 3.26. Theoretically this should be more accurate than measures (1)-(5) however it might be less accurate than (6) because the average speed of a flock is likely to not change as frequently in comparison to the average position of a flock.

$$S_i(t) = \frac{1}{|N_i|} \sum_{j \in N_i} |v_j(t)|$$
(3.26)

The result of using these measures is a collection of time series $X_1, X_2, X_3...X_n$ which correlate to the behaviour of the *n* indicator birds, which should allow for an analysis of the flocking behaviour demonstrated by the whole flock. During the trial stages of experimentation it was discovered that the number of birds significantly impacted the results. In other words the graph produced using 10 birds would not have the same curve as a graph produced using 1000 birds. This could simply be due to the fact that a larger number of birds results in greater social forces which in turn directly affects the causal density demonstrated. As such, two primary experiments were conducted; one with 10 birds, 3 of which were used as indicators and one with 1000 birds, 10 of which were used as indicators.

3.2.2 From Data To Causal Density

Once the necessary collection of time series have been generated the causal density is analysed within small interleaving windows. Essentially meaning that rather than using the whole time series to measure the causal density, the causal density is measured in a small window which moves across the whole time series. The window itself spans 500 ms of the simulation and moves by an offset of 250 ms. As a result the analysis returns two components:

- 1. The maximum causal density measured in any window.
- 2. The average causal density across all the windows.

The size of the window was determined via a series of trials. With a large window size the results tend to span out more and be less consistent (Which can be seen in Figure 5c). Smaller sized windows did make the results more consistent, however if the window was set to be too small then that would result in very low causal density (As shown in Figure 5a). Also very small windows often made it difficult to make a time series covariance stationary due to lack of data being available from which differencing can be applied. Thus an average is chosen (Which is shown in Figure 5b).

Measuring the causal density in small windows is preferable firstly because it is more likely that data is covariance stationary within each window and secondly the results tend to be more accurate. If a time series within a window is not covariance stationary; which can be easily determined using the Augmented Dickey Fuller test; it might be possible to make it so by 1. removing linear trends and 2.computing the difference of the time series $(Y_i(t) = X_i(t) - X_i(t-1))$. If necessary,



Figure 5: Effects of different sized analysis windows. The above dot plots are just experimental trials which were used to find the ideal sized window.

differencing is applied several times to the time series.

The measurement of the Granger Causality and ultimately the Causal Density is done using the *Granger Causal Connectivity Analysis (GCCA) Toolbox* by Seth [28] with a model order of 10. The model order or number of lagged observations was determined via experimentation. The Granger Causality analysis was only able to detect the leader bird (i.e. a bird which all the others follow) with a model order of 10.

3.2.3 Experiments Conducted

With everything set in order experimentation was possible, but due to there being many questions associated with this study for example: Does the flocking factor control behaviour correctly? Are the time series generating methods correct? Is the causality itself being measured correctly?

Before proceeding with the actual experimentation a set of "acceptance" tests were conducted in order to verify that the measurement methods are accurate. For each acceptance test a leader bird was explicitly defined with its own predetermined path. The leader bird itself ignored all other forces and merely followed this path. All the other birds also ignored the primary forces and only followed the leader bird. This test was conducted 6 times for each measurement method. 3 of those test used 10 birds, 3 of which were indicators and the other 3 tests used 1000 birds, 10 of which were indicators.

A measurement method only "passed" the acceptance test if the analysis of the time series generated by that measurement method allowed for the detection of the leader bird (I.e. if the Granger Causality analysis showed the leader bird has having the greatest causal influence over all other birds). In other words if the leader bird is not detected, then it is likely that the measurement method is useless. This however was not the case and all the measurement methods were able to detect the unique leader bird. However, for the large scale test with 1000 birds detection did become increasingly difficult. Having passed these preliminary acceptance test all measures were then used for the 3 primary tests. The 3 primary sets of experiments were conducted each with the goal of minimizing uncertainty, confirming the reliability of the measurement method and of course measuring the causal density:

The first set of experiments were:

Similar to the acceptance test a leader bird was set and all the other birds followed the leader bird. Except in this case the leader bird did not have a predetermined path to follow, and all birds were confirmed to the model forces. One may argue that this is a repetitive test however the dynamics of these tests were far greater because of the inclusion of the normal model rules. These tests were conducted in order to determine if the leader bird could still be detected and could there be a situation where the other birds cause the leader bird due to the normal model forces. Remember

that in this the leader bird is not driven by a path but by the 3 primary social forces (alignment, separation and cohesion). This also served as a confirmation of the results of the acceptance test.. In this case 70 simulation runs were conducted (5 small and 5 large for each measurement method).

The small simulations involved 10 birds, 5 of which were indicator birds and it ran for 10000 ms. The large simulations involved 1000 birds, 10 of which were indicator birds, these simulations ran for 50,000 ms. The first 2000 ms were removed as as a precaution to ensure that the initial instability of the flock at the beginning of every simulation was not analysed. This instability simply refers to the amount of time the flock takes to stabilize into its desired behaviour. **Note:** For these "leader" experiments, the windowing technique was not used, simply because graphing out the causal relationships for a specific window only represents a specific time frame of behaviour. Ideally regardless of the time frame, the leader bird should cause all the other birds, but the leader bird is still driven by the rest of the flock according to the other social rules, this means that at specific points in time other birds may cause the leader bird.

The 2nd set of experiments were to determine the causal density measured at different flocking factor values. In this case 350 simulations runs were conducted (50 for each measurement method) each with 10 birds, 3 of which were "indicator" birds and each simulation run lasting 15,000 ms. Each simulation run used the exact same settings except for the flocking factor ξ , which was chosen randomly [0,10] at the beginning of each new simulation. In this case the first 5,000 ms of each simulation run were discarded. More of the beginning of the simulation is discarded in this case because the flocking exhibited in these trials was more complex, and as such it would take the flock a longer time to stabilize.

The 3rd and final set of experiments were to determine the causal density measured at different flocking factor values. In this case the flocks were much larger and flocking lasted for much longer. A total of 350 simulations were conducted (Again 50 for each measurement method) each with 1000 birds, 10 of which were indicator birds and each simulation lasting up to 50,000 ms. A large simulation time was allowed so that the behaviour over longer periods of time could be determined. Similar to the 2nd set of experiments, each simulation run used the exact same settings except with a randomly chosen flocking factor [0,10] also in this case, the first 5000 ms of the simulations were discarded.

4 Results

The following section is going to cover the results of each of the three experiments outlined in Section 3.2. The presentation of each set of results will be accompanied by an analysis of each of those results and what they signify. Due to the measurement methods being experimental, each method has to be carefully and critically analysed in order to determine if it accurately quantifies a bird's behaviour. The final discussion and comparison of data is left to Section 5.

It was found that the number of birds caused significant variations in the analysis of the dynamical complexity of the flock. In order to validate the outcome of the raw data, each experiment had to be conducted in both large and small scale to verify the correlation between dynamical complexity and the increase in the number of birds, as well as, to confirm the accuracy of the measurement methods.

4.1 Detecting The Leader

This section presents the results of the case where a leader bird (1) is set and all other birds follow this leader(1). The leader bird itself and all other birds still conform to the normal model rules (alignment, cohesion and separation). Meaning that the only difference between the leader bird and the rest of the flock is that the leader bird does not receive and guidance as far as its destination goes.

For the small scale test (10 birds, 5 of which are indicator birds) a sample of the results are shown in Figure 6 on page 30. As it can be seen in most cases, the leader bird (1) is detected by each measurement method with varying degrees of success.

According to these results, the weakest measurement method is the Average Difference of Velocity(Figure 6c). In this case, not only the leader bird is not detected, but another bird (4) is detected to be causing the leader bird. This could be perceived as this measurement method being inaccurate due to its inherent weakness (discussed in Section 6).

It is important to note that strong casual relations between the leader bird and all the other birds are still detected - thus, the underlying primary casual links are detected. Yet, the notion of the leader bird being caused by another could be viewed as inaccurate due to there being too few birds. If there were more birds however, the social forces could affect the leader bird more than the leader bird causes the rest of the flock .



Figure 6: Small Flock with Leader Results: The above figures shows which indicator birds have the greatest causal influence. A green arrow from x to y indicates x causes y. Red arrows indicate a strong equivalent causal relationship between two variables. All omitted connections mean that there is no significant causal relationship.

The Average Velocity(Figure 6a) and Average Difference of Position(Figure 6d) are the next best leader detectors. In both cases no other bird is causing the leader. The leader is causing at least one other bird ((5) and (4) respectively) and there are strong causal relations with most of the other birds. At this point we cannot say that a strong causal relationship between the leader bird and another bird is wrong because even just the act of a bird chasing the leader bird might be enough to create an equivalent (note: not greater) causal relation.

The Average Position(Figure 6b) and Difference of Speed(Figure 6e) are actually similar to the Average Velocity(Figure 6a) and Average Difference of Position(Figure 6d). Even though the results provided demonstrate them as stronger leader detectors (remember these are sample results), on average their performance is not any more accurate in terms of detecting the leader.

The Difference of Position From Home(Figure 6f) and Difference of Position From the Flock(Figure 6g) seem to be the best leader detectors. Every test agreed with the sample results being shown. The former detecting that the leader bird causes 4 out of 5 of the other birds and the latter showing that the leader bird causes all of the other birds (which is essentially the ideal result). Clearly there is a reason that Seth [27] used the center of the environment (i.e. home position) as a reference point. So far the results of this study indicate that the best way to measure a flock in order to generate a time series, is by using the home position and mean position of the flock as a point of reference.

The results of the large scale tests (1000 birds, 10 of which are indicators) are shown in Figure 7 on page 32. The first thing noticed is that every type of measurement returns a significant number of causal links between all the variables. This may seem counter intuitive to some but it is actually expected. By increasing the number of birds in the simulation, the social forces increase and since the social forces guide the correlation between each bird it is only obvious that the number of causal relations increases.

Increasing the social forces also impacts the direction of the causal relations. Essentially the leader bird is being chased by a much larger number of birds. Meaning that the repulsive force is substantially stronger. The question we are left with is, what does this mean in terms of detecting the leader?

If it is expected that the leader should still be found then the weak measures from the small scale are more useful, and the strong measures are no longer accurate *(This is might be unlikely)*. If on the other hand it is expected that the leader should not be found due to the increase in the repulsive force experienced, then the weak measures from the small scale remain inaccurate and the strong measures might still be reflecting what is happening in terms of the actual flocking behaviour.

In order to narrow down which measure is truly and accurately reflecting the behaviour of the flock in terms of locating the leader, a further analysis of how Granger Causality works, and how non-measured birds can skew the results must be looked into. This discussion is left to Section 5. The only way to truly prove if the strong measures from the small scale are still accurate for the large scale is to measure every bird in the flock. Unfortunately in flocks containing up to 3000 birds such an analysis would be infeasible due to the computational requirements.

Due to the contradictory results being shown from small to large scale, a series of mid scale tests were conducted. These mid scale tests were used to estimate if a strong measure in small scale remained so in large scale. Each of these test used flocks of 15, 20 and 25 birds under the same conditions which the small and large scale tests were conducted. Except in this case all birds were measured, which is why using larger flocks is not feasible.


Figure 7: Large Flock with Leader Results: The above figures shows which indicator birds have the greatest causal influence. A green arrow from x to y indicates x causes y. Red arrows indicate a strong equivalent causal relationship between two variables. All omitted connections mean that there is no significant causal relationship.

The results shown in Figure 8 are a sample of the midscale tests. These results were achieved using the *Difference of Position From Home* measurement method. Similar results were achieved using the *Difference of Position From the Flock* measurement method. These results demonstrate



Figure 8: Relationship Between Leader(1) and Followers(2-25) - As the number of birds increases, the social forces increase as well. Meaning that the leader bird is caused more by the other birds.

how the causal relationship between leader and followers slowly changes. As the number of birds increases, the birds perceived in each neighbourhood increase, which in turn means that the social forces are much stronger. This simply means that not only do more causal links start to occur, but the leader is more affected by its followers. This can be seen by comparing Figure 8b with Figure 8c. In the former the leader bird (1) is causing 5 out of 19 birds, so we can assume that it is still the primary driver. On the other hand in the latter case, the leader bird(1) is only causing 4 out of 24 other birds. It has become a significantly weaker driver of the flock. In fact another bird (25) is more involved in the flock dynamics. This gives credence to the theory that as the number of birds increases, detecting the leader should be more difficult. This issue is addressed in more detail in Section 5.

The causal analysis of the other measures showed no such behaviour. They essentially produced the same results as the small scale, i.e. the leader bird was detected but very weakly. At this point we can start to assume that even though all the measures passed the simple "acceptance" test, not all of them possess the necessary accuracy needed to measure the true flock dynamics.

4.2 Causal Density

This next section is the main focus of the study. As mentioned previously all this work has been don in order to simply measure the dynamical complexity of the flocking behaviour. The previous section has shown that there is a correlation between the number of birds and the social forces which occur within the model. The question we are left with is, since more social forces may result in more causal links what does this do for the causal density?

In order to answer this question, once again the testing had to be conducted on a large and small scale. The following sections will present an analysis of the results returned by those tests. By now one may ask as to why all the measures are still being used, even though the only ones which seem to be showing some accuracy are the *Difference of Position From Home* and *Difference of Position From the Flock* measurement methods.

The reason for continuing to use all the measurement methods is not only for the sake of being thorough, but for the fact that in the large scale leader tests they were more successful in detecting the leader. Whether this detection can be viewed as anomaly or inaccuracy, it may still be argued that these measures only become accurate when done on a large scale. This issue will be addressed in more detail in Section 5.

4.2.1 Small Scale Results

This section will present the results of the tests conducted in small scale. As mentioned in Section 3.2.3 the small scale tests lasted 15,000 ms, used 10 birds, 3 of which were indicator birds and used a randomly selected flocking factor [0,10] for each simulation run. The results will be presented in a graphical form plotting the flocking factor ξ of a specific simulation against the resulting causal density of that simulation (The average causal density and the maximum causal density of all the windows, will be shown in the figures respectively). Each dot in every graph represents a different simulation. The polynomial curve fits are used to show the relationship between the flocking factor (essentially the flocking behaviour) and the causal density.



Figure 9: Small Scale: Average Velocity Measured

The results shown in Figure 9 show the unit causal density returned when the average velocity is measured. The first half tends to agree with initial assumptions. With a low flocking factor, there is a random behaviour thus resulting in a low causal density. This causal density increases as the flock becomes more and more coordinated (i.e. as ξ increases).

At higher values for ξ the results dont seem very intuitive. At high values for ξ the flock moves as one. Meaning that it is essentially over integrated, thus a low causal density should (theoretically) be measured. According to Figure 9 this is not the case. Apparently as the flock moves more and more as one unit, the causal density increases. This could be due to some internal dynamics within the flock, but with a flock of such a small size no dynamic behaviour was visually observed.

Furthermore it could be argued that these results are inaccurate because the causal density is at its maximum when the flocking factor, ξ is between 7-8. This does not agree with the visual observations. The behaviour of the flock in the small scale tests tends to start stagnating with $\xi \approx 6$. Meaning there is very little difference between the behaviour of the flock for $6 \leq \xi \leq 10$. Essentially the only difference is that the flock tends to stary closer to the home position for high ξ values.

So the causal density for $6 \le \xi \le 10$ should be roughly the same, meaning there should be a curve at this stage. However this is not the case, rather it continues to increase. One possible reason for this increase is due to the social forces increasing (Again due to ξ). The average velocity might be inadvertently measuring the social forces, not the actual flocking behaviour caused by those social forces. Another reason is that the one of the limitations (discussed in Section 6) of the model or measure maybe skewing the results.

Figure 10 shows the causal density when the indicator birds are measured according to their average position. This result tends to agree with the results shown in Figure 9. In fact for $6 \le \xi \le 8$ the results seem to favour a high causal density.

At this stage if we take both the results of the measures together we get a rough maximum causal density when $6 \le \xi \le 8$. It was mentioned before that visually the flock behaviour differs little for $\xi \ge 6$. These result tend to agree with those visual observations. At the point where $\xi \ge 6$ the causal density has the same spread of values, roughly between 0.2 and 0.4 (Figure 10a).



Figure 10: Small Scale: Average Position Measured

The next set of results in Figure 11, which are retrieved using the difference of velocity shows an extreme climb. This is confirmed when comparing the curve fits. The curve fit equation in Figure 11a is very close to a perfect quadratic equation. (Relative to the other curve fits thus far).



Figure 11: Small Scale: Velocity Difference Measured

In terms of the average (Figure 11a) there seems to be no significance for when $\xi = 6$; the causal density continues to increase. However when observing the maximum (Figure 11b) causal density per window, we can see that there are many stages at which there is a high causality within the flock, most of which occur for $\xi \geq 6$.

The next set of results (Figure 12) which used the difference of position measurement are the opposite of the results in Figure 11. The former has a much more gradual climb in causal density.



Figure 12: Small Scale: Position Difference Measured

The reason for this sudden change in the slope is unknown. It might be due to a velocity being able to change quite drastically, while a bird's position can only change relative to where the bird currently is. However this change in slope is difficult to determine in the Large scale tests due to a few portion of birds being measured. So it could possibly be due to the different change rates but this cannot be confirmed since a low portion of the birds in the large scale are not measured.

If we once again observe the maximum causal density (Shown in Figure 12b) we notice that the causal density tends to remain at its highest (0.8) from $\xi \approx 6$. Clearly at $\xi \approx 6$ the behaviour of the flock begins to change and remains changed. The visual simulation agrees with this theory. At $\xi \geq 6$ the flock stagnates more and more.



Figure 13: Small Scale: Speed Difference Measured

The results of analysing the causal density according to the speed difference (Shown in Figure 13) seem to agree with all the findings and theories at this point. The question is how reliable are all these results?

It is noted in the previous section (Section 4.1), that the measures we have discussed up to this point are not very accurate at detecting the leader. Which could mean that all the results up to now are inaccurate. The next two results are from the (assumed) accurate measurement methods.

Figure 14 shows the causal density when the home/roost position is taken into account. In the previous cases the spread of the results at later values for ξ was quite minimal. However in this case the spread becomes very obvious. By looking at Figure 14a, we notice that for $\xi \geq 6$ there seems to be contradicting results with regards to what should be the causal density. According to these result, when the flock stagnates the causal density can be as high as 0.5 or as low as 0.1.



Figure 14: Small Scale: Difference from Home Measured

Furthermore by observing the maximum causal density (Figure 14b) there seems to be a random spread of causal values. Which is obviously not what we should be seeing, as all previous maximum measures correlated (at least to some extent) with their average counterparts. Note that there is a maximum causal density for both low and high ξ values. The question is, why is this measure returning such unexpected data?

The answer might lie in the flocking factor ξ . Remember as shown in section 3.1.2, the flocking factor directly influences a bird's desire to fly home. Normally this should not skew the behaviour due to there being a significant number of birds, which as has been stated, results in greater social forces. Which in turn results in a good balance between the social forces F_{social_i} and the other forces F_{other_i} . If we significantly reduce the number of birds, this balance is lost due to the social forces being significantly lessened. Meaning that increasing the flocking factor ξ will no longer result in a balanced force, but rather a stronger force to fly home.

The visual representation of the model confirms this. With a low number of birds, the birds tend to circle around the roosting site when ξ is high. While with a high number of birds the desired flocking is achieved. To summarise, due to the flocking factor being developed with a large number of birds in mind, it does not scale well to a very low number of birds. The reason this limitation is not shown in the other measures is because the other measures do not use the home position as a reference point, and this issue primarily concerns the home position.

The next results shown in Figure 15 are achieved when measuring the birds according to how far away each one is from the mean position of the perceived flock. The results in this case, seem to be more intuitive. The highest average causal density (Figure 15a) is achieved when the flocking factor ξ is between 4-6. Visually this makes sense since the most interesting flocking behaviour is observed between these values.

Once again there seems to be some significance when $\xi = 6$, except that in this case the causal density returns to a low value for $\xi > 6$. In the previous results maximum causal density is achieved for

 $\xi \ge 6$. The question that remains is, which is the correct causal density, i.e. when is true dynamical complexity achieved? Once the flock has stagnated, or before when there is a more organized yet not stagnated behaviour?

As already mentioned, intuitively we would expect the causal density to be low when the flock stagnates (Due to being over integrated). Meaning that causal density should be maximal when natural flocking occurs. In this case, Figure 15 agrees with the intuitive results, but none of the other analysis results agree with this. Even if we eliminate the results in Figure 14 due to the scal-



Figure 15: Small Scale: Position Difference from Main Flock Measured

	Measurement Method			
Causal Density	Average Velocity (Fig	Average Position		
-	9)	(Fig 10)		
Average	0.5929	0.40354		
Max	0.95798	1.0203		
	Difference of Velocity	Difference of Position		
	(Fig 11)	(Fig 12)		
Average	0.63801	0.61155		
Max	0.91585	1.0359		
	Difference of Speed	Difference of Home		
	(Fig 13)	(Fig 14)		
Average	0.61933	0.5838		
Max	0.92366	1.2723		
	Position Difference	From Flock (Fig 15)		
Average	0.4083			
Max	1.3179			

ing issue with ξ , we are still left with significant evidence against what one would expect. Clearly further analysis is needed.

Table 2: Small Scale: Norm of the residuals of the curve fits applied to data points. Values closeto 0 indicate an accurate curve fit.

As the reader may have noticed, each of the results has associated with it, a polynomial fit. The norm of the residuals of each of these fits is given in Table 2. Essentially the closer to 0, the more

accurate the fit. The purpose of associating a fit with data from these results is to determine if there is an accurate mathematical correlation between the behaviour of the flock (controlled by ξ) and the causal density.

Please note that results with a curve fit which have a low norm of residuals, does not immediately make it a more accurate analysis than one with a high norm. Rather an accurate fit means that all the data point are more consistent. The fit is simply to show a possible correlation between the flocking behaviour and causal density.

4.2.2 Large Scale Results

The following section will present the results of the large scale tests. As previously noted, for these tests each simulation lasted 50,000 ms, used 1000 birds, 10 of which were measured and each simulation used a randomly selected flocking factor [0,10]. These results are presented in a similar fashion to the Small Scale Results.

It is worth noting however, that due to there being more birds, the social forces within the model will be substantially stronger. This in combination with the limitation that the flocking factor does not scale well to small simulations; due to the imbalance discussed; means there will not necessarily be a direct correlation between the results of the small scale and the results of the large scale. In essence, the curves and causal density values may be substantially different.

The first set of results shown in Figure 16 are based on using an indicator bird's average position. The first thing that should be noted is that the causal density is substantially lower. With the highest average being 0.015 (Figure 16a) and the maximum in any window being 0.1 (Figure 16b). These values might seem strange in comparison to the small scale, but remember that in the small scale there were only 10 birds, 30% of which were measured. In the large scale only a tiny 1% of birds is measured.



Figure 16: Large Scale: Average Velocity Measured

This means that a large amount of non-measured birds affect the causal density in the following two ways:

1. Indicator birds can fly very far away from each other due to being pulled away by nonindicator birds. Which as a consequence means that these indicator birds can have little or no influence between each other. Since only the causal relations between the indicator birds are being analysed, low causal density will always be measured. 2. Even if indicator birds are not pulled away from each other; such as in the case of the stagnated flock (Refer to Appendix B.3 for a visual example); there can still be little causal density. This is because an indicator bird i is likely to be more influenced by all the non-measured birds between itself and another indicator bird j. It could be said that these non-measured birds break the causal relation between the indicator birds.

Obviously, the reasons above do not apply to the small scale. These low causal density values occur in all the other results as well, due to the above reasons being relevant to any large scale test using a small portion of measured birds.

Looking back at the results in Figure 16 we see that causal density peaks at $\xi = 4$, and for any greater values of ξ the accuracy is lost (i.e. some low causal values and some high causal values). If we observe the model we notice that at $\xi \approx 4$, the flock begins to stagnate (though not entirely). Just like the small scale result, the average velocity measure peaks at the point where the flock stagnates. However unlike the small scale tests there is a greater spread in the resulting causal density. The reason for this spread might be due to a particular limitation of the model (Discussed later in Section 6) which is emphasized when there are a large number of birds present.



Figure 17: Large Scale: Average Position Measured

The next set of results shown in Figure 17 show the causal density measured when an indicator bird's average position is used. These results are similar to the previous results. At $\xi = 4$ there is a greater spread in the causal values. In this case there is a *slight* favour towards a low causal density as the flock stagnates, but this could just be due to lack of simulation runs. They are also similar in that there is a similar spread of causal density, approximately between 0 and 0.015.



Figure 18: Large Scale: Difference of Velocity Measured

The causal density measured when a bird's difference of velocity is used is shown in Figure 18. These results show a much higher causal density then the ones discussed up to this point in the large scale tests. The 2nd highest being 0.04 when $\xi = 5$. In this case, at high ξ values the



Figure 19: Large Scale: Difference of Position Measured

the causal density tends to be low (with the one maximum exception at $\xi = 9.8$). As previously stated this is what should be seen. Once the flock stagnates the causal density should be low. However there seems to be a contradiction with these results. These results have a maximum causal density at $\xi \approx 5$ not 4. If the model is observed at $\xi = 5$, the only difference is that at $\xi = 4$ there tends to be more straggler birds (usually 1-10) than there is when $\xi = 5$. In essence the majority of the flock is still moving as one. Detecting a high causal density at $\xi = 5$ could be due to the spread of results that are seen once the flock starts to stagnate. This spread can be seen in nearly all the results, when $\xi = 6$ and $\xi = 4$ for the small and large scale tests respectively.

The next set of results shown in Figure 19 bring back an interesting observation. The slope of velocity measures is far greater than the slope of position measures. This was noted in the small scale results and now again in the large scale results. The rate of change may indeed affect the causal density results. However this is just an observation. There is no sufficient evidence in other studies or this one to prove that the rate of change or variation in data can impact the causal density.



Figure 20: Large Scale: Speed Difference Measured

The causal density measured when the difference of speed is used (Figure 20) once again, validates most of the points brought up at this stage. Causal density is maximum at the point when stagnation begins.

As mentioned in the previous section, due to the flocking factor not scaling well to very small sized flocks (≈ 100 birds), the results are not truly accurate. This was further confirmed by observing the spread of the maximum causal density at both low and high flocking factors. This was further analysed by conducting a series of mid range (500 birds, 10 indicators) tests using the home position as the reference point. These midrange tests showed similar results to the ones in Figure 21, the only difference being that stagnation occurred at larger values for ξ . In the large scale, this measure is more reliable due to the balance of forces being restored. As can be seen in Figure 21, there is no random spread of data points.



Figure 21: Large Scale: Position Difference from Home Measured

Notice how the results in Figure 21a and Figure 22a are very similar in terms of the overall relationship between the flocking factor and causal density. These results agree with each other



Figure 22: Large Scale: Position Difference from Flock Measured

more than any of the previous results. In this case maximum causal density (≈ 0.09) is achieved when the flocking factor is between 1 and 2 (Also note for Figure 21b: (≈ 0.5), and for Figure ??: (≈ 0.7)). The causal density is at its minimum when the flock stagnates ($\xi \ge 4$) and also when there is a random flocking pattern ($\xi < 1$). This agrees with the model visually, (As can be seen in Figure 23) the flock moves as one unit with a few stragglers.

However for the results shown in Figure 21a and Figure 22a, the causal density when the flock is moving randomly $\xi < 1$, is greater than the causal density once the flock begins stagnating $\xi > 4$. This is actually an expected result. For low values of ξ the birds no longer align with each other and don't seek the mean position, however they still repulse each other when they get too close. Meaning that to some degree they still influence each other. On the other hand, for high values of ξ they align with each other so much that that the repulsive forces become negligible. All this means that the flock can easily become more integrated rather than more segregated. Thus the difference in causal density.

A final observation about these 2 results is that the causal density is much higher than that made by the previous measures; not only in the average



Figure 23: Large Scale: Flocking behaviour with $\xi = 4$

causal density but also in the maximum per window causal density. Indeed the previous measures were strangely low. Lower than expected even taking into account the fact that larger flocks result in lower causal density due to a greater distance between indicator birds.

Just as with the small scale tests, a curve fit was associated with each of the results. The norm of the residuals of these curve fits is shown in Table 3. Notice that in this case most of the curve fits are more accurate. This accuracy might be due to the flocking factor being used in such a way that it does not affect the balance between the social forces and the roosting/home force. Please note that for the difference of home and the difference of the position from the mean of the flock, a polynomial of the 8th order was fitted to the data points.

	Measurement Method				
Causal Density	Average Velocity (Fig	Average Position			
-	16)	(Fig 17)			
Average	0.021192	0.025138			
Max	0.095911	0.11816			
	Difference of Velocity Difference of Positio				
	(Fig 18)	(Fig 19)			
Average	0.055774	0.021846			
Max	0.22533	0.15178			
	Difference of Speed	Difference of Home			
	(Fig 20)	(Fig 21)			
Average	0.032643	0.07854			
Max	0.12444	0.50409			
	Position Difference From Flock (Fi				
Average	0.054992				
Max	0.59654				

Table 3: Large Scale: Norm of the residuals of the curve fits applied to data points. Values closeto 0 indicate an accurate curve fit.

4.2.3 Combining The Results

Taking the large and small scale tests into account, leaves us with two contradicting theories about the dynamical complexity exhibited by the flocking behaviour of the model. These theories are:

- 1. Maximum Dynamical Complexity is achieved at the point where the flock starts to stagnate (Small: $\xi \approx 6$ and Large: $\xi \approx 4$) but then trails off once the flock has over stagnated, to the point where the whole flock circles around the roosting position (As shown in Appendix B.3). This case is slightly unlikely due to only some of the large scale tests supporting it.
- 2. Maximum Dynamical Complexity is achieved from the point at which the flock begins to stagnate, and this maximum continues to be measured well past that point (Small: $6 \le \xi \le 10$ and Large: $4 \le \xi \le 10$). This case is even more unlikely, as only the small scale results provide support for such a scenario.
- 3. Maximum Dynamical Complexity is achieved when the flock shows a minor degree of coordination, dynamically forming various sub-flocks but never merging into one large flock (Small: $\xi \approx 5$ and Large: $1 \le \xi \le 2$). An example of this behaviour is shown in Section B.2. This case is only supported by two measures, however those measures were estimated to be accurate (from the leadership tests) and this case appeared in both large and small scale tests.

With three contradicting theories, a deeper discussion is necessary.

5 Discussion

This section will address the various conflicting issues arising from the contradicting results. While some of the issues which have been brought up will be addressed fully, others may require further study.

The first issue that must be addressed is brought up in Section 4.1. It was noted that only two measures (the difference from home and the position difference from the flock) were able to detect the leader when all the model forces were used. Why then were the other measures able to detect the leader during the small acceptance tests ?

These acceptance tests were very basic in nature and the model used to conduct them even more so. In these acceptance tests all of the forces defined in Section 3.1 were removed. Meaning that all of the non-leader birds only had one force to drive them i.e. the desire to follow the leader. The leader itself also only had one force, which was to follow an explicitly defined path.

A significant issue in the model is brought up at this point; the model does not explicitly include collision detection. Meaning that all collision avoidance is achieved implicitly by the separative social force. By removing this force all birds can collide, and even overlap each other perfectly given some time.

This means that in these acceptance tests every non-leader bird would follow the leader in the exact same way; using the same velocities, being in the same positions etc. . .

What this means is that the time series produced by all the indicator birds (except the leader) became equivalent, due to all non leader birds overlapping each other after only a small amount of time ($\approx 500ms$). As such the Granger Causality is essentially only being measured between two variables; the leader bird and all the other birds who have compressed into the same position. Regardless of how many birds were included in the model, the leader would always be detected due to the analysis being very simple. Meaning that every measure could pass the acceptance test easily.

By reintroducing the social forces, the birds no longer overlap each other perfectly. In other words, the Granger Causality is being performed between all the indicator birds as designed. It is at this point that most of the measures fail to detect the leader bird.

A second issue regarding the leader detection must now be addressed. When more birds were added to the model, the two "accurate" measurement methods did not result in a collection of time series which allowed for the detection. Furthermore the measures which failed to detect the leader in the small scale were able to detect the leader in the large scale.

Due to the natural difference between large flocks and small flocks, leader detection using Granger Causality should actually become more difficult. As Shown in Figure 8 on page 33, as the size of the flock increases, leader detection becomes more difficult. This is not only due to the fact that not all birds can be measured but is also due to the natural behaviour in the model.

Observe Figure 24a, which shows an example of what a typical leader orientated flock might look like in the model. (Coincidentally the leader orientated model seemed to favour a 'V' shape flying pattern, similar to that demonstrated by ducks, when migrating). What can immediately be seen is that there is a relatively small number of birds, and the direct distance between a leader bird(red) and an indicator bird(black) is minimal. Also notice that there are very few non-measured birds(grey) between the leader and the indicators. This means three things:

1. The leader bird can influence all of the birds more easily because there are too few birds to create a substantial force to influence the leader bird. It is important to remember that the leader bird is still driven by cohesion, alignment and separative forces which impact where the bird goes. If there are too few birds then those forces are minor in comparison to the



Figure 24: *Flocks With a Leader.* The red dot is the leader bird being followed, the grey dots are non-measured birds, the black dots are measured indicator birds and the shaded blue region represents the free area with little repulsive forces, which the leader bird is able to through. The angle θ represents the angle of the surrounding area of the leader bird, which the leader bird is able to move through.

other forces.

- 2. The leader bird also has more freedom of movement (As shown by $\theta = 270^{\circ}$). Meaning that it can turn and move more freely, essentially making it a highly independent bird.
- 3. The causal relation between the leader and the indicators is stronger due to there being little external influence from non-measured birds. Making it easier for the leader bird to have a direct influence on the indicator birds.

All of the above factors make leader detection simple in the small scale case. Now observe Figure 24b, in this case there is a variety of differences.

In this case there are a larger number of non-measured birds, which as a consequence result in there likely being a greater distance between the leader and the indicators.

This entails the following:

- 1. The leader bird can easily receive more influence than it puts out. Due to a large number of birds, the social forces are greater. This means that the leader bird experiences larger cohesion, separative and alignment forces. As a consequence the leader is driven more by these forces (especially the repulsion of the separative force). Therefore, the leader is more driven by the flock.
- 2. The leader has less freedom of movement (as indicated by $\theta = 180^{\circ}$, which essentially makes it a less independent bird.
- 3. Due to the greater distance between the leader and any indicators, as well as many nonmeasured birds possibly occurring between the leader and the indicators, it is difficult for the leader bird to influence the indicator birds. Furthermore, just as the leader bird receives more social forces, the other flock members receive more social forces as well, meaning that an indicator bird is more likely to be influenced by a close neighbour rather than a far away leader.

Taking all of the factors above into account it is easy to see how detecting the leader becomes unlikely. The "leader bird" itself is also less of a driver for the flock. In other words the causal influences from the leader to the other birds become significantly lessened. To summarise, large flocks make the "leader bird", less of a leader all together.

This explains why detecting a leader becomes difficult for large flocks. The issue which remains is, why some measures were able to detect the leader in the large scale but not the small scale. This required further investigation and upon careful analysis an anomaly was discovered.



Figure 25: Leader analysis using a window of 500 ms with a 250 ms offset. These result are based on using the average velocity measurement.

Due to this question, each of the measures under went additional analysis. For this additional analysis the same data used to produce the small and large scale results shown in Figure 6 and Figure 7 were analysed using the windowing technique. For each window a graph was produced to indicate the causal relations between each of the variables at different stages in time (A sample of which is shown in Figure 25).

The purpose of this analysis is to determine more accurately if the leader can be detected. The difference from home and the position difference from the flock, retrieved similar results to the final ones shown in the Figures (i.e. the leader was detected in small scale but not large scale). However, the other measures showed that they were not able to detect the leader in small or large scale when using the windowing technique (An example of this lack of detection in large scale is shown in Figure 25). This could mean that these measures detecting the leader in the large scale could possibly be an inaccuracy or anomaly. If so, then it seems that the only measures used in this study which truly reflect the behaviour of the flock might be the difference from home and the position difference from the flock.

Having dealt with the issues brought up in detecting an explicitly defined leader, all the issues now

facing the causal density analysis must be addressed. At the end of Section 4.2.3, three contradicting theories were noted based on the results given. If we eliminate the "non-accurate" measures according to the conclusions of the Leader tests then we are left with only two measures (the difference of home and the position difference from the flock). Consequently this leaves us with only one possible conclusion - that dynamical complexity is maximal when the flock dynamically forms various sub-flocks which coordinate with each other in a subtle but significant manner, resulting in a "beautiful" aerial demonstration and balance between integrated and segregated movement. While this would be the theoretically expected result we cannot simply ignore the other measures. Rather some further investigation is required.

The first two theories are at odds as to whether causal density remains high once the flock stagnates, (as shown by Figures 9 - 13) or rather causal density begins to tail off to a minimum once the flock has completely stagnated (as shown by Figures 16 - 22). Both have some merit to them: It could be assumed that causal density remains high when the flock stagnates due to some subtle dynamics occurring within the flock. This could simply be a result of the extreme social forces experienced when all the birds are tightly packed together. On the other hand, it is also possible that due to the flock moving as one unit, it is essentially over integrated resulting in a low dynamical complexity.

In order to determine what actually happens to the causal density at the point of stagnation, one more set of experiments was conducted. In this case, the model was once again simplified by removing all the social and home forces, leaving only the speed control force to maintain velocity at a realistic pace. This essentially removes any ability for birds to influence each other. At this point all the birds were given the exact same path to follow (Note the similarities to the leader test).

In these tests 10 birds were used, all of which were measured and the simulation lasted 10,000 ms. The result of this simulation is a trail of birds moving from one point to the next, all relatively close to each other and some even overlapping. It can safely be said at this stage that this behaviour is similar to the behaviour exhibited at a high flocking factor ξ ; the only difference being that the primary forces are not used and the path is explicitly set. It can also be clearly stated that at this stage no dynamics are occurring and the flock is clearly over integrated, meaning, that a low causal density should be measured. This above scenario was performed and the causal density analysed for each measurement method. The results of this are shown in Table 4.

Causal Dansity	Measurement Method			
Causal Density	Average Velocity	Average Position		
Average	0.0117283951	0.0154320988		
Max	0.0333333333	0.0333333333		
	Difference of Velocity	Difference of Position		
Average	0.0555555556	0.0086419753		
Max	0.0666666667	0.044444444		
	Difference of Speed	Difference of Home		
Average	0.0117283951	0.0007209877		
Max	0.0333333333	0.0111111111		
	Position Difference From Flock			
Average	0.000617284			
Max	0.0111111111			

Table 4: Causal Density: The above shows the causal density measured when the flock is forcibly stagnated. Such that all the flock members follow the exact same predetermined path

Observe that no value in table 4 is 0 as expected, however some values are higher than others. All of the values are not 0 most likely due to every measurement method having weaknesses. From

these results we can clearly see that a low causal density value should be returned when the flock stagnates, which means that for high values for the flocking factor ξ , the results in Figures 9 - 13 are inaccurate. Meaning that the theory which states that causal density remains high once the flock stagnates is not feasible.

Now the question we are left with is, when does the flock demonstrate maximum dynamical complexity? At the point just before stagnation begins or at the point where the birds have merged into sub flocks but not one large flock (Large scale: $1 \le \xi \le 2$; Small Scale: $4 \le \xi \le 6$). In order to finally answer this question, an accurate measurement method must be agreed upon. Thus far there is a significant amount of evidence which indicates that the home difference and position difference from the flock measures are more accurate than the other measures. There is one final piece which shows that these measures are likely the most accurate, meaning that the results achieved by using them (except for the results in Figure 14 due to ξ) are the correct results.

By incrementally reducing the length of the simulation to 6000 ms, thereby limiting the complexity of the path followed by the birds, a causal density of 0 is measured when the flock is measured according to how far the indicator birds are from the mean flock position. For all the other measures however, a non-zero causal density was not measured. Taking this into account together with all the previous conclusions about the measures, it is very likely that the accurate method of measuring the behaviour of a flock to produce a collection of time series is to use each bird's current distance from either the perceived mean position of the flock or the roosting position (As similarly done here and by Seth [27]). Please Note: That is that the difference from home measurement method, while not measuring a causal density of 0, does result in substantially lower causal densities in comparison to the others.

So it is very likely that the flocking behaviour demonstrates a good balance between integrated and segregated activity (i.e. a high dynamical complexity) when it is the most visually compelling. That is, the point where various dynamically changing sub flocks are formed; where each acts independently but still influences the whole flock in a subtle manner, but never so much as to cause all the sub flocks to merge into one.

Another observation which can be noted, is that firstly, it seems the best way to measure a flock in order to generate a collection of time series is to use a point of reference such as the mean position of the flock or a roost position. Secondly the measurement results should preferably be unbounded and continuous. It is probably because of this second case that the speed difference measure was inaccurate; a flock possesses a maximum and a minimum speed (real or simulated), measuring the flock based on such terms, limits the unique states each bird can produce. However when measuring the position in a non-restricted environment a bird is not limited to a finite set of states, and as such is more likely to return an accurate measurement of the bird's behaviour.

6 Limitations

It was mentioned several times throughout this paper that the model as well the the measurement methods used to generate the time series have some weaknesses. Some of these weaknesses may have already been mentioned or briefly discussed. This section will cover them in detail.

The biggest issue with the model developed, is that in many ways it is not a biologically accurate model (at least not as biologically accurate as the model by [14]). Firstly, the roll is omitted, while this may not affect a bird's field of view it does mean that the aerodynamic forces associated with banking are also omitted. Obviously this limits a bird's true freedom of movement. Another limitation is that a global cohesive force had to be added to the model in order to allow for stagnation to occur. While this is actually a necessity for the purpose of the study, it does introduce unrealistic forces.

Another weakness which was briefly mentioned, is that the model does not explicitly use collision detection. Rather collision avoidance is achieved via the separative social rule. While this rule does limit the amount of collisions which do occur, it is still possible for birds to collide and even overlap each other. As a consequence this can, and as previously discussed did, skew the results in a minor way.

The roosting position also introduces a few weaknesses. Firstly it should be noted that the roosting position was a necessity in order to give birds a reason to change direction, and without this force natural flocking could not be achieved. However the introduction of this force did allow for more stragglers to occur in the model for mid flocking factor values $4 \le \xi \le 7$. The problem is for mid level ξ values, there is not enough force to drive a bird to indefinitely follow the flock, nor at the same time is there enough force for the bird to be pulled away from the home when no other birds are around. As a consequence it is possible for a few birds, roughly 1-10, to be left behind at the roosting position when the whole flock moves as one. The primary reason is simply due to these birds not perceiving other birds enough to be drawn away from the home position.

The home position itself is not biologically accurate. Real birds flock around a roosting site which covers a large area and not one specific position. While one may argue that this is not much of a limitation due to it only being possible to simulate a finite of number of birds, making large roosting area unnecessary, a roosting area as opposed to roosting point (0, 0, 0) would be more accurate.

Whilst discussing the environment itself, it is worth noting that a biologically real model should include some restrictions. Namely there should be a "ground" signifying the lowest Y value a bird's position can have, as well as a "sky limit" indicating the maximum height (Y) the bird can move in the model. The introduction of this model would introduce more non-social forces and as previously noted, restricting the flight area of the birds can produce interesting flocking patterns.

While a maximum velocity and momentum were added to limit unrealistic movement. Observing the simulation reveals it is still possible for some unrealistic movements to occur. For example when a bird is drawn back to the roosting position it is drawn back with so much force that the bird will rapidly decelerate and turn with a very low turning circle.

Another issue which has already been mentioned is the lack of balance in forces when there are few birds in the model. This limitation is only shown when the flocking factor is high; for a low or middle flocking value, this limitation is not relevant. The reason is that as the flocking factor grows, its effect on the social forces depends on the number of birds, while its effect on the roosting force is consistent. As a consequence when there are too few birds (≈ 100), as the flocking factor grows the roosting force grows faster than the social forces. On the other hand if there are a sufficient amount of birds (roughly 100 birds) then the growth rate for the social forces is equivalent to the growth rate of the roosting force. That is to say both growth rates are at their maximum.

This issue is simply due to the fact that achieving complex control over a flock of birds, where each bird essentially has the ability to act independently is difficult using only one parameter. Even in the study conducted by Seth [27], a single parameter was not used; rather, Seth used 3 explicitly defined settings to achieve random, natural and stagnated flocking. Each of these settings affected a wide range of factors, much like the flocking factor in this model but each setting was compromised of various parameter values, each affecting a different component of the model. A final limitation of the model used in this study is that aerodynamic forces such as lift and drag were not incorporated

A large question which arose in this study was, what is the best way to measure the birds in order to generate a time series which quantifies the behaviour of each bird, in such a way so that each bird is uniquely quantified. This in essence was the backbone of the study. If the flock could not be measured accurately, the dynamical complexity in turn could also not be measured accurately. This study proposed seven different methods of measuring the flock, all of them with limitation which will now be covered.

The measures all share a common problem, each bird's unique identification usually depends on a minimum of three values. Compressing these three values into a singular number which uniquely identifies the bird introduces ambiguity. For example, if Figure 26 represents a simple two dimensional world with only two birds it is easy to see the weaknesses of some measures. The average position takes a bird's n-dimensional position and calculates the average by summing each coordinate and dividing by n. So for the case in Figure 26 the red bird and the blue bird are identified as 2. In this case the birds are not uniquely quantified. The average velocity posses the exact same problem, except that the velocity vector is used rather than the position vector. So if a bird i has a velocity $v_i(t) = (5,3)$ and a bird j has a velocity $v_i(t) = (3,5)$ then both are measured as 4.



Figure 26: Simple 2D representation of two birds in a grid world. The red dot's position is (3, 1). The blue dot's position is (1, 3)

The difference of position posses a similar problem; let bird i (the red dot) have the velocity $v_i(t) = (1, 1)$ and let bird j (the blue dot) have the velocity $v_j(t) = (-1, -1)$. At the next time step their positions will respectively be $p_i(t+1) = (4, 2)$ and $p_j(t+1) = (0, 2)$. At this stage they are both in different positions and used different velocities to get to there. However, from one time step to the next, they covered the same distance and as such, are measured in the same way. The difference of velocity, has the exact same problem. For example if at time step t the velocities are $v_i(t) = (1, 1)$ and $v_j(t) = (3, 5)$, and at the next time step t+1 the velocities change to $v_i(t) = (0, 1)$ and $v_j(t) = (4, 5)$, the difference of velocities are the same, and once again two birds who could essentially be at opposite ends of the world are measured to be at the same.

The other 3 measures are less likely to cause ambiguity due to using a reference point, however there are circumstance in which ambiguity can be still occur. For the difference of home, the situation in Figure 26 reflects such a case. Both birds are an equal distance from the home and as such are

measured to be the same. For the difference of speed, two birds are likely to be measured the same when all birds have a maximum or minimum velocity (which occurs often in the stagnated case). This is because even the perceived mean velocity will be the same for both birds, which in turn means the difference of speeds will be the same.

The final measure, which takes the difference between a bird's current position and the mean position of the flock, is highly unlikely to introduce ambiguity. This is because the mean position of the flock is a perceived mean (i.e. all the **other** birds, not including the bird itself, are taken into account). So every movement by every bird affects how all the birds are measured, which means birds are measured more uniquely. However even this measure might posses a limitation.

7 Future Work

As previously mentioned the field of collective animal behaviour is an interesting one. Various models exist which simulate different animal group characteristics and behaviours. Dynamical complexity is also an interesting field with the potential to be applied to many topics. Taking these two facts into account, this section will highlight some future work which is possible by combining these two fields once more.

The most obvious in terms of avian flocking is to incorporate a more biologically realistic model, such as the one by [14]. An interesting challenge would be to analyse in more detail how flocking can be controlled using a single parameter which affects various factors. This paper introduced the "flocking factor", but unfortunately this sweeping parameter does not scale well to smaller flocks. As such, the development of a scale free behaviour altering parameter would be interesting, not only for the sake to make similar analysis like this more accurate, but also to better understand the fine relationship between the social forces which drives flocking.

It was mentioned in this paper that there seems to be some correlation between dynamical complexity and the size of the flock. This was never analysed in detail, and it was just observed that greater flocks create stronger social forces which may result in an earlier stagnation point. Determining accurately how the number of birds correlates to dynamical complexity would allow us to better understand how the size of a system can affect the analysis of its complexity.

In Section 2.2 several different techniques for measuring dynamical complexity were studied. This means that this study could essentially be "redone" using a different complexity measure (for example Interaction Complexity). Since the most critical issue of how to measure a flock accurately has already been answered, such a study could easily be performed. Even the simulation environment developed could be reused to aid in such a study. Furthermore the results of the study using Interaction complexity could possibly further confirm the findings in this paper.

There is no reason why analysing dynamical complexity should be limited to this specific type of flocking. The analysis of other flocking models, such as the predator-prey model would extend the possibilities to another dimension. It would be interesting to determine if there is a correlation between the complexity of the prey flock and the complexity of the predator flock. For example, would the dynamical complexity of both be high at the same time, or would one flock having a high dynamical complexity result in a low dynamical complexity for the other?

This could be easily measured in terms of the degree to which the predators chase the prey. Would the predators chasing the prey result in high dynamical complexity for the predators due to the coordination required ? Secondly would the dynamical complexity of the prey be low due to being scattered by the predators?

The answers to these questions would be interesting and would possibly introduce a new element to typical dynamical complexity analysis (for lack of a better term, competitive dynamical complexity or correlating dynamical complexity).

There is also no reason why dynamical complexity analysis should be limited to avian flocking. It was mentioned before that the swarming behaviour of bees and the shoaling of fish demonstrate similar properties; as such an analysis for those animal aggregations would also be possible. The results of analysing those models could possibly serve as further agreement with the results in this paper.

Though not explicitly "flocking" an interesting behaviour to study is the "pre-flocking" behaviour. Just before birds take off and commence their flocking, there is a period of "uncertainty" where only a few birds take off and then land again due their small number. The question is, at which point in this "uncertainty" does the whole flock decide to take off and commence flocking, and secondly is there some complex behaviour in this preliminary stage. These questions not only produce the

problem of discovering point at which flocking commences, but also if this point correlates to a significant dynamical complexity.

The various studies suggested in this section are not only based on fixing the limitations of the model defined, but also on taking the concept further. These suggestions are only a few of many which might be possible. Since dynamical complexity is still a relatively new concept and the techniques used to measure dynamical complexity are also relatively new, there are many possibilities to the fields to which this measure can be applied.

8 Conclusion and Summary

In this project natural flocking and swarming was studied in order to better understand why this behaviour occurs in nature, and what is the significance of this behaviour. While natural flocking is beautiful to observe a simulation model was needed in order for a detailed analysis to be conducted.

Many different simulation models of varying complexity and realism were studied. Having found that none of these existing models provided a complex control over flocking using only one parameter a new model was developed to allow for such control. The model developed drew much inspiration from the current models in existence, as such, while not being the most biologically accurate model, it did allow for the flocking behaviour to be swept between random, natural and stagnated using a single newly designed parameter, the flocking factor.

In order to analyse the dynamical complexity demonstrated by the flock, various complexity measures were studied. Having studied Interaction Complexity, Information Integration and Causal Density the measure chosen was Causal Density. The reason for choosing this measure is because it uses Granger Causality which allowed for a mid-level set of tests to be conducted to determine if an explicitly set "leader bird" could be detected.

Before being able to conduct any analysis relating to dynamical complexity the flock members had to be measured in an accurate way. Due to there being a lack of material indicating exactly what is an accurate technique for measuring flocking, various experimental techniques were used. These experimental measurement methods resulted in contradicting causal densities.

However through carefully planned simulations, tests, studies and analysis of the various measurement methods, some were slowly proven to be inaccurate. By process of eliminating the inaccurate measurement methods while strengthening the accurate methods. Two measurement methods remained which could not be proven to be inaccurate. Both of them producing similar results and one of which was used in a similar study by Anil Seth [27].

By proving these measurement methods to be correct, this study can conclude with the following statement regarding flocking and dynamical complexity.

Maximum dynamical complexity is achieved when various constantly changing sub-flocks are dynamically created by the birds. Each of these sub-flocks having the ability to act independently whilst still influencing the flock in a subtle but strong manner, never to the point where the subflocks merge thereby stagnating the flocking patterns.

Acknowledgements

I would like to thank the following people without whom this study would have not been possible.

My supervisor Professor Murray Shanhan goes my most sincere gratitude for his endless patience, guidance and advice, and for giving me the opportunity to take part in this interesting field, allowing me to conduct a truly unique and diverse study. It has been a fascinating journey which combined my passion for computational intelligence with the theory of swarming and flocking.

A special thanks to Chris Reynolds for the flocking model he developed and Anil Seth for the Granger Analysis toolkit which greatly helped me with this study.

My mother Aldina Peliz for her endless support and motivation, as well for helping me proof read my thesis.

Ian McLaren, my stepfather for his huge help and support throughout this journey.

My father Carlos Pinto Teixeira for the motivation and financial support which enabled me to study at this prestigious institution.

Christine Seifried, for the motivation, support and valuable advice regarding the structure and overall format of this document.

Last and not least, Imperial College. This has been the most interesting, challenging and most enjoyable experience of my academic life.

References

- [1] Iztok Lebar Bajec, Nikolaj Zimic, and Miha Mraz. Simulating flocks on the wing: the fuzzy approach. *Journal of Theoretical Biology*, 233(2):199 220, 2005.
- [2] Michele Ballerini, Nicola Cabibbo, Raphael Candelier, Andrea Cavagna, Evaristo Cisbani, Irene Giardina, Alberto Orlandi, Giorgio Parisi, Andrea Procaccini, Massimiliano Viale, and Vladimir Zdravkovic. Empirical investigation of starling flocks: a benchmark study in collective animal behaviour. Animal Behaviour, 76(1):201 – 215, 2008.
- [3] Partridge BL. Structure and function of fish schools. Scientific America.
- [4] S. Camazine, P. K. Visscher, J. Finley, and R. S. Vetter. House-hunting by honey bee swarms: collective decisions and individual behaviors. *Insectes Sociaux*, 46:348–360, 1999. 10.1007/s000400050156.
- [5] Andrea Cavagna, Alessio Cimarelli, Irene Giardina, Giorgio Parisi, Raffaele Santagati, Fabio Stefanini, and Massimiliano Viale. Scale-free correlations in starling flocks. *Proceedings of the National Academy of Sciences*, 107(26):11865–11870, June 2010.
- [6] Marco Dorigo, Eric Bonabeau, and Guy Theraulaz. Ant algorithms and stigmergy. Future Gener. Comput. Syst., 16:851–871, June 2000.
- [7] Walter. Enders. Applied Econometric Time Series. 1995.
- [8] A.P. Engelbrecht. Computational intelligence: an introduction. John Wiley & Sons, 2007.
- C. W. J. Granger. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3):424–438, 1969.
- [10] C. W. J. Granger. Causality, cointegration, and control. Journal of Economic Dynamics and Control, 12(2-3):551 – 559, 1988.
- [11] Charlotte K. Hemelrijk and Hanno Hildenbrandt. Self-organized shape and frontal density of fish schools. *Ethology*, 114(3):245–254, 2008.
- [12] F. Heppner. Three-Dimensional Structure and Dynamics of Bird Flocks. In J. K. Parrish and W. M. Hamner, editors, *Animal Groups in Three Dimensions*, pages 68–89. Cambridge University Press, Cambridge, UK, 1997.
- [13] F. Heppner and U. Grenander. A stochastic nonlinear model for coordinated bird flocks. In E. Krasner, editor, *The ubiquity of chaos*, pages 233–238. AAAS Publications, 1990.
- [14] H. Hildenbrandt, C. Carere, and C.K. Hemelrijk. Self-organized aerial displays of thousands of starlings: a model. *Behavioral Ecology*, 21(6):1349–1359, 2010.
- [15] Daniel J. Hoare, Graeme D. Ruxton, Jean-Guy J. Godin, and Jens Krause. The social organization of free-ranging fish shoals. *Oikos*, 89(3):546–554, 2000.
- [16] Jr. John T. Emlen. Flocking behavior in birds. The Auk, 69:160–170, April 1952.
- [17] Dervis Karaboga. An idea based on honey bee swarm for numerical optimization. Technical Report TR06, 346, 2005.
- [18] Dervis Karaboga and Bahriye Akay. A survey: algorithms simulating bee swarm intelligence. Artificial Intelligence Review, 31:61–85, 2009. 10.1007/s10462-009-9127-4.
- [19] J. Kennedy and R. Eberhart. Particle swarm optimization. In Neural Networks, 1995. Proceedings., IEEE International Conference on, volume 4, pages 1942 –1948 vol.4, nov/dec 1995.

- [20] Martin Lindauer. Schwarmbienen auf wohnungssuche. Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 37:263–324, 1955. 10.1007/BF00303153.
- [21] Millonas MM. Swarms, phase transitions, and collective intelligence. Articial life III, page 417445, 1994.
- [22] S. Momen, B.P. Amavasai, and N.H. Siddique. Mixed species flocking for heterogeneous robotic swarms. In EUROCON, 2007, pages 2329 –2336, sept. 2007.
- [23] H. Rackham. Plinys Natural History Book X:XXXV. Harvard University Press, 1933.
- [24] William T. Reeves. Particle systems a technique for modeling a class of fuzzy objects. ACM Transactions on Graphics, 2:359–376, 1983.
- [25] Hauke Reuter and Broder Breckling. Selforganization of fish schools: an object-oriented model. *Ecological Modelling*, 75-76:147 – 159, 1994. State-of-the-Art in Ecological Modelling proceedings of ISEM's 8th International Conference.
- [26] Craig W. Reynolds. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '87, pages 25–34, New York, NY, USA, 1987. ACM.
- [27] A. Seth. Measuring emergence via nonlinear Granger causality. In S. Bullock, J. Noble, R. Watson, and M. A. Bedau, editors, Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems, pages 545–552. MIT Press, Cambridge, MA, 2008.
- [28] Anil K. Seth. A MATLAB toolbox for Granger causal connectivity analysis. Journal of neuroscience methods, 186(2):262–273, February 2010.
- [29] Anil K. Seth, Eugene Izhikevich, George N. Reeke, and Gerald M. Edelman. Theories and measures of consciousness: an extended framework. *Proceedings of the National Academy of Sciences of the United States of America*, 103(28):10799–10804, July 2006.
- [30] Murray Shanahan. Dynamical complexity in small-world networks of spiking neurons. Phys. Rev. E, 78(4):041924, Oct 2008.
- [31] H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Flocking in fixed and switching networks. Automatic Control, IEEE Transactions on, 52(5):863-868, may 2007.
- [32] H.G. Tanner, A. Jadbabaie, and G.J. Pappas. Stable flocking of mobile agents, part i: fixed topology. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 2, pages 2010 – 2015 Vol.2, dec. 2003.
- [33] Giulio Tononi. An information integration theory of consciousness. BMC Neuroscience, 5:1–22, 2004. 10.1186/1471-2202-5-42.
- [34] Giulio Tononi, Gerald M. Edelman, and Olaf Sporns. Complexity and coherency: integrating information in the brain. Trends in Cognitive Sciences, 2(12):474 – 484, 1998.
- [35] Gerhard Venter. Particle swarm optimization. In AIAA Journal, pages 1583–1589, 2002.
- [36] January Weiner. Norberg u. m. vertebrate flight. mechanics, physiology, morphology, ecology and evolution. springer verlag (xiv) + 291 pp. dm 238. isbn 3-540-51370-1. Journal of Evolutionary Biology, 5(3):541–543, 1992.
- [37] Vladimir Zhdankin and J. C. Sprott. Simple predator-prey swarming model. *Phys. Rev. E*, 82(5):056209, Nov 2010.

Appendix

A Simulation User Guide

A.1 Getting Started

In order to run the application in Linux, Windows or Mac you must ensure that you have installed the latest Java Runtime Environment. Which can be downloaded here:

http://www.oracle.com/technetwork/java/javase/downloads/jre-6u25-download-346243.html

If the Java Runtime Environment is already installed then you can proceed to accessing the command line. Once the command line has been accessed navigate to the directory where ExtBoid-Sim.jar and Core.jar are saved and input into the command line:

- Windows "java -cp .;core.jar;ExtBoidSim.jar sim.Simulation"
- Linux and Mac OS X "java -cp .:core.jar:ExtBoidSim.jar sim.Simulation"

The welcome screen shown in Figure A.1 should then appear. From this screen the main control screens can be accessed via the side buttons.



Figure A.1: The Welcome Screen

A.2 The User Interface

The following sections will cover the various UI screens and their usage. Please note it is possible to switch between these windows by using the control panel on the right of each of these screens.

A.2.1 The Simulated World



Figure A.2: Simulation Window

Figure A.2 is an example of the simulated world which visually shows the flocking behaviour of the birds. The while lines coming out from the center of the screen represent the 3D axes of the world, and the intersection point of these lines is the roosting point which the birds focus around. When this window has focus it is possible to interact with the simulation environment with the following controls:

- 1. Mouse Can change the direction the camera is facing by either clicking an dragging an in the appropriate direction or by simply moving the mouse if (8) is switched on.
- 2. W key Moves the camera forward.
- 3. S key Moves the camera backward.
- 4. A key Moves the camera left.
- 5. **D** key Moves the camera right.
- 6. **E key** Moves the camera up.
- 7. C key Moves the camera down.
- 8. **M key** Switches manual mode on/off. If it is on then the camera's facing direction can be adjusted by simply moving the mouse, if it is off then the camera's facing direction can be changed by clicking and dragging with the mouse in the appropriate direction. Default is off.
- 9. H key Hides or reveals the simulation manager (Shown in Figure A.3).

Please note that a reminder of the controls is given by accessing Help -¿ Controls (Shown in Figure A.7 on page 66).

A.2.2 The Simulation Manager



Figure A.3: Main Simulation Controls

The main screen to be used when conducting a range of simulations in order to gather many data points is the simulation screen shown in Figure A.3. The controls of this screen are as follows:

- 1. By pressing File the options Save Settings, Load Settings and Exit appear. Save Settings allows the user to save the current simulation settings to the hard drive. The simulated birds are often difficult to control and as such getting them to act in a certain manner is often a tedious parameter hunt, once the ideal parameters have been located this option simply allows the user to save all the settings for future use. The Load Settings just allows the user to load previously saved settings ending in the extension .sset. Exit just allows the user to quit the application. **Note:** any unsaved results or settings will be permanently lost.
- 2. By pressing Help the user can access the Controls screen (Shown in Figure A.7 on Page 66) to be reminded of the simulation window controls and the About screen (Shown in Figure A.8 on Page 66) to learn more about the simulation program.
- 3. Shows the current simulation being conducted, once this number is equal to that input into (5), the simulations will stop.
- 4. Shows the current Cycle which simply refers to the amount of time the current simulation has been running. Once this number equals that of (6) the current simulation will stop (also meaning that all results regarding this simulation have been safely recorded) and either the next simulation will start with a new random flocking factor (If (3) has not yet reached (5)) or nothing will happen meaning the end of the simulation runs.
- 5. Refers to the number of simulations to be conducted.
- 6. Refers to how long each simulation should run.
- 7. This controls how often the flocking should be measured (according to (11)) and the results saved. For example if set to one the flock will be measured every cycle, if set to 100 the flock will be measured every 100 cycles.

- 8. Is used to set how many birds should be simulated. **Note:** This value must be divisible by 5 for optimization and threading reasons, also note that if this value is set too high then the simulation will run very slowly.
- 9. Controls the number of "Indicator" birds, i.e. the number of birds from which to conduct the measurements (Specified by (11)), the higher this number the more data produced per simulation. Setting this value too high might mean having an overload of data and slowing down the simulation.
- 10. Controls what is referred to as the analysis scan range, which simply controls how many other birds of the specified distance from an indicator bird should also be measured. For example if set to 0 then only the indicator bird will be measured, if set to 1000 then all birds which fall into an Indicator bird's perception radius of 1000 radius will also be taken into account.
- 11. Controls which type of data about the birds will be measured:
 - *Vel* Each bird's average velocity (taken by summing the speed of each bird in each direction and dividing by 3) will be recorded.
 - *Pos* Each bird's average position (taken by summing the position of each bird in each axes and dividing by 3) will be recorded.
 - *Vel Diff* Each bird's average velocity change from one save to the next (the length of which is specified by (7)) will be recorded.
 - *Pos Diff* Meaning that the distance a bird travels from one save to the next (the length of which is specified by (7)) will be recorded.
 - Home Diff Which records the distance each bird is from the roosting site.
 - $Mean \ Diff$ Which records the distance each bird is from mean position of all the other birds.
 - *Mean VelM* Each bird's average velocity magnitude change from one save to the next (the length of which is specified by (7)) will be recorded. In essence this records the difference between the distance a bird has travelled and the mean distance all the other birds have travelled.
- 12. Controls how often a bird should consider its current state relative to all other birds. Essentially this controls how often a bird should reconsider whether it is going to fly away from the roost or towards a roost. High values means birds will linger in their current states for longer periods of time. Low values mean means birds will change their behaviour often.
- 13. This controls the most important attribute of the simulations i.e. the Flocking Factor. A high value means birds will likely all move as one, a low value means birds will flock more randomly and anything in between will result in more interesting flocking behaviour. One could argue this controls the amount the social rules impact each bird. **Note:** Upon the start of each new simulation a new flocking factor will be randomly chosen, only the first simulation will use the user specified flocking factor.
- 14. Stops the current simulations and loses any unsaved progress.
- 15. Starts or restarts the current simulations (Resetting (3) and (4) to 1). Warning: when starting a new set of simulations any old simulation data in the directory "AnalysisData" will be deleted.

For more specific control with regards to the behaviour of the flock please refer to Sections A.2.4 and A.2.5.

A.2.3 Camera Controls

٠						
File He	lp		Sim Nu	mber: 0	Cycle: 0	
Con	nponents	Camera Controls:				
W	🔘 Welcome	Camera Speed:	0		100	
**	 Simulation 	Y Speed:				
	⊙ Camera	Z Speed:	0			:
\bigcirc	 Drives 	Camera Spawn Poin	t: Requires Simulation	Restart		
(\checkmark)	O Neighbours	Y Position:	-3,000	▼ ₹		
	Start	Z Position:	7,000	Ť		
	Stop	Modifying the above camera within the si	will affect the user co mulation	ontrolled		

Figure A.4: Camera Properties

Using the screen shown in Figure A.4 it is possible to adjust how the simulated world is viewed and how it is navigated by the user. The following camera properties can be adjusted:

- 1. Controls the speed at which the camera moves on the X-axes. I.e. it controls the speed of moving the camera left and right.
- 2. Controls the speed at which the camera moves on the Y-axes. I.e. it controls the speed of moving the camera up and down.
- 3. Controls the speed at which the camera moves on the Z-aces. I.e. it controls the speed of moving the camera backward and forward.
- 4. Controls where in the simulated world the camera spawns on the X-axes at the start of the simulation. Positive values means the camera spawns towards the right and negative values means the camera spawns towards the left.
- 5. Controls where in the simulated world the camera spawns on the Y-axes at the start of the simulation. Positive values means the camera spawns towards the bottom and negative values means the camera spawns towards the top.
- 6. Controls where in the simulated world the camera spawns on the Z-axes at the start of the simulation. Positive values means the camera spawns with the front of the flock in view and negative values means the camera with the back of the flock in view.

A.2.4 Drive Forces

٠						
File He	lp			Sim Number: 0	Cycle: 0	
Con	nponents		Drive Factors:			
W	🔘 Welcome	Cohesion:	0)	.0	-1
x.x	O Simulation	Seperation:	()		-2
	🔘 Camera	Alignment:	(-	-3
$\mathbf{\Theta}$	• Drives	Random: Home:	()		-4
$(\not\prec)$	O Neighbours	Leader:)	<u> </u>	-6
	Start	Global Cohesi	on: 0_ (.0	-7
	Stop	Each parameter associated with	above is a sacling	factor for the rule	e it is	

Figure A.5: Flock Drives

It is through the screen shown Figure A.5 that a detailed control over the flock can be achieved. The following properties of the flock can be modified:

- 1. This controls the local cohesive force of the sub flocks which form. A high value means that any sub flocks which form will be closely bundled together into a ball, while low values mean that birds will be more inclined to not stick close to their locally perceived neighbourhood.
- 2. This controls how strongly birds will fly away from each other when they get too close. A high value means that birds are highly repelled when they get too close to each other and a low values means that birds can move quite close to each other an even freely overlap. Note: This simulation model does not incorporate collision detection, thus overlapping of birds is possible.
- 3. This controls to what degree birds which to move at the same rate and in the same direction. A high value will mean that more and more birds will fly at the same velocity and a low value means that birds will be less inclined to match velocities and therefore will act more independently.
- 4. This controls the strength of the random force birds experience. The random force is used to model imperfect flight. A high random force means birds will fly more chaotically and a low value will cause birds to fly in a more controlled path, thus the flocking will be more deterministic.
- 5. This controls the drive which birds have to head towards the roosting site. A high value means birds will be more inclined to stay around the center of the simulation and a low values mean that birds will have little or no desire to remain around the center.
- 6. This controls how strongly the birds will follow one specific bird. A low value means that the flock has no leader while a high value will result in all the birds chasing after an automatically chosen leader.

7. This is similar to (1) except it refers to entire flock rather than the emergent sub flocks

It is important to note that different combinations of settings will result in various flocking. For example if setting (1) is very high will it only have a large effect if all the other drives are very low. Each drive influences each bird relative to all other drives.

A.2.5 Neighbourhood Definitions



Figure A.6: Neighbourhood Definitions

With the screen shown in Figure A.6 it is possible to control how birds' perceive each other which controls how the neighbourhood of each bird is defined. Too some extent modifying the properties as stated below controls the degree to which sub flocks form as well as the extent to which social rules affect the whole flock.

- 1. The field of view controls the the viewing radius. If the value is 360° that means that the bird is fully aware of it's surroundings, if on the other hand the value is 90° for example then the bird can only see exactly what is in front of it. 0° means that the bird is visually blind. Note: A bird fly away/towards the center based on the number of birds it can see. If a bird can see more than it desires it will fly away from the roosting area on the basis that the roosting area is "overcrowded" if on the other hand a bird sees to few birds then it will fly towards the roosting area signifying it is "scared" to be alone and far away from home.
- 2. The interpolation factor controls how rapidly a bird expands and collapses it perception range. A bird will expand it's perception range to approximately the maximum specified in (4) if it perceives too few birds, if on the other hand it perceives too many bird then it will collapse it perceptive range. The goal is to keep the percentage of birds specified in (5) in view. The interpolation factor controls this rate of expansion and contraction.
- 3. This value specifies the minimum range a bird's perception is allowed to drop too.
- 4. This value specifies approximately the maximum range a bird's perception is allowed to expand too.

5. This specifies the percentage of birds each bird tries to keep in view, thus directly influences the expansion and contraction of a bird's perceptive range.

The above settings provides very detailed control of how each bird perceives all the other birds. Using these properties it is possible to accurately control the sub flocks which form.

A.2.6 Other Windows

•				
Controls				
Moves the camera forward W E Moves the camera up				
Turns the camera to the left 🛛 S 🗩 Turns the camera to the right				
Moves the camera back Moves the camera down				
H Hides or reveals the managment window				
Sets Manuel Control - When on the camera is controled by dragging the mouse when off the camera is controled by moving the mouse				

Figure A.7: Controls Window



Figure A.8: About Window

A.3 Usage Notes and Conducting an Analysis

When using the application the following must be kept in mind.

Whenever a specific simulation is completed two text files in the directory AnalysisData are created as a result; called "Results_Sim_#" and another "Settings_Sim_#", where # refers to the simulation number. The former contains the results of the simulation while the later contains the settings in a readable manner used to achieve those results.

It is very important to note that whenever a new set of simulations is run, all the old text files in the directory "AnalysisData" will be overwritten, as such please ensure to move the data to a safe location upon completing the simulations.

Before an analysis can be conducted please ensure that the Granger Causal Connectivity Analysis (GCCA) toolbox developed by Anel Seth (Seth, A.K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *Journal of Neuroscience Methods*.**186**:262-273) is installed in Matlab.

Once the toobox is installed it is then possible to conduct an analysis by first moving all the data into a directory such as "SimRun" and then opening Matlab, navigating to the directory "MATLAB Analysis Code" and then running from the Matlab console: *Analyse('etc/SimRun', Max, nLags, WindowSize, Overlap);* where etc is the path to the directory, nLags is the model order to be used for analysis, Max is the toal number of simulations conducted, WindowSize is the size of the analysis windows to be used and Overlap is the offset to be used for the interleaving windows. The results of this will be 5 vectors being saved:

- FlockingFactor A vector where each element indicates the flocking factor which was used for that simulation run. For example FlockingFactor(1) = 2 means that for the 1st simulation the flocking factor was 2.
- AverageCD A vector where each element indicates the average causal density across multiple analysis windows for that simulation run.
- **MaxCD** A vector where each element indicates the maximum causal density across multiple analysis windows for that simulation run.
- AccumulatedCD A vector where each element indicates the total causal density across multiple analysis windows for that simulation run.
- **FinalCD** A vector where each element indicates the causal density across multiple analysis windows for that simulation run, which is computed by using the maximum Granger causality values and the Causal Relational matrix at the end of the analysis. The Granger causality and Causal Relational values themselves are the maximums taken across each analyses window.

The results of the final analysis can then be seen graphically by plotting the **FlockingFactor** against one of the other vectors by using in the Matlab console createfigure(FlockingFactor, X) where X is one of the other vectors.

A.4 Closing Notes

Please note that this application is not perfect and it may contain some errors. This application was developed and submitted in partial fulfilment of the requirements for the MSc Degree in Computing Science/Artificial Intelligence of Imperial College London.

It is meant as a research aid. If this application is ever used under academic or professional circumstances please ensure to give credit to the developer as well as reference Imperial College London. Also note that the analysis in Matlab would not be possible without the excellent Matlab Ganger Causality toolbox developed by Anil Seth

Seth, A.K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *Journal of Neuroscience Methods*.**186**:262-273

Please ensure you cite the above if you make use of the analysis code.

Thank you for using this application and for being patient with any errors which may have occurred, please feel free to contact the developer should you have any questions.
A.5 Technical Details

The Simulation Environment

Version 1.6 Developed In: Java SE 6 and Processing Developed By: Filipe Peliz Pinto Teixeira Email: fppintoteixeira@gmail.com

Matlab Analysis Code

Developed By: Filipe Peliz Pinto Teixeira Using: Granger Causal Connectivity Analysis (GCCA) Toolbox by Anil Seth Seth, A.K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. Journal of Neuroscience Methods. **186**:262-273

B The 3 Primary Flocking Patterns

The following section presents a series of screen-shots at different time steps and under different flocking factors used to show the 3 primary flocking patterns which are of interest, namely:

- **Random Flocking** Where the flocking factor is 0 and the only social aspect driving the birds is the separative force.
- **Natural Flocking** Where the flocking factor is set to 1 and the birds move slightly more together thus creating a more visually appealing and interesting simulation.
- Stagnated Flocking Where the flocking factor is 10 an all birds move together as one.

Other than the flocking factors specified above, the following settings were used to achieve the behaviour demonstrated:

Flock Size:	1000
Number of Indicators:	1
Flock Update Rate:	1.0
Separation Drive:	2.0
Cohesion Drive:	2.0
Global Cohesion Drive:	0.0
Alignment Drive:	1.0
Random Drive:	1.0
Home Drive:	5.0
Leader Drive:	0.0
Field of View:	90.0
Interpolation Factor:	0.1
Minimum Perception Range:	0.0
Maximum Perception Range:	1000.0
Percentage of Birds Kept In View:	10%

Table B.1: Settings used to the achieve flocking presented

B.1**Random Flocking Pattern**



(a) ms = 5000

(b) ms = 5500

(c) ms = 6000



(d) ms = 6500

(e) ms = 7000

(f) ms = 7500



(g) ms = 8000

(j) ms = 9500(k) ms = 10000(l) ms = 10500

Figure B.1: Random Flocking: Flocking Factor = 0

B.2 Natural Flocking Pattern





(d) ms = 6500

(e) ms = 7000

(f) ms = 7500



(g) ms = 8000

- (h) ms = 8500
- (i) ms = 9000



Figure B.2: Natural Flocking: Flocking Factor = 1

B.3 Stagnated Flocking Pattern





(d) ms = 6500

(e) ms = 7000

(f) ms = 7500





Figure B.3: Stagnated Flocking: Flocking Factor = 10

C Primary Code Snippet

The following is a small snippet of code which shows the calculation of the social and roosting forces defined in Section 3.1

```
* Class Name:
                  Rule Applier
                  Main class which computes social
 * Purpose :
                   and other forces which a bird receives
* Developed By: Filipe Peliz Pinto Teixeira
                  September 2011
* Date:
*/
package sim;
import java.util.Random;
import processing.core.PVector;
public class RuleApplier implements Runnable {
    . . .
    private PVector calculateSocialForces(int index)
        PVector result = new PVector (0, 0, 0);
        {\tt PVector seperation} = {\tt new PVector} \left( 0 \,, \ 0 \,, \ 0 \right);
        PVector neighbourhoodPos = new PVector (0, 0, 0);
        PVector cohesion = new PVector (0, 0, 0);
        \texttt{PVector globalPos} = \texttt{new} \texttt{PVector}(0, 0, 0);
        PVector globalCohesion = new PVector(0,
                                                     0, 0);
        PVector alignment = new PVector (0, 0, 0);
         float distance = 0;
         float cohesionNeighbourhood = 0:
         float alignmentNeighbourhood = 0;
         //Changes the birds perception radius such that the bird tries to
         //keep a certain amount of birds in its neighbourhood
         //(think of it as allowing the eyes to adjust //to view things at different distances)
        parent.birds[index].perceptionRadius =
             (float) ((1 - parent.getInterportationFactor()) *
             parent.birds[index].perceptionRadius +
             parent.getInterporlationFactor() *
                  (parent.getMaxPerceptionRadius() -
                     (parent.getMaxPerceptionRadius() *
                           (parent.birds[index].birdsInNeighbourhood
                           / parent.getBirdsInNeighbourhood())) ));
         if(parent.birds[index].perceptionRadius < parent.getMinPerceptionRadius())
             \verb|parent.birds[index].perceptionRadius = \verb|parent.getMinPerceptionRadius();|
         parent.birds[index].birdsInNeighbourhood = 0;
         //Iterating through the perceived flock
         for (int j = 0; j < parent.birds.length; j ++)
         {
             if(index != j)
             ł
                  distance = PVector.dist(parent.birds[j].pos,
                              parent.birds[index].pos);
                  if (distance >= 0 \&\&
                     distance <= parent.birds[index].perceptionRadius)
                  {
                      parent.birds[index].birdsInNeighbourhood++;
                      //Separation Rule: First
                      //checks if safety radius is breached
                      if (distance > 0 && distance <= parent.getBirdBody())
                      {
                           {\tt PVector temp} \ = \ {\tt PVector.sub} \left( {\tt parent.birds} \left[ {\tt index} \right] . {\tt pos} \ , \right.
                                                         parent.birds[j].pos);
                           temp.normalize();
                           temp.div(distance);
                           seperation.add(temp);
                      }
                  }
```

```
//Alignment Rule
        float adjust = 1;
        if (parent.getFlocking() >= 1)
            adjust = parent.getFlocking();
        if (distance >= 0 \&\&
           parent.birds[index].rot, parent.birds[j].pos))
        {
            alignmentNeighbourhood++;
            alignment.add(parent.birds[j].vel);
        }
        //Local Cohesion rule
        if (distance >= 0 \&\&
           distance <= 2* parent.birds[index].perceptionRadius)
        {
            if (inView(parent.birds[index].pos,
                      parent.birds[index].rot, parent.birds[j].pos))
            {
                cohesionNeighbourhood++;
                neighbourhoodPos.add(parent.birds[j].pos);
            }
        }
   ł
   //Global Cohesion Rule
   globalPos.add(parent.birds[j].pos);
}
globalPos.div(parent.getBirdCount()-1);
globalCohesion = PVector.sub(globalPos, parent.birds[index].pos);
globalCohesion.limit(0.1f);
if(cohesionNeighbourhood > 0)
{
    neighbourhoodPos.div(cohesionNeighbourhood);
    cohesion = PVector.sub(neighbourhoodPos, parent.birds[index].pos);
    cohesion.limit(0.1f);
}
if(alignmentNeighbourhood > 0)
{
    \verb"alignment.div(alignmentNeighbourhood);"
   \texttt{alignment.limit}(0.1 \texttt{f});
}
//Scaling drives
//Note parent.getGlobalCohesionDrive() =
//Flocking Factor / 5.0
globalCohesion.mult(parent.getGlobalCohesionDrive());
\verb|cohesion.mult(parent.getCohesionDrive())|;|
seperation.mult(parent.getSeperationDrive());
alignment.mult(parent.getAlignmentDrive());
//Strengthened Force to Fly Home
currentCheckIt ++;
PVector roost = PVector.sub(parent.getHome(), parent.birds[index].pos);
roost.limit(0.1f);
//Checking if the strengthened force should be applied
if (parent.getBirdsInNeighbourhood() != 0 &&
   currentCheckIt \% parent.getCheckIt() == 0 & &
   (alignmentNeighbourhood/parent.getBirdsInNeighbourhood())
      < parent.getReturnHomeProbability())
{
   currentCheckIt = 0;
   roost.mult(parent.getHomeDrive());
    result.add(roost);
}
globalCohesion.mult(parent.getFlocking());
roost.mult((float) (parent.getFlocking()/10.0));
{\tt result.add}({\tt roost});
alignment.mult(parent.getFlocking());
result.add(globalCohesion);
```

}

```
result.add(cohesion);
    result.add(seperation);
    {\tt result.add(alignment)};
     //Leader Rule
     if (index != 0 && parent.getLeaderDrive() != 0)
     {
         {\tt PVector \ leader \ = \ PVector \ . \ sub} \left( \, {\tt parent \ . \ birds} \left[ \ 0 \ \right] . \ {\tt pos} \ , \right.
                                            parent.birds[index].pos);
         leader.limit(0.1f);
         leader.mult(parent.getLeaderDrive());
         result.add(leader);
    }
    return result;
}
// Determines if one object is in the field of view of another.
//\,{\rm This} is based off of Nic Foster's XNA
//3d tutorial code http://nfostergames.com/Lessons/3DFieldOfView.htm
private boolean inView(PVector posFirst, PVector facingFirst, PVector posSecond)
{
    PVector test = PVector.sub(posSecond, posFirst);
    test.normalize();
    PVector facing = new PVector(facingFirst.x, facingFirst.y, facingFirst.z);
    facing.normalize();
     float dot = PVector.dot(facing , test);
      i\,f \ (\,\texttt{dot}\,>=\,\texttt{parent}\,.\,\texttt{getFOV}\,(\,) \ ) \\
         return true;
    return false;
}
```