Strategy dynamics in groups of prey

Joseph William Barrass Butterick

Doctor of Philosophy

University of York

Department of Mathematics

August 12, 2021

Abstract

Group living is a widespread and ubiquitous phenomenon in the animal kingdom. Animals in groups are subjected to numerous costs and benefits. Of the benefits, increased protection from predators is the most theoretically and experimentally researched. Three "principal" effects explain the anti-predation benefits of grouping: collective vigilance, dilution of risk, and the confusion effect. In this thesis, new theoretical models are developed to study the behaviours that emerge from groups of prey. Using a mathematical approach, this research analyses how the principal anti-predation effects interact to influence the behaviours of grouping animals.

Through an analytical approach, the first part of this thesis examines the role of anti-predation effects on individuals with discrete behavioural choices. The outcomes of group interactions are classified by exact analytical conditions. Results are provided which show how factors relating to predation risk, such as group size, affect anti-predator behaviours.

Another component of this research uses field data to assess the theoretical results that are presented. The relative influence of each anti-predation effect, and how the anti-predation effect interact to influence individual behaviours within groups is assessed. Consideration of how predation risk and other factors of grouping affect behaviours is also analysed.

The final section of research considers highly flexible and continuous anti-predator behaviours. This shift in theory is relevant in groups of prey, for example in models of vigilance. Intuitive results are presented, and conditions are provided which determine qualitatively distinct behavioural dynamics when groups are characterised by continuous behaviours. Equivalences to the discrete behaviour analogue case are shown. It is also shown that individuals may choose neither of the discrete behavioural options. Conditions are derived which specify the occurrence of multiple distinct anti-predator behaviours emerging from a group initially composed of one behaviour, and when the reserve process occurs.

Contents

Acknowledgements

I would first and foremost like to express my gratitude to my thesis advisors, Dr. Jamie Wood and Dr. Daniel Franks. Their combined knowledge have been the determinants of the progression of this research. Dan's professionalism, experience and constant availability to answer my (many) questions has been an anchor for this research. Jamie's insight, intuition and enthusiasm cannot be matched. Both supervisors have supported me more than I could have expected. It has been an honour to have worked with them.

I am also especially grateful to Dr Will Cresswell, who early on in this research, provided much advice and guidance. Without Will's encouragement and generosity a large section of this research would not have been possible.

I am sincerely grateful to the Engineering and Physical Sciences Research Council (EPSRC) and the University of York for providing the facilities for the undertaking of this research.

Lastly, I would like to thank my friends and family for offering whatever support they could have. Over the writing up period, my cat Ollie has also been a reliable friend, often sat on the laptop, checking my spelling.

Declaration of Authorship

I, Joseph William Barrass Butterick, declare that this thesis titled, "Strategy dynamics in groups of prey" and the work presented in it are my own.
Signed: Joseph William Barrass Butterick

Date: 12/08/2021

Chapter 1

Introduction

The risk of predation is a major force in shaping the behaviours and physical characteristics of animals. Predation presents a challenge for many species because both the non-lethal effects (such as avoiding the predator) [1], as well as the more obvious lethal effects (such as being killed) [2], reduce an animal's reproductive success. To increase the probability of unsuccessful predation, prey have developed a variety of anti-predator characteristics, from physiological and morphological defences, to behaviours. For examples of physiological changes, in the fish guppies, populations exposed to predators that predominantly feed on large individuals have developed reproductive maturity at a smaller size (and earlier age) compared with populations exposed to weaker predation on small individuals [3]. Anti-predator behaviours ranging from forming groups [4], to stunning dynamics in bird flocks [5] and fish shoals [6], to sentinel behaviour [7, 8] or raising alarm calls [9], are all believed to be adaptations to reduce predation [4, 10, 11].

Adaptations generally change over a very long period of time, i.e., over generations. Occasionally individuals are born with altered characteristics, caused by the genetic changes (or mutations) in the reproduction process. The changes cause differences in the mutant individuals ability to survive and reproduce. If the mutant individuals have a survival and reproductive advantage then they reproduce at higher rates and pass the altered characteristics to their offspring. If the mutant individuals are at a disadvantage then they have a lower chance of survival. The ability of an individual to survive and reproduce is referred to as that individual's "fitness". The theory of how mutation and selection explain the changing characteristics in populations is now considered classic, and has had success in many applications, for example the work of Fisher [12], Wright [13] and Haldane [14] showed that selection was consistent with Mendel's theory of genetic inheritance.

Some anti-predator characteristics, and in particular behaviours, can change on a much faster timescale. For example, prey exhibit adaptive responses when in the presence of a predator, such as changing habitat use, altering temporal activity and responding to predator cues [2, 15–18]. This thesis is interested in the behaviours which prey adopt to reduce predation risk. Behaviours (and thus anti-predation behaviours) are defined through the individual's phenotype, i.e., the traits whether behavioural or morphological, that distinguish an individual and that are to some extent determined by genotype. Although phenotypic traits are the result of complex interactions between an individual's inherited genes and the environment that they experience, even when the effect genotype has on phenotype is complex or unknown (which is nearly all of the time [19, 20]), one can focus on the implication's of phenotype on fitness [21]. Thus, to explain the certain behavioural responses of prey individuals subject to predation, one can study how anti-predation behaviours

and other phenotypic traits of the prey interact and give rise to some measure of fitness, which is dependent on predation risk.

Behaviour is widely considered to be phenotypically plastic (able to alter across environmental conditions) [22], and prey responses can vary according to the type of predator [23], the responses of other prey [24, 25] and the environmental surroundings [26]. Moreover, differences in prey behaviours can also reflect connections to other phenotypic traits relating to vulnerability. For example morphological defences and aggressiveness can covary within individuals [27], and underlying metabolic conditions can prevent individuals from making costly defences [28, 29]. In a broad sense, this thesis attempts to bridge complex anti-predator behaviours and how they interact with other aspects of the prey's phenotype, to explain the dynamics of prey behaviour.

It is believed that anti-predator behaviours combine with other phenotypic traits to maximise predator evasion [18, 30–32]. Yet the ways in which multiple behavioural, morphological, life history traits – or a combination of the above – act in synergy, compensation, co-dependence or co-specialisation is largely under-investigated and varies on a taxonomic scale. What is certain is that behaviours such as responding appropriately to predator cues are vital in contributing to an evasion mechanism [17, 27], and thereby an individual's fitness.

The use of mathematical models allows the study of phenotypic change. Popular modelling frameworks include evolutionary game theory and adaptive dynamics. The appropriateness of each framework largely depends on the specific problem. In this thesis, evolutionary game theory [33–35] is predominantly used along with adaptive dynamics. In the following sections of this introduction, first the basic concepts of static evolutionary game theory will be introduced, and then the shift to dynamic evolutionary game theory will be outlined. A brief account of the relevance of stochastic game theory follows, and finally, the framework of adaptive dynamics will be given.

1.1 Deterministic Evolutionary Game Theory

Game theory was formalised by von Neumann & Morgenstern [36], and developed by Nash [37] amongst others as a framework to analyse decisions in human behaviour. Any situation in which at least two agents interact through anticipating their opponents actions, either consciously or implicitly is known as a "game". The agents are referred to as "players" whose available set of actions is called their "strategy". A strategy played in the game produces a payoff, i.e., how successful the strategy is. Mathematically, game theory analyses how the payoff of one strategy depends on the other strategies played. Nash [37] showed that assuming only rational agents interact, there exists a "Nash equilibrium" for each game in which no agent can do better by changing their strategy. Hamilton [38] and Maynard Smith [39] applied the ideas of game theory to explain the ecological problems of sex-ratio and animal conflicts. In evolutionary game theory, the idea of rationality is replaced with fitness: payoffs are equated to fitness, and strategies to phenotypes. It is useful to think of biological interactions in terms of games. One can then associate the actions that a player takes with given probabilities, i.e their so-called strategies, with different behaviours, and the payoffs that one receives with fitness. If the fitness that one receives is constant regardless of the strategies of others, fitness is frequency-independent, where as if fitness depends on the actions of others it is frequency-dependent. In order to establish an evolutionary

argument, strategies have to be able to spread in the population. The way in which strategies spread depends on the biological problem: for example the more successful strategies may spread through imitation [40] or reproduction [35].

In this section, cases of games in a non-dynamic setting are first introduced. Static evolutionary arguments are defined with regards to the ESS concept. Following this, dynamic game theory is introduced in a deterministic form.

1.1.1 Two player games

From an economic perspective it is assumed that players have some pre-defined capacities that compose their "rationality". A focal player's actions are then based on their expectations of the actions of other players in the game. At the precise same time, the action's of other players are based on their expectations of the focal player's actions. Because each player's actions are based on expectations, it is quite possible for the outcome of a game to have not been intended by any agent. If every player in the game has perfect knowledge of the actions taken in the game up to that point, then agents have "perfect information". In games with perfect information and all players being rational, each player knows which actions lead to which outcomes and can therefore discriminate from alternative actions to achieve their preferred outcome. In this case players can rank-order outcomes of actions, know which sequence of events lead to which outcomes, and choose actions from alternatives (their strategy space) according to their rank-order of outcomes. Games fall into the categories of either "normal form" or "extensive form." The extensive form game is more general than the normal form and allows for sequential rather than simultaneous decision making. For the scope of this thesis an introduction to normal form games is sufficient.

Normal form games consist of n players, a set of all strategies available to each player, and a payoff function for each player [33]. Technically the payoff function for a player is a mapping from the cross product of all players' strategy spaces to that player's payoff set. The payoff function takes the so-called strategy profile, i.e., the n-tuple of strategies chosen by each player, and produces a real number (payoff). Often games are assumed to have only two players. In this case a representation by a payoff matrix A with elements describing the payoffs to each player suffices to encompass all potential outcomes.

Let there be M fixed distinct behaviours are referred to as "pure strategies". The strategy space is given by $S = \{S_1, S_2, ..., S_M\}$. Players are allowed to either choose a pure strategy $S_i \in S$ or a mixed strategy specified by a vector of probabilities $\mathbf{p} = (p_1, p_2, ..., p_M)$ where each p_i gives the probability of using strategy S_i. Generally, two player games can be simply represented by a payoff matrix in which a "row" player whose strategies are $\{R_1, R_2, ..., R_M\}$ and a "column" player with strategies $\{C_1, C_2, ..., C_M\}$ have respective payoff functions $\alpha_1(R_i, C_i)$ and $\alpha_2(R_i, C_i)$:

$$
\begin{array}{c}
 \\
R_1 \\
R_2 \\
\vdots \\
R_M
\end{array}
\begin{array}{c}
\begin{array}{cccc}
C_1 & C_2 & \dots & C_M
\end{array} \\
\begin{pmatrix}
\alpha_1(R_1,C_1), \alpha_2(R_1,C_1) & \alpha_1(R_1,C_2), \alpha_2(R_1,C_2) & \dots & \alpha_1(R_1,C_M), \alpha_2(R_1,C_M) \\
\alpha_1(R_2,C_1), \alpha_2(R_2,C_1) & \alpha_1(R_2,C_2), \alpha_2(R_2,C_2) & \dots & \alpha_1(R_2,C_M), \alpha_2(R_2,C_M) \\
\vdots & \vdots & \ddots & \vdots \\
\alpha_1(R_M,C_1), \alpha_2(R_M,C_1) & \alpha_1(R_M,C_2), \alpha_2(R_M,C_2) & \dots & \alpha_1(R_M,C_M), \alpha_2(R_M,C_M)
\end{pmatrix}
\end{array}
$$

Here, when the row player chooses strategy R_i and the column player chooses C_j, the row player gets a payoff $\alpha_1(R_i, C_j)$ the column player receives $\alpha_2(R_i, C_j)$. An obvious condition on the probability vector $\mathbf{p}$ is that $p_1 + p_2 + ... + p_M = 1$. Therefore each mixed strategy belongs to the M (number of strategies) dimensional simplex

$$\Delta^M = \Big\{\mathbf{p} = (p_1, p_2, ..., p_M) \in \mathbb{R}^M : p_i \geq 0 \text{ and } \sum_{i=1}^{M} p_i = 1\Big\}. \tag{1.1}$$

The vertices of the simplex are the unit vectors whose i^{th} component is equal to one and represent each of the $S_i \in S$ pure strategies. The interior of the simplex, denoted by $\text{int}\Delta^M$ contains only (and all of the) mixed strategies $\mathbf{p}$ and the boundary $\text{bd}\Delta^M$ (sometimes $\partial\Delta^M$) consists of all the $\mathbf{p}$ with support $\text{supp}(\mathbf{p}) = \{i : 0 \leq i \leq M \text{ and } p_i > 0\} \subset \{1, 2, ..., M\}$. We interpret the support of a mixed strategy $\mathbf{p}$ as the set of pure strategies with positive probabilities p_i.

Let $p_R(R_i)$ be the probability that the row player chooses the strategy R_i, and $p_C(C_i)$ be the probability that the column player chooses strategy C_i. The p_R and p_C describe probability distributions (mixed strategies), for the row and column players respectively. The expected payoff for the row player with strategy R_j is $p_C(C_1)\alpha_1(R_j, C_1) + p_C(C_2)\alpha_1(R_j, C_2) + ... + p_C(C_M)\alpha_1(R_j, C_M)$, i.e., the sum of the conditional payoff probabilities that the column player plays a particular strategy, given the row player's strategy choice.

In symmetric games both of the row and column players have the same strategic choices, $S = \{S_1, S_2, ..., S_n\}$. The payoff obtained by using any of the available strategies is then irrespective of the player that uses it. As such, the game can be characterised by an $n \times n$ square matrix, A, whose elements in the i^{th} row and the j^{th} column (here denoted a_{ij}) are the singular entries: $a_{ij} = \alpha_1(S_i, S_j)$. A player choosing the pure strategy S_i has the expected payoff $\pi_{S_i} = (A\mathbf{q})_i$ when its opponent chooses $\mathbf{q}$, and a player choosing a mixed strategy $\mathbf{p}$ obtains the expected payoff $\mathbf{p}A\mathbf{q}^T$ when its opponent plays a mixed strategy $\mathbf{q}$. In the following it should be clear that the transpose vector is assumed if not notationally explicit.

An important concept in game theory is the Nash equilibrium. The "best response strategy" is a strategy such that when used against a given strategy offers the "best" or highest possible payoff. If the best response strategy is unique it is called the "strict" best response. A "dominant strategy" is a strategy which is a strict best response to every other strategy. A Nash equilibrium is a set of strategies consisting of a strategy for each player for which each player's strategy is the best response to the other players' strategy, i.e., if any focal player alters its strategy while the other players do not, then the focal player's payoff will either decrease or remain the same. Strategies can be pure or mixed. If the only result of a focal player changing strategy is a decrease in payoff, the Nash equilibrium is "strict".

Mathematically, a strategy $\mathbf{p}$ is a Nash equilibrium if the following holds: if player 1 chooses a strategy $\mathbf{p}$, then any other player with an alternate strategy $\mathbf{q}$ cannot achieve a higher payoff. The Nash equilibrium is written as

$$\mathbf{p}A\mathbf{p} \geq \mathbf{q}A\mathbf{p}, \quad \forall \mathbf{q} \in \Delta^M \tag{1.2}$$

and is said to be strict (and unique) if

$$\mathbf{p}A\mathbf{p} > \mathbf{q}A\mathbf{p}, \quad \forall \mathbf{q} \in \Delta^M \tag{1.3}$$

To illustrate the Nash equilibrium concept, consider a symmetric 2 player game (players 1 and 2) with two available strategies given by S_1 and S_2. The payoff matrix, A, can be written as

$$\begin{array}{c} \\ S_1 \\ S_2 \end{array} \begin{array}{c} \begin{array}{cc} S_1 & S_2 \end{array} \\ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \end{array} \tag{1.4}$$

The strategy S_1 is a pure Nash equilibrium if $a_{11} \geq a_{21}$ since neither player can improve their payoff by choosing strategy S_2, given that their opponent chooses S_1. By the same reasoning, S_2 is a pure Nash equilibrium if $a_{12} \leq a_{22}$. If $a_{11} < a_{21}$ and $a_{12} > a_{22}$ then there is a mixed Nash equilibrium where the players choose strategy S_1 with probability u and S_2 with probability $1-u$. To find the mixed equilibrium we use the expected payoffs $\pi_{S_1} = ua_{11} + (1-u)a_{12}$ and $\pi_{S_2} = ua_{21} + (1-u)a_{22}$. The probability $u \in (0,1)$ is found by finding the root of $\partial_q(q\pi_{S_1} + (1-q)\pi_{S_2})$, which is equivalent to solving $\pi_{S_1} = \pi_{S_2}$ and yields $u = (a_{22} - a_{12})/(a_{11} + a_{22} - a_{12} - a_{21})$.

The solutions to classical strategic games are based on assumptions of full information, full rationality and simultaneous decisions. However, in reality (and in biological interactions) players may act irrationally. In the next subsection, a brief overview of an important concept in biology is given. Similar to classic game theory, the concept of evolutionary stable strategies does not involve dynamics, but an important development is made in that a population of interacting strategies is considered rather than two players.

1.1.2 Evolutionary Stable Strategies

The research of Maynard Smith & Price [39] addressed the question of why animals of the same species escalate conflicts. Their principle assumptions have since created a new paradigm of evolutionary game theory. Evolutionary game theory shifted from an individual perspective, i.e., a focal player choosing some strategy, to a population perspective: a large (possibly infinite) number of individuals with a set of pre-defined strategies. Within the population strategies are hard-wired into each individual. Completely new strategies can arise through mutations. Individuals are randomly and repeatedly paired, and play a game using their hard-wired strategies.

The idea of an evolutionary stable strategy (ESS) is concerned with the success of invading mutant strategies. In the population setting, the issue addressed by Maynard Smith is as follows: given a monomorphic population in which all individuals play the same strategy, can an alternative strategy enter the population and do better? That is, can a mutant strategy invade? An ESS is theoretically explained by Maynard Smith [35] p.10 as a "behavioural phenotype... [which] if all members of a population adopt it, then no mutant strategy can invade the population." From a mathematical perspective, consider a symmetric M-strategy game between two players, let the pure strategies be given by $S_1, S_2, ..., S_M$, and let a_{ij} be the payoff for strategy S_i played against strategy S_j (and a_{ji} for S_j played against S_i). The strategy S_i is an ESS for $\forall k \neq i$ if $a_{ii} > a_{ki}$ holds or $a_{ii} = a_{ki}$ and $a_{ik} > a_{kk}$. Players are not restricted to play the pure strategy S_i all of the

time, and instead can play (probability distributions) mixed strategies $\mathbf{p}$. The (possibly mixed) strategy $\hat{\mathbf{p}}$ is an ESS if

$$\hat{\mathbf{p}}A\hat{\mathbf{p}} \geq \mathbf{q}A\hat{\mathbf{p}}, \quad \forall \mathbf{q} \in \Delta^M \tag{1.5}$$

with the condition

$$\text{for } \mathbf{q} \neq \hat{\mathbf{p}} \text{ and } \mathbf{q}A\hat{\mathbf{p}} = \hat{\mathbf{p}}A\hat{\mathbf{p}} \text{ then } \mathbf{q}A\mathbf{q} < \hat{\mathbf{p}}A\mathbf{q} \tag{1.6}$$

Condition (1.5) is simply the Nash equilibrium condition. When applied in solitary, condition (1.5) does not guarantee non-invasibility (since another strategy $\mathbf{q}$ can also be a best reply). When (1.6) is applied in conjunction with (1.5), the strategy $\hat{\mathbf{p}}$ cannot be invaded by any mutant strategy $\mathbf{q}$ because $\hat{\mathbf{p}}$ played against $\mathbf{q}$ fares better than $\mathbf{q}$ does against itself [33, 41]. Strict Nash equilibrium implies ESS and ESS implies Nash equilibrium. For two strategy two player matrix games with payoff matrix A as in (1.4) there is a mixed strategy ESS if $a_{11} < a_{21}$ and $a_{22} < a_{12}$.

The concepts of the Nash equilibria and the ESS are closely linked. All strict Nash equilibrium are ESS's and every ESS is a weak Nash equilibrium. To illustrate, consider the "Hawk-Dove" game [35] that served as motivation for the ESS concept. In the two strategy game, individuals compete over some resource V, and can choose to play hawk H, or dove D. Encounters involve pairwise contests. If a hawk meets a dove then the dove gets zero payoff and the hawk receives the full resource V. When the encounter is between two hawks the fight escalates, with probability one half each wins the resource V devalued by the costs of fighting C, both receive expected payoff $(V-C)/2$. When two doves meet, each wins the resource with a probability of one half, there are no costs, both expect a payoff $V/2$. This leads to the symmetric payoff matrix

$$\begin{array}{c} \\ H \\ D \end{array} \begin{array}{c} \begin{array}{cc} H & D \end{array} \\ \begin{pmatrix} \frac{1}{2}(V-C) & V \\ 0 & \frac{V}{2} \end{pmatrix} \end{array} \tag{1.7}$$

where for now either of $V > C$ or $V < C$ are possible. It is clear that playing dove can never be an ESS since $V/2 < V$: a population of doves can always be invaded by a hawk. H is an ESS if $(V-C)/2 > 0$ or $V > C$ since choosing the hawk strategy against another hawk is always better than playing the dove strategy against a hawk. Alternatively, if $V < C$ the game has one ESS which is a mixed strategy. To find the mixed ESS, let the strategy $\mathbf{p} = (p, 1-p)$ represent the probability p of playing H, and $1-p$ of playing D. By the Bishop Cannings Theorem [42], the payoff from playing either pure H or pure D against the mixed strategy, is equal to the payoff from playing the mixed strategy against mixed strategy: the ESS is found by setting $\pi_H = p(V-C)/2 + (1-p)V$ equal to $\pi_D = (1-p)V/2$. It is easy to check (but not done so for space constraints) that $\hat{\mathbf{p}} = (V/C, 1-V/C)$ is a mixed ESS and satisfies both of the ESS conditions (1.5) and (1.6). The best possible strategy, i.e., the ESS, is not to play H all of the time, but rather to play H with probability V/C, and to play D the rest of the time.

1.1.3 Population Games

So far, games have been assumed to be random pairwise encounters which subsequently specify the success of strategies. In many biological problems, the success of a strategy may depend on the population state (frequencies of other strategies) rather than the probability that one opponent plays a specific strategy. A classic example of this is the sex ratio game for which the success of a sex ratio depends not on the sex ratio of any particular opponent, but rather on the average sex ratio within the population. Maynard Smith [35] explained a correspondence between a monomorphic population in which all individuals play a (mixed) strategy of choosing strategy S_1 with a probability p, and S_2 with probability $1-p$, and a polymorphic population in which a frequency x adopt S_1 and a frequency $1-x$ adopt S_2. The equivalence is that p and x can be used interchangeably [33, 35].

Symmetric population games are a generalisation of matrix games with an infinitely large population with a finite set of pure strategies. The population state $\mathbf{x} = (x_1, x_2, ..., x_M)$, through each x_i, gives the population frequency of the strategy S_i. Since the frequencies of strategies sum to one, i.e., $x_1 + x_2 + ... + x_M = 1$, each population state lies in the Δ^M simplex which here is referred to as the set of "mixed population strategies" [33, 43]. For the population to evolve continuously, the outcomes of individual encounters are described by a payoff that is a function of the population state. If the payoff function is linear the population state, the game can be represented via a payoff matrix, whereas nonlinear payoff games are so-called "playing-the field" games [35]. Mathematically, population games are equivalent to the 2-player games of section 1.1.1, the only difference being that in population games, the vector $\mathbf{p}$ of probability distributions of each pure strategy is replaced by the state vector $\mathbf{x}$ of population frequencies of each pure strategy.

As such, there is a concept similar to the ESS definition, but applied to a population of strategies rather than an individual strategy. Namely, an "evolutionary stable state" is a population equilibrium in which the composition of pure strategies is resistant to perturbation, i.e., will be restored after disturbance. Define $f_i(\mathbf{x})$ as the expected payoff to an individual playing strategy S_i which has a frequency x_i in a population with state $\mathbf{x}$. Next, consider a potential dissident group of individuals of which a proportion y_i play strategy S_i and define $\sum_i y_i f_i(\mathbf{x}) = \mathbf{y} f(\mathbf{x})$ as the average payoff in that group (which in the case of a linear payoff function can be expressed as $\mathbf{y}A\mathbf{x}$). Whenever $\mathbf{y} \in \Delta^M$ and $\text{supp}(\mathbf{y})$ is contained in $\text{supp}(\mathbf{x})$ where $\mathbf{x}$ is an equilibrium population state, one has that $\mathbf{y} f(\mathbf{x}) = \mathbf{x} f(\mathbf{x})$ [42, 44]. To find the evolutionary stable state one has to consider the effects of a perturbation of the state $\mathbf{x}$. If a small subpopulation of density ϵ adopt a state $\mathbf{y}$, the state $\mathbf{x}$ is stable if the average fitness of $\mathbf{x}$ in the perturbed state is greater than the average fitness of the remaining population [44]. The evolutionary stable state, $\hat{\mathbf{x}}$ is defined by $\hat{\mathbf{x}} f(\bar{\mathbf{x}}) > \mathbf{y} f(\bar{\mathbf{x}})$ for all $\mathbf{y} \neq \hat{\mathbf{x}}$ where $\bar{\mathbf{x}} = (1-\epsilon)\hat{\mathbf{x}} + \epsilon\mathbf{y}$ for sufficiently small ϵ.

To see the mathematical connection between the mixed (individual) ESS in a monomorphic population, and the equivalent mixed evolutionary stable (population) state, consider the hawk dove game with $V < C$. Let x denote the frequency of hawks in the population, and let $1-x$ be the frequency of doves. Assuming that the strategies H and D are paired off at random, then the expected payoff to H (which depends on the frequencies of hawks and doves in the population) is: $\pi_H = x(V-C)/2 + (1-x)V$ and to D is $\pi_D = (1-x)V/2$. The condition $\pi_H = \pi_D$ results in

$\hat{x} = V/C$. We see that $\hat{x}$ and $1 - \hat{x}$ are the evolutionary stable frequencies of hawks and doves in the population since $V/2 < V$ and $(V - C)/2 < 0$.

1.1.4 Multiplayer Games

In this section we consider games in which each individual simultaneously interacts with more than one other individual at each instance [43, 45–48]. Such games are known as "multiplayer games". Despite multiplayer games receiving less attention compared two player games (partly due to the wide applicability and simplicity of two player games) within the field of evolutionary biology, in many ecological contexts multiplayer games are appropriate. For example when considering group-wide interactions [31, 49–51]. Here, we only introduce symmetric multiplayer games: for which the ordering of opponents does not matter.

Let the (pure) strategy set be fixed, finite and of size M. Consider interactions involving n players and let the vector $\mathbf{x} = (x_1, x_2, ..., x_M)$ represent the frequency distribution of individuals playing each strategy. Payoffs are nonlinear in the population state $\mathbf{x} \in \Delta^M$ and are found by averaging over weighted probabilities of rewards that arise through specific compositions of opponents. Thus, the payoff to an individual depends on its strategy, and the composition of strategies of the other $n - 1$ players. An $n \times M-$dimensional matrix can be used to represent payoffs. As an example, the simplest case of a three player, two strategy game is considered. Let the two pure strategies be S_1 and S_2 with frequencies x and $1 - x$ respectively so that the population state $\mathbf{x} = (x, 1 - x) \in \Delta^2$. The payoff matrix is

$$\begin{array}{c} \\ S_1 \\ S_2 \end{array} \begin{array}{c} \begin{array}{ccc} S_1S_1 & S_1S_2 & S_2S_2 \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \end{array}. \tag{1.8}$$

The matrix (1.8) describes the respective payoffs for an individual playing strategies S_1 or S_2 as dependent on the composition of the strategies played by its (two) opponents. If the row player chooses S_1 then either; both opponents may choose S_1 (column one), one opponent plays S_1 while the other plays S_2 (column two), or both opponents play S_2 (column three). Hence, because each opponent player chooses strategy S_1 with probability x and S_2 with probability $1-x$, the expected payoffs to each strategy are now polynomials:

$$\begin{aligned} \pi_{S_1} =& x^2 a_{11} + x(1-x)a_{12} + (1-x)^2 a_{13} \\ \pi_{S_2} =& x^2 a_{21} + x(1-x)a_{22} + (1-x)^2 a_{23} \end{aligned} \tag{1.9a}$$

The above payoffs for two strategy games can be extended to an arbitrary number of players. Consider a population in which there are two available strategies, S_1 and S_2, and with state $\mathbf{x} = (x, 1 - x)$ respectively representing the frequencies of individuals playing each strategy. As the population is assumed to be infinite and well mixed (every individual can interact with any other individual), randomly formed groups have compositions of strategies defined by the binomial distribution. In a group of size n, any focal player faces $n - 1$ opponents, each of which chooses strategy S_1 with probability x and S_2 with probability $1 - x$. Let $\pi_{S_i}(\mathbf{x}; S_1^l, S_2^{n-1-l})$ be the payoff

for a focal individual playing S_1 in the population defined by $\mathbf{x}$, when it finds itself in the group with l opponents who play strategy S_1 and $n-1-l$ opponents who choose strategy S_2. The expected payoff from choosing S_1 depends on the probabilities of facing certain strategic compositions of opponents, i.e., is summed over all possible compositions of opponent strategies in the group

$$E[\pi_{S_i}(\mathbf{x}; S_1, S_2)] = \sum_{l=0}^{n-1} \binom{n-1}{l} x^l (1-x)^{n-1-l} \pi_{S_i}(\mathbf{x}; S_1^l, S_2^{n-1-l}). \tag{1.10}$$

In this case, a polymorphic population in which a frequency $0 < x < 1$ play S_1 and $0 < 1 - x < 1$ play S_2 is equivalent to a monomorphic population in which each individual plays strategy S_1 with probability x and S_2 with probability $1 - x$.

Palm [45] provided the general groundwork to analyse the concepts of a Nash equilibrium and an ESS for multiplayer games. The ESS condition becomes a significantly more complicated for multiplayer games, see Bukowski & Miekisz [52] or Broom & Rychtář [53] for further details.

A motivating example of a multiplayer game is the n-player public goods game [54, 55]. In the game each player is given some initial endowment $\rho > 0$ and must then choose some fraction γ of this endowment to contribute to a public good. All of the contributions are multiplied by a factor $r > 1$ and distributed equitably between the group. Each player payoff is $\rho - \gamma$ nuanced by their share of the public good. Consider that the focal player contributes γ_1 and the other players contribute $\gamma_2, \gamma_3, ..., \gamma_n$ then the payoff to the focal playing γ_1 with opponents concurrently playing $\gamma_2, ..., \gamma_n$ is

$$\pi_{\gamma_1}(\gamma_2, \gamma_3, ..., \gamma_n) = \rho - \gamma_1 + \frac{r}{n} \sum_{i=1}^{n} \gamma_i \tag{1.11}$$

So long as $r < n$ each player is returned with only a fraction of their contribution. There is only one Nash equilibrium in the game, and that is to contribute nothing. We see this because $\pi_0(\gamma_2, \gamma_3, ..., \gamma_n) > \pi_{\gamma_1}(\gamma_2, \gamma_3, ..., \gamma_n)$ for all $\gamma_1 > 0$. The Nash equilibrium is however at the detriment of the group as each player receives their maximum payoff if all contribute ρ. The public goods game has become the standard game for investigating social dilemmas and the evolution of cooperation (here defined as a behaviour that is costly to the actor but benefits all other group members) when defecting is attractive [55, 56]. Similar games considering the evolution of cooperation under multiple and simultaneously interacting group members are the n-player prisoners dilemma [57], volunteers dilemma [58, 59] and snowdrift game [60, 61].

Multiplayer games provide a rich tool to study biological interactions in which the strategy chosen by each member of a group has a direct consequence for its own fitness and for the fitness of other group members. Of interest in this thesis, are model's such as Pulliam's [49], as to be introduced in section 1.5.1, which study the evolutionary stable strategies of anti-predator behaviours in groups of prey. Numerous other multiplayer games have researched the strategic compositions of animal groups [31, 32, 46, 51, 62, 63].

1.1.5 Replicator Dynamics

The research of Maynard Smith [39] can indicate whether a strategic composition is prone to mutant strategies invading or not. A limitation of the ESS concept, however, is that it assumes an initial

state where all of the members of the population play the same strategy. That is, the Maynard Smith understanding of stability does not consider how the population state has been reached.

In contrast to the concept of evolutionary stable strategies, Taylor & Jonker [44] introduced a set of differential equations that describe how the frequencies of strategies within the population change in time, i.e., the dynamics from any initial population state to an evolutionary stable one, or vice versa. Largely developed by Taylor & Jonker [44], and named by Schuster & Sigmund [64], the so-called replicator equation deterministically describes the relative spread of continuously changing frequencies of different strategies within a population playing an underlying game. In applying the replicator dynamics, one can model the changing frequencies of genotypes or phenotypes [41, 65]. For the concerns of this thesis which is focused on behavioural strategies, the replicator equation can model situations whereby strategies spread through imitation or learning [60, 65, 66]. Here the basic assumptions and simple applications of the replicator dynamics are given.

Consider a population of M different types represented by $S_1, S_2, ..., S_M$ strategies. Denote by x_i the frequency (proportion of individuals) of those playing strategy S_i within the population. A key assumption is that the population size N is infinite so that in an evolutionary setting, $x_i \in [0, 1]$ is continuous and differentiable in time. An equivalent understanding is that x_i is the probability that a randomly drawn individual has strategy S_i [41]. A natural constancy implies that

$$\sum_{i=1}^{M} x_i = 1 \tag{1.12}$$

To measure the evolutionary change in the relative frequencies of strategies, one needs to ascribe a fitness to each strategy. The initial derivation of the replicator equation equates the fitness f_i of the strategy S_i to the expected number of offspring of that strategy [44]. In this sense, individual fitness can be thought of as the product of vitality; the probability that the individual reaches productive age, and fecundity; the number of offspring that the individual has. We assume that all population members with strategy S_i possess the same vitality and fecundity, and thereby fitness f_i. Accordingly, the average fitness (or growth rate) in the population is defined by the scalar

$$\langle f \rangle = \sum_{i=1}^{M} x_i f_i \tag{1.13}$$

where $f_i : \Delta^M \to \mathbb{R}$ is generally a function of the population state $\mathbf{x} = (x_1, ..., x_M) \in \Delta^M$, i.e., $f_i = f_i(\mathbf{x})$. We would like to capture the idea that the higher the fitness of strategy S_i in comparison to the mean fitness of the entire population, the more advantageous it is to the growth of that strategy. As such, the relative rate of growth of the frequency of x_i is governed by how well this strategic type is doing relative to the population average. One arrives at $\dot{x}_i/x_i = f_i - \langle f \rangle$ as a measure of the advantage or disadvantage of strategy S_i over other population strategies in terms of reproductive success. In this case, the growth rate of the population frequency using strategy S_i is given by the replicator equation:

$$\dot{x}_i = x_i(f_i(\mathbf{x}) - \langle f \rangle) \ , \ \ i = 1, ..., M. \tag{1.14}$$

Equation (1.14) leaves both $\mathrm{bd}\Delta^M$ and $\mathrm{int}\Delta^M$ of the simplex invariant. This means that any trajectory starting from the interior of the simplex will remain there for all time and never reach the boundary [41]. Likewise, trajectories starting on the boundary stay on the boundary. This property allows one to find lower dimensional subsets of the simplex, for which any trajectory starting from must stay there. The roots of equation (1.14) provide equilibrium frequencies of the strategy S_i and characterise the "population dynamics".

Suppose that individuals of the population are paired randomly and engage in a symmetric game as described in section 1.1.1. In this case, an individual's fitness can be found from a payoff matrix of the form

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & \cdots & \cdots & a_{2M} \\ \vdots & & \vdots & \\ a_{M1} & \cdots & \cdots & a_{MM} \end{pmatrix} \tag{1.15}$$

An individual playing the pure strategy S_i against an ensemble of strategies $\mathbf{y} = (y_1, y_2, ..., y_M)^T$ has an expected payoff $\pi(i, \mathbf{y}) = (A\mathbf{y})_i$. Hence, the fitness of a pure strategy S_i is simply $f_i(\mathbf{x}) = (A\mathbf{x})_i = \sum_j a_{ij}x_j$. The average fitness in the population with state $\mathbf{x}$ is given by $\langle f \rangle = \mathbf{x}A\mathbf{x} = \sum_i \sum_j x_i a_{ij} x_j$. In evolutionary dynamics modelled by the replicator equation, the rate of increase or decrease of a strategy is proportional to fitness, and typically depends on the relative fitness of the strategy as compared to the average fitness in the population. This gives rise to the standard the replicator equation (1.14) which describes the per-capita growth rate (logarithmic derivative) of frequency of individuals playing strategy S_i:

$$\dot{x}_i = x_i\big((A\mathbf{x})_i - \mathbf{x}A\mathbf{x}\big) \quad i = 1, ..., M. \tag{1.16}$$

A fixed point $\hat{\mathbf{x}}$ of equation (1.16) is stable (in Lyapunov sense) if for every neighbourhood U of $\hat{\mathbf{x}}$ there exists a neighbourhood V of $\hat{\mathbf{x}}$ such that $\mathbf{x} \in V \implies \mathbf{x}(t) \in U \ \ \forall t \geq 0$. The fixed point $\hat{\mathbf{x}}$ is said to be attracting if it has a neighbourhood U such that $\mathbf{x}(t) \to \hat{\mathbf{x}}$ as $t \to \infty$ and $\forall \mathbf{x} \in U$. The fixed point $\hat{\mathbf{x}}$ is asymptotically stable if it is both stable and attracting.

If the replicator dynamics a posteriori are obtained from matrix games, then there is a natural mathematical connection between the stability properties of the dynamics' fixed points, and the stability properties of strategies in the underlying game. These features are summarised in the Folk theorem of evolutionary game theory [41]. A Nash equilibrium population state, $\mathbf{v} \in \Delta^M$ is such that self interactions of $\mathbf{v}$ are at least as or more than successful than any other state $\mathbf{y} \in \Delta^M$, [43], i.e.,

$$\pi(\mathbf{v}, \mathbf{v}) \geq \pi(\mathbf{v}, \mathbf{y}) \tag{1.17}$$

A point $\mathbf{v} \in \Delta^M$ is an evolutionary stable state if

$$\pi(\mathbf{v}, \mathbf{y}) > \pi(\mathbf{y}, \mathbf{y}) \tag{1.18}$$

for all $\mathbf{y} \neq \mathbf{v}$ [43]. The main features of interest of the Folk, and other theorems (which can be found in Cressman & Tao [43] or Hofbauer & Sigmund [41]) are as follows: if a rest point of the replicator dynamics $\mathbf{v}$ is the convergent limit of an orbit contained in $\mathrm{int}\Delta^M$, then it is a Nash equilibrium in

the sense of condition (1.17). If a rest point $\mathbf{v}$ of the replicator dynamics is asymptotically stable then it is a Nash equilibrium. If the state $\mathbf{v}$ is an evolutionary stable state as in condition (1.18) then it is a locally asymptotically stable rest point of the replicator dynamics (1.16) [41].

In more detail, it is useful to connect the replicator dynamics' equilibrium frequencies x_i of each pure strategy S_i in the population, and the probabilities that in a single game, an individual will play a mixed strategy in which each pure strategy S_i is played with probability x_i. Let $\pi(\mathbf{p}, \mathbf{q})$ now represent the payoff to a player with mixed strategy profile $\mathbf{p}$ playing an individual with mixed strategy $\mathbf{q}$. Consider the rest point of the replicator dynamics as $\mathbf{p}$ such that all $\dot{p}_i = 0$ where p_i is now interpreted as the frequency of the strategy S_i. Then, if $\hat{\mathbf{p}}$ is a strict Nash equilibrium of the game, then $\hat{\mathbf{p}}$ is locally asymptotically stable solution of the replicator equation (where the probabilities of playing a pure strategy in the game equal the equilibrium proportion of pure strategy in the population). An ESS $\hat{\mathbf{p}}$ in the interior of Δ^M is a globally stable rest point of the replicator dynamics. The converse is not true [67]. A Nash equilibrium $\hat{\mathbf{p}}$ of the underlying game with matrix payoff matrix A is a rest point of the replicator dynamics [33, 41].

A natural extension to the replicator equation is to allow for mutations between the fixed set of strategies. The replicator-mutator equation considers mutations in the sense of replicas of a strategist being characterised by different strategies. Selection is accounted for by supposing that the replication rate of a strategy is proportional to its success in the underlying game. The probability that during the replication process, an individual choosing S_i produces an individual choosing S_j is given by q_{ji}. The replicator-mutator equation reads as

$$\dot{x}_i = \sum_{j=1}^{M} x_j f_j(\mathbf{x}) q_{ji} - x_i \langle f \rangle \;\; i = 1, ..., M. \tag{1.19}$$

Equation (1.19) is applicable to a broad range of biological problems, for example it has been used on both evolutionary and ecological timescales [68].

In this research, the replicator-mutator equation is appealing as one can apply it to fast behavioural timescales. In a behavioural setting, the dynamics have been used to model the change in dominant languages [69, 70] as well as popular dance trends [71]. In this context, mutations are equivalently interpreted as mistakes in an imitation process: instead of imitating the better performing strategies, an individual chooses to randomly "explore" its choices of strategy [40]. Including exploration or "random sampling" of an individual's strategic choices is important because animals rarely act rationally, i.e., do not always choose the optimal behaviour. It is far more likely that animals demonstrate a small but non-zero propensity to randomly sample from a number of alternate behaviours [21].

The replicator is conceptually simple and reasonable way to model a range of biological interactions. This however comes at a cost. As noted by Nowak & Sigmund [65], only selection in the population (sometimes "demographic") dynamics is accounted for in the equation. That is, because the replicator equation only allows for only a finite set of fixed strategies which are invariant over time, i.e., does not incorporate mutation, then the evolutionary dynamics driven by mutation and selection cannot be captured. Another issue when using the replicator dynamics is an unrealistic assumption of an infinitely large population. Later in this introduction (section 1.2) we briefly

introduce game dynamics in finite populations which address this issue.

Replicator dynamics for two strategy games

Here we introduce some simple game dynamics using the replicator equation. The simplest case of a game is when there are two strategies S_1, S_2 and a payoff matrix game representation of

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{1.20}$$

The relative frequencies of strategies S_1, S_2 are respectively given by x_1 and x_2. The population state $\mathbf{x} = (x_1, x_2)^T$ allows the respective fitness for each strategy to be given by $f_1(\mathbf{x}) = (A\mathbf{x})_1$ and $f_2(\mathbf{x}) = (A\mathbf{x})_2$ and as such the replicator equation gives $\dot{x} = x(1-x)\big((A\mathbf{x})_1 - \mathbf{x}A\mathbf{x}\big)$ as we can set $x = x_1, 1 - x = x_2$. As the assumption of a well mixed population means that the expected payoff of each strategist are evaluated according to random pairwise interactions with the entire population, one can write

$$\pi_1(\mathbf{x}) = (A\mathbf{x})_1 = a_{11}x + a_{12}(1-x) \tag{1.21a}$$
$$\pi_2(\mathbf{x}) = (A\mathbf{x})_2 = a_{21}x + a_{22}(1-x) \tag{1.21b}$$

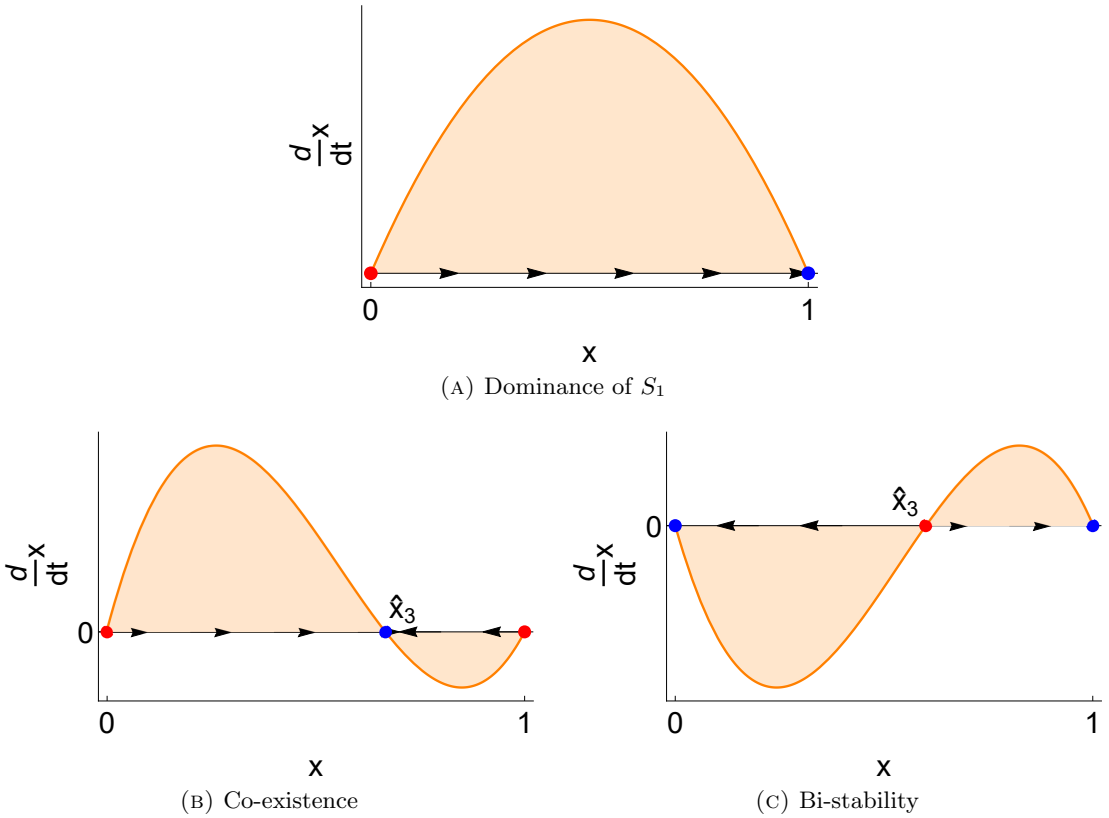
and the average payoff in the population $\langle \pi \rangle = \mathbf{x}A\mathbf{x} = x\pi_1(\mathbf{x}) + (1-x)\pi_2(\mathbf{x})$. Substituting equations (1.21) into the replicator equation (1.16) allows a simplification of the frequency dynamics such that

$$\dot{x} = x(1-x)\big(a_{11}x + a_{12}(1-x) - a_{21}x - a_{22}(1-x)\big) \tag{1.22a}$$
$$= x(1-x)(cx + d) \tag{1.22b}$$

where $c = a_{11} + a_{22} - a_{12} - a_{21}$ and $d = a_{12} - a_{22}$. Clearly there are up to three fixed points of equation (1.22b), one on each boundary of the simplex Δ^2 given by $\hat{x}_1 = 0$ and $\hat{x}_2 = 1$ and a possible interior solution on the interior of the simplex $\hat{x}_3$. The interior solution depends on the payoff matrix through the values of c and d: if physical, (existing in the interval $(0, 1)$) one finds the interior fixed point $\hat{x}_3$ by solving $cx + d = 0$. The values of the payoffs in the matrix (1.20) specify the stability properties of each fixed point. Stability can be determined by standard methods of non-linear dynamics, namely the sign pattern of the eigenvalues of the linearised system [72]. By denoting the Jacobian matrix (which is not really a matrix for one-dimensional systems but is for $M > 2$) by $J(x)$ a quick check finds that $J(0) = a_{12} - a_{22}$, $J(1) = a_{21} - a_{11}$ and $J(\hat{x}_3)$ is straightforward to calculate (but not explicitly written for space).

If $a_{11} = a_{21}$ and $a_{12} = a_{22}$ then $\dot{x} = 0$ for any $x \in [0, 1]$ and one speaks of neutral evolution; no strategies do not change. In all other scenarios there are three possibilities; dominance, bi-stability and co-existence. If $a_{11} > a_{21}$ and $a_{12} > a_{22}$ then clearly $J(0) > 0, J(1) < 0$ and the strategy S_1 dominates (by consistency of stability $\hat{x}_3$ cannot be physical). See Fig 1.1a for an illustration.

If both $J(0) > 0$ and $J(1) > 0$ then by fundamentals of dynamic systems the interior must lie between the two boundary states and be stable. A situation of coexistence of the strategies persists



(A) Dominance of S_1

(B) Co-existence

(C) Bi-stability

FIGURE 1.1: For illustration the unstable fixed points are shown by red points and stable fixed points are blue. (A): $\hat{x}_2 = 1$ is a stable fixed point of the dynamics and $\hat{x}_1 = 0$ unstable. For all initial states $x_0 \in [0, 1)$ the time change given by $\dot{x}$ is positive and x increases up to the point where the entire population adopts S_1. (B): The interior fixed point $\hat{x}_3$ is stable and each of population states $\hat{x}_1$ and $\hat{x}_2$ are unstable. Any initial state converges to the mixed population state: for $x \in (0, \hat{x}_3)$ we see $\dot{x} > 0$ (x is increasing towards $\hat{x}_3$) and for $x \in (\hat{x}_3, 1)$ we see $\dot{x} < 0$ (the frequency of S_1 strategies is decreasing towards $\hat{x}_3$). (C): Both edge states of all adopt S_1 and all adopt S_2 are stable. Depending on which side of $\hat{x}_3$ the initial state x_0 lies, the entire population plays either strategy S_1 (if $x_0 > \hat{x}_3$) or strategy S_2 (if $x_0 < \hat{x}_3$)

as shown by the stable fixed point $\hat{x}_3$ in Fig 1.1b. This case is representative of $a_{11} < a_{21}$ and $a_{12} > a_{22}$.

Bi-stable dynamics occur when the payoff matrix entries are such that $a_{11} > a_{21}$ and $a_{12} < a_{22}$. A class of games known as coordination games [57] describe such cases where each strategy requires some threshold density to have an advantage over the other strategy. Both of the vertices of the simplex $\hat{x}_1 = 0$ and $\hat{x}_2 = 1$ are stable so that the population is at a stable state if all adopt S_1 or S_2, and it is the location of the unstable interior solution that then determines the threshold density required for either strategy to dominate. Denote x_0 the initial state of x. If $x_0 < \hat{x}_3$ then the initial population state lies in the basin of attraction of $x = 0$ and thus the strategy S_2 dominates, where as if $x_0 > \hat{x}_3$ then S_1 dominates as shown in Fig 1.1c.

Cyclic 3 strategy replicator dynamics

Perhaps the most common example of the replicator dynamics with three strategies is the Rock-Paper-Scissors game [33, 34]. The idea behind the game is that of cyclic dominance: that the

strategy R beats strategy S, and strategy S beats strategy P, but then strategy P beats strategy R. In a pairwise encounter, the winning strategy obtains payoff 1, and the losing strategy gets $-b$. Thus, the payoff matrix is

$$A_{RPS} = \begin{pmatrix} 0 & -b & 1 \\ 1 & 0 & -b \\ -b & 1 & 0 \end{pmatrix} \tag{1.23}$$

Any payoff matrix demonstrating cyclic dominance can be transformed to one of this form by linear algebra operations, which do not affect the dynamics orbits [41]. The payoff for each of the pure strategies R, P and S, with respective frequencies x_R, x_P and $x_S = 1 - x_R - x_P$ are

$$\pi_R = -bx_P + (1 - x_R - x_P) \tag{1.24a}$$
$$\pi_P = x_R - b(1 - x_R - x_P) \tag{1.24b}$$
$$\pi_S = -bx_R + x_P \tag{1.24c}$$

and the replicator dynamics can be fully specified by the frequencies of x_R and x_P:

$$\dot{x}_R = x_R(\pi_R - \langle\pi\rangle) \tag{1.25a}$$
$$\dot{x}_P = x_P(\pi_P - \langle\pi\rangle) \tag{1.25b}$$

where $\langle\pi\rangle = x_R\pi_R + x_P\pi_P + (1 - x_R - x_P)\pi_S$ is the mean population payoff.

Before analysing the fixed points and their stability, it is useful to note that the Poincaré-Bendixon theorem applies here such that there cannot be an isolated periodic orbit, i.e., there are no limit cycles.

The fixed points of the dynamic system (1.25a)-(1.25b) can be fully specified by (x_R, x_P): three of which exist on the boundary of Δ^3 given by $\mathbf{e}_S = (0,0), \mathbf{e}_P = (0,1), \mathbf{e}_R = (1,0)$. There exists an interior fixed point given by $\hat{\mathbf{x}}$. The stability of each boundary solution is determined by the (two dimensional) Jacobian matrix:

$$\begin{pmatrix} \frac{\partial}{\partial x_R}\dot{x}_R & \frac{\partial}{\partial x_P}\dot{x}_R \\ \frac{\partial}{\partial x_R}\dot{x}_P & \frac{\partial}{\partial x_P}\dot{x}_P \end{pmatrix} \tag{1.26}$$

which produces eigenvalues $\lambda_1^{\mathbf{e}_i} = 1, \lambda_1^{\mathbf{e}_i} = -b$ for $i \in \{R, P, S\}$. Thus, each of the vertices of Δ^3 are saddle fixed points and form what is called a heteroclinic cycle, together with the edges of Δ^3 [41]. There are three generic cases of $b = 1, b < 1, b > 1$ that determine the stability of $\hat{\mathbf{x}}$.

For $b > 1$ the determinant of A_{RPS} is negative then $\hat{\mathbf{x}}$ is unstable. Any orbit starting from the interior of Δ^3 converges to the heteroclinic cycle on the boundary, although since $\mathrm{int}S^3$ is invariant, the trajectory never quite reaches it.

If $b < 1$ then the determinant of A_{RPS} is positive, and $\hat{\mathbf{x}}$ is an asymptotically stable fixed point. In this case all orbits inside the simplex converge to it.

When $b = 1$ one finds that $\langle\pi\rangle = 0$, and thus

$$\dot{x}_R = \pi_R x_R = x_R(1 - x_R - 2x_P) \tag{1.27a}$$
$$\dot{x}_P = \pi_P x_P = x_P(2x_R + x_P - 1) \tag{1.27b}$$

By multiplying equation (1.27a) by $(1 - x_p - 2x_R)$ and equation (1.27b) by $(2x_P + x_R - 1)$, one finds that

$$\dot{x}_R\left(\frac{1}{x_R} - \frac{x_P}{x_R} - 2\right) + \dot{x}_P\left(\frac{1}{x_P} - \frac{x_R}{x_P} - 2\right) = 0. \tag{1.28}$$

Multiplying equation (1.28) by $x_R x_P$ and applying the product rule, one then finds

$$0 = x_R(1 - 2x_P - x_R)\dot{x}_P + x_P(1 - x_P - 2x_R)\dot{x}_R = \tag{1.29a}$$

$$\frac{d}{dt}\big(x_R x_P(1 - x_R - x_P)\big) =: G(x_R, x_P) \tag{1.29b}$$

and $H(x_R, x_P) = \int G(x_R, x_P)dt$. The function $H(x_R, x_P) = x_R x_P(1 - x_R - x_P)$ is a constant of motion [41] and since

$$\frac{\partial}{\partial x_R}H(x_R, x_P) = x_P(1 - x_P - x_P) - x_P x_R \tag{1.30}$$

equals zero at $x_R = x_P = 1/3$ and that $\partial_{x_R^2}H(x_R, x_P) = -2x_P < 0$ the center $\hat{\mathbf{x}} = (1/3, 1/3)$ is a maxima. The constant level sets, $\{(x_R, x_P) \in \text{int}\Delta^3 : H(x_R, x_P) = \text{Constant}\}$, of which there are infinitely many, are closed curves surrounding $\hat{\mathbf{x}}$. Each closed periodic orbit has a time average of $\hat{\mathbf{x}}$ [41]. In other words all trajectories inside the simplex return to their starting place.

1.1.6 Continuous strategies and the G-function

Up to this point, the strategies in the considered game dynamics have been discrete and finite. Later in this introduction the interpretation of continuous strategies is outlined by the application of adaptive dynamics. Here, it seems appropriate to introduce the G-function approach pioneered by Vincent & Brown [73] and co-authors as a framework for converting discrete strategy matrix games to continuous strategy games. The key concept is a so-called fitness generating, or G-function (see [74, 75] and references therein) which generates the fitness of any individual within a population given a dynamic ecological system.

The way to convert matrix games into continuous strategy games is shown in great detail in [73, 74] where the reader is referred for a more comprehensive outline. Here the basic idea is presented. Consider a population in which there are $i = 1, 2, ..., \rho$ strategies. Each strategy, denoted u_i, is limited to some feasible set S, and each individual in the population adopts some strategy. Note that in general strategies can be vectors, but in this thesis only scalars are considered. Each strategy u_i has a population density x_i. The individuals using strategy u_i have some fitness $H_i(\mathbf{u}, \mathbf{x})$, where fitness depends on the population strategies $\mathbf{u} = (u_1, u_2, ..., u_\rho)$ and densities $\mathbf{x} = (x_1, x_2, ..., x_\rho)$. Similar to the replicator dynamics of section 1.1.5, one can use a differential equation to model the change in the population densities:

$$\dot{x}_i = x_i H_i(\mathbf{u}, \mathbf{x}) \text{ for } i = 1, ..., \rho \tag{1.31}$$

The definition of a G-function according to Vincent & Brown [73] is "a fitness generating function if and only if $G(v, \mathbf{u}, \mathbf{x})|_{v=u_i} = H_i(\mathbf{u}, \mathbf{x})$." The strategy v is referred to as a virtual variable, which upon substitution with any other $u_i \in \{u_1, u_2, ..., u_\rho\}$ generates the fitness of those individuals using strategy u_i.

Unlike Maynard Smith & Price [39] who distinguish between the fitness rates of distinct strategies, but not between a focal strategy and other distinct strategies, the G-function separates the focal strategy from all other strategies at play via the virtual variable v. Thus, fitness is modelled from an individual selection point of view (the strategy maximises individual fitness and not the fitness of all those adopting the strategy). Without the virtual variable fitness is not frequency dependent with respect to all individuals adopting the same strategy. The G-function, roughly speaking, can be redefined as the fitness of a very small population of mutants [76–79] which provides a direct link between evolutionary (matrix) game dynamics and the invasion analysis of the field of adaptive dynamics (see section 1.3).

The G-function is able to model density or frequency dependence. For the interpretation of different strategies competing within a single population it is most convenient to represent the G-function in terms of the frequency of each strategy. This thesis applies a frequency dependent approach to allow a proportional representation of each strategy. As such one can write $G(v, \mathbf{u}, \mathbf{x})$ with $x_1 + ... + x_\rho = 1$ and

$$\frac{\partial}{\partial \mathbf{x}} G(v, \mathbf{u}, \mathbf{x}) \neq 0. \tag{1.32}$$

In the methodology of Vincent & Brown [73] the notion of an adaptive landscape is what governs the strategy dynamics. The adaptive landscape is a plot of $G(v, \mathbf{u}, \mathbf{x}) - \bar{G}$, i.e., fitness if the virtual strategy is adopted, relative to the average fitness in the population with existing strategies given by $\mathbf{u}$ and frequencies $\mathbf{x}$. Using the idea of a fitness landscape, the G-function provides an alternative ESS definition, namely the "ESS Maximum principle". A coalition of strategies $\mathbf{u}^* = (u_1^*, u_2^*, ..., u_s^*)$, assuming that there are $\rho - s$ strategies not being played, (we can order such that $x_{s+1} = ... = x_{\rho-1} = x_\rho = 0$), is an ESS if

$$\max_{v \in S} \left\{ \left(G(v, \mathbf{u}, \mathbf{x}) - \bar{G} \right)|_{\mathbf{u}=\mathbf{u}^*, \mathbf{x}=\mathbf{x}^*} \right\} = 0 \text{ at } v = u_1^*, u_2^*, ..., u_s^*. \tag{1.33}$$

According to equation (1.33), each strategy of an ESS must be a (global) maximum of adaptive landscape. That is, any possible strategy $v \in S$ attains its global maximum value at $u_1^*, u_2^*, ..., u_s^*$. This property ensures that the coalition $\mathbf{u}^*$ is resistant to invasion since it resides at a peak of the adaptive landscape [73]. Clearly in a monomorphic population, the ESS Maximum principle can be thought of as an invasion landscape where u^* lies at a maxima (of value zero) of the invasion fitness of v, and is therefore resistant to invasion. We shall see later on in section 1.3 that this concept is very similar to the "invasion fitness" of adaptive dynamics. Further conditions on strategies such as dynamic attainability, i.e., able to invade other populations can be found in [73, 80, 81].

Thus, using equation (1.31) and the adaptive landscape, the G-function models the population and strategy dynamics of the strategy u_i according to

$$\dot{x}_i = x_i \left(G(v, \mathbf{u}, \mathbf{x}) - \bar{G} \right)|_{v=u_i} \tag{1.34a}$$

$$\dot{u}_i = k \frac{\partial}{\partial v} G(v, \mathbf{u}, \mathbf{x})|_{v=u_i} \tag{1.34b}$$

where k is a simple timescale constant. Generally, the method does not specify a need for a separation in timescales, although implicitly does so through assuming that the population dynamics at equilibrium $\mathbf{x}^*$, i.e., $G(v, \mathbf{u}, \mathbf{x}^*) - \bar{G} = 0$, whenever strategy adaptations occur.

The G-function is a useful tool and is closely connected with the theoretical synthesis of adaptive dynamics (which are introduced in section 1.3 of the introduction). In fact, Vincent & Brown [81] applying the G-function methodology, and Doebeli *et al.* [61] applying computer simulations to an adaptive dynamics model, found the same qualitative results when studying the continuous snowdrift game. Moreover, the G-function is flexible enough to model a range of ecological and evolutionary dynamics. Dercole & Geritz [79] used the G-function to comprehensively classify degenerate evolutionary outcomes in a resident and mutant competition model. Krivan & Cressman [82] applied the G-function to general predator-prey models to find whether mutant strategies can invade populations in which strategies change instantaneously, i.e., are plastic. Despite intermittent periods of relative success, the G-function has not quite become a dominant force in the modelling of evolutionary dynamics.

1.2 Stochastic dynamics

In section 1.1.5 the replicator and replicator-mutator dynamics where introduced as way to describe the changing frequencies of strategy types in an infinitely large population. In some biological problems, the interactions of phenotypes do not precisely fit a continuous model: the deterministic equation is a (mean field) approximation. For the scope of this thesis the deterministic description is a sufficient tool to tackle the subsequent research problems. However, it can be advantageous to derive the deterministic equation using a stochastic framework. In this way one can specify biologically appropriate mechanisms which describe how strategies change at a specific time, i.e., through births and deaths, or learning, or imitation. The resulting dynamics emerge naturally and should be a good description of the biological problem: the biological problem should not be assumed to fit some arbitrary equation, e.g., the replicator dynamics.

Stochastic dynamics considers that a population contains a finite number of individuals. In this case, rather that the relative frequencies of strategies which are continuous variables, one considers integer quantities of the number of certain strategies in the population. The change in strategies in finite populations are governed by probabilistic laws, i.e., through a master equation. The master equation depends on specific transition probabilities that are derived from the underlying game. Thus, the complexity of the stochastic process depends on the complexity of the biological problem.

In this section, the intention is not to give a comprehensive introduction to stochastic game dynamics, but rather to identify the important concepts that are used as a way to derive deterministic equations. The important features for this thesis (i): the master equation shortly introduced in section 1.2.1 (ii): how the underlying biological problem should set transition probabilities that enter the master equation, and (iii): how one can expand the master equation to recover a deterministic characterisation of the problem.

1.2.1 From stochastic dynamics to a deterministic description

The approach used in this thesis is to represent the number of individuals adopting different strategies by a discrete set of variables referred to as the system. Continuous time descriptions are also possible but for the scope of this thesis, a discrete description suffices. The probability that at a certain time the variables take on some specific value is characterised by a probability density

function with a discrete support. Values which the random variable assume at the time is known as the systems state. Generally the systems state changes according to a transition matrix which specifies the probability that from one time to another the state jumps to another state. Interactions between strategies in the underlying game determine the nature of the transition probabilities and thus the probabilistic movement of the systems state.

If the probability that the system moves to a state at a given time does not depend on the systems previous states, i.e., the system is Markovian, then a master equation describes the change in time of the probability density function. From the master equation follow a series of differential equations for each moment defining the stochastic process. The moments scale with the systems size, in particular, the average step size and variance of step size are inversely proportional to the system size and therefore decrease with increasing system size [83, 84].

Assuming a large enough system size (or population) only the mean and variance respectively, are required to model the macroscopic (deterministic) and fluctuating dynamics respectively. In the subsequent sections the van Kampen expansion is introduced [85]. The expansion results in a continuous approximation for the discrete random variables. Usually a Fokker-Planck equation [83] shows the average directional changes of the system through the "drift" term and fluctuations though the "diffusion" term (where the diffusion term within the physics literature is often known as "genetic drift" in the biology literature). A population large enough for the expansion to give rise to scaled variables that are approximately continuous is essential.

The master equation

Here, an introduction to one-dimensional Markov chains that are common throughout evolutionary game theory is given. In this thesis, we only treat one dimensional discrete state, discrete time Markov chains. Given two strategies (A and B) and a population of size $N < \infty$, at any time step, the number of A strategists, $i \in \{0, 1, ..., N\}$ fully specifies the systems state (since there are $N - i$ B strategists). The change in state i with each time step defines a stochastic process. In the subsequent chapters of this thesis we consider populations in which the number $0 \leq i \leq N$, of A strategies, can only move to $i \pm 1$ or remain unchanged as visualised in Fig 1.2. These processes are known as one-step [84], birth-death [85], or Generation-Recombination (G-R) processes [83]. The so-called generation rate $G(i, t)$ is the probability per unit time for a transition from i to $i + 1$, and the recombination rate $R(i, t)$ is the probability per unit time for a jump from from i to $i - 1$.

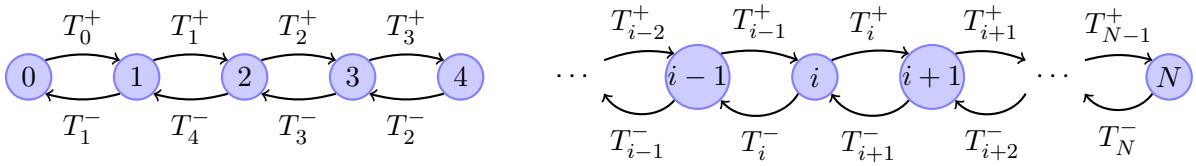


FIGURE 1.2: The one-step process and transition probabilities. The population state $i \in \{0, 1, ..., N\}$ increases by one ($i \to i + 1$) with probability T_i^+ and decreases by one ($i \to i - 1$) with probability T_i^-. A natural condition is $T_N^+ = T_0^- = 0$ so that the population size remains constant. If there are no mutations (i.e., $\mu = 0$) between the A and B strategies, one finds that $T_N^- = T_0^+ = 0$ due to absorbing boundary conditions, whereas when $\mu > 0$ this no longer holds.

The generation and recombination rate define a transition matrix,

$$T_{i\to i+1} = T_i^+ = G(i,t), \qquad T_{i\to i-1} = T_i^- = R(i,t), \tag{1.35a}$$

$$T_{i\to i} = T_i^0 = 1 - T_i^+ - T_i^-, \qquad T_{i\to i\pm k} = 0, \text{ for } k \in \mathbb{Z}, k > 1 \tag{1.35b}$$

which specifies the permitted jumps of the systems state (see Fig 1.2). The probability of finding the system in state i at time t given by a master equation:

$$\frac{d}{dt}P^t(i) = G(i-1,t)P(i-1,t) - G(i,t)P(i,t) + R(i+1,t)P(i+1,t) - R(i,t)P(i,t) \tag{1.36a}$$

$$= T_{i+1}^- P^t(i+1) + T_{i-1}^+ P^t(i-1) - T_i^- P^t(i) - T_i^+ P^t(i). \tag{1.36b}$$

The boundary conditions are important to consider. So that the population size is constant, the transitions $T_0^- = 0$ and $T_N^+ = 0$ are necessary. As such, the transition matrix is zero everywhere except the main diagonal, and the diagonals above and below. Other properties of the boundary, for example absorbing, reflecting of natural depend on the details of the transition probabilities, and in particular whether mutational effects are incorporated into the process. These issues are addressed in the next section.

Traditionally, G-R process applies to births and deaths, and associates with G the birth of a new A strategist (and subsequent death of a B strategist) and with R the birth of a new B strategist (and death of an A strategy). Mutations traditionally correspond to the probability that a birth is not a clonal copy of the reproducing individual. In the absence of mutations, the states $i = 0, N$ are absorbing boundaries, whereas, with mutations the states are reflecting. Later in this thesis, we follow Antal & Nowak [66], and Nowak *et al.* [40] and use one-step processes to model imitation. Behavioural mutations, or "exploration" [40] in the transition probabilities are equivalent to the genetic mutations of the birth-death interpretation.

How interactions specify transitions

This section is motivated to outline the most appropriate choices for strategy updating mechanisms in the underlying game. Choosing a suitable microscopic biological process to model the increase or decrease in strategy with fitness is important for the stochastic dynamics of the system [86–91]. Microscopic interactions govern updating processes, and thus specify the transition probabilities and the complexity of the master equation of section 1.2.1. Here, in a finite population setting with two strategies A and B, the frequency dependent Moran process and the pairwise update process are introduced.

The interactions between strategies in the underlying game produce payoffs which specify the transition probabilities of the system. Let π_A and π_B be the respective payoffs for each strategy in the underlying game. It is important to specify a payoff to fitness mapping in finite systems. In the mapping, a parameter ω, reflects the intensity of selection: the relative contribution of the underlying game to individual fitness [92]. Fitness, $f_A = f(\pi_A\omega)$ (and $f_B = f(\pi_B\omega)$), can be modelled by any product of selection intensity and payoff so long as fitness is non-negative. The increase in strategy is proportional to the increase in fitness.

The frequency-dependent Moran process: a classical model used in population genetics and more recently been applied to evolutionary game theory [88, 93]. Fitness is usually set to a background fitness of 1, plus the payoff from the game with a mechanism describing the intensity of selection ω: $f_A = 1 - \omega + \omega\pi_A$. If $\omega = 0$ then selection is neutral (frequency independence) and if $\omega = 1$ then fitness is fully determined by the payoffs of the game.

At each time step an individual is chosen from the population with probability proportional to their fitness. The chosen individual produces an offspring which, to keep the population constant, then replaces another randomly chosen individual. A mutation rate μ, i.e., the probability that an individual produces offspring that are of a different strategy, can be included. In this case, the probability that the population gains one A strategist (and loses one B strategist) is the probability that the individual chosen for elimination is a B strategist, and that the offspring is an A strategist. This can occur in two ways: with a probability $1 - \mu$, an A (B) strategist is chosen to reproduce with probability ρ^A (ρ^B):

$$\rho^A = \frac{if_A}{if_A + (N-i)f_B} \quad , \quad \rho^B = \frac{if_B}{if_A + (N-i)f_B} \tag{1.37}$$

proportional to that individual's fitness relative to the average population fitness. The probability that population transitions from i number of A strategies to $i+1$ is $\rho^A(N-i)/N$ where $(N-i)/N$ is the probability that the individual chosen to be removed is a B type. Alternatively, with a probability μ, a B strategy is chosen to reproduce with a fitness proportionate probability, but has an offspring A strategy. The transition probabilities in the limit of $\mu \to 0$ are shown in Table 1.1. The adjusted replicator equation of section 1.1.5 is recovered in the infinite population limit, and for $\mu > 0$ the replicator-mutator equation (section 1.1.5) is recovered (see Table 1.1). Of interest to this thesis, is that the frequency dependent Moran process can describe the spread of behaviours in a cultural or learning setting [40, 66]. Viewing the process as a model of imitation, one can think of a focal individual being selected to "update" its strategic choice, and choosing a strategy with a probability equal to the total fitness of individuals using that strategy divided by the total fitness in the population [94].

Local (or pairwise) update : often used in situations where it may be unrealistic for individuals to compute the average fitness in the population [91]. Two random individuals from the population; a focal and a role model are chosen. The focal imitates the strategy of the role model with a probability depending on the difference in payoffs they receive. The focal does not necessarily adopt a strategy with higher payoff as imitation is probabilistic, and worse strategies may be imitated.

Traulsen *et al.* [93] proposed that the probability of imitation could be linear with the payoff difference, on top of a baseline probability of 1/2. This model limits the intensity of selection measured by the rate of probability increase. To account for strong intensities of selection, the Fermi function (from statistical physics) is more appropriate (see [89, 90, 95] for examples). Using the Fermi function, the relative fitness of an individual is given by the exponential $f_A = \exp[\beta\pi_A]$, and the difference between strategies, $\exp[\beta(\pi_B - \pi_A)]$ determines the probability that an individual imitates another. $\beta \in [0, \infty)$ measures the strength of selection, ranging from 0 and the case of neutral drift, to extreme selection for $\beta \to \infty$.

Moran	Local update	Fermi
$\rho^A = \frac{if_A}{if_A+(N-i)f_B}$	$\rho^A = \frac{1}{2}+\frac{\omega}{2}\frac{\pi_A-\pi_B}{\Delta}$	$\rho^A = \frac{1}{1+e^{-\omega(\pi_A-\pi_B)}}$
$T^+ = (1-\mu)\rho^A\frac{N-i}{N}+\mu\rho^B\frac{i}{N}$	$T^+ = \rho^A\frac{i}{N}\frac{N-i}{N}+\mu\frac{N-i}{N}$	$T^+ = \rho^A\frac{i}{N}\frac{N-i}{N}$
$0 \le \omega \le 1$	$0 \le \omega \le 1$	$0 \le \beta$
In the deterministic limit ($N \to \infty$) when $\mu = 0$		
Adjusted replicator eq: $\dot{x} = x(f_A - \langle f\rangle)\frac{1}{\langle f\rangle-\frac{1}{\omega}-1}$	Replicator eq: $\dot{x} = x(\pi_A - \langle\pi\rangle)\frac{\omega}{\Delta\pi}$	Replicator eq in* $\dot{x} = x(1-x)\tanh\left(\frac{\beta}{2}(\pi_A-\pi_B)\right)$
* Traulsen *et al.* [97]		

TABLE 1.1: Update processes and their deterministic limits

The ρ^A, ρ^B are interpreted as the probability that a role model is chosen to be imitated. These values can take on many forms, but the most common examples are based on linear [93], or Fermi functions. In the linear case used Traulsen *et al.* [96], one finds

$$\rho^A = \frac{1}{2}+\frac{\omega}{2}\frac{\pi_A-\pi_B}{\Delta\pi} \quad , \quad \rho^B = \frac{1}{2}+\frac{\omega}{2}\frac{\pi_B-\pi_A}{\Delta\pi} \tag{1.38}$$

where $\Delta\pi$ is the maximum payoff difference between two strategies. The probability of population transition from i number of A types to $i+1$ is $\rho(i/N)(N-i)/N$. Although behavioural mutations are likely in animal imitation, for general μ it is not possible to define a microscopic process with the deterministic limit of the replicator-mutator equation. Various transition probabilities have been proposed. An example used by Traulsen [96] is shown in Table 1.1. For this example, in the deterministic limit with $\mu = 0$, the replicator dynamics are recovered Table 1.1, but in the regime of $\mu > 0$ the replicator-mutator equation is not.

van Kampen's expansion to find the deterministic limit

Depending on the complexity of the payoff structure in the underlying game, and thus the transition probabilities, often the master equation (1.36a) cannot be solved analytically. To find a suitable description of the time evolution of the probability density function one can apply an expansion to the inverse of the system size. This approximates the spatial and temporal variables such that $x = i/N$ and $\Delta t = 1/N$ are assumed continuous. Changing from the extensive variable i, the number of individuals adopting strategy A, to the density x being an intensive variable of those individuals results in an approximation of a continuous stochastic process on the unit interval $[0, 1]$.

Starting with the transition probabilities, T_i^+ and T_i^-, and assuming a large enough population, the master equation (1.36b) governs the time evolution of the system. The intuitive idea is that the deterministic part of the system x should be of order N where as the fluctuating part of the system should be of a lower order $N^{1/2}$. Introducing the shift operators $\mathbb{E}^{\pm}f(n) = f(n \pm 1)$ and

the probability density $p(\xi,t) = P^t(Nx(t) + \sqrt{N}\xi)$ where the van Kampen ansatz

$$i = Nx(t) + \sqrt{N}\xi. \tag{1.39}$$

is used, one then Taylor expands the transition probabilities and shift operators (see [85] or [83] for details). After the Taylor expansions, and taking the time derivative of $p(\xi,t)$, one collects coefficients of order $N^{1/2}$ which yields the deterministic dynamics:

$$\frac{dx}{dt} = T^+(x) - T^-(x) \tag{1.40}$$

and collecting terms of order N^0 the stochastic dynamics are given by:

$$\frac{\partial p(\xi,t)}{\partial t} = -\frac{\partial}{\partial x}\Big(\underbrace{T^+(x) - T^-(x)}_{A(x,t)}\Big)\frac{\partial}{\partial \xi}\xi p(\xi,t) + \frac{1}{2}\Big(\underbrace{\frac{T^+(x) + T^-(x)}{N}}_{B^2(x,t)}\Big)\frac{\partial^2}{\partial \xi^2}p(\xi,t). \tag{1.41}$$

Equation (1.41) is a linear Fokker-Planck equation whose coefficients depend only on $x(t)$. The drift and diffusion terms are respectively given by $A(x,t)$ and $B^2(x,t)$.

The Fokker-Planck equation serves as a foundation for further analysis of the systems asymptotic behaviour [96]. Equation (1.41) is relevant to the subsequent chapters in this thesis because it provides a natural derivation of deterministic equations used to describe changing frequencies of strategies in infinite populations, based on microscopic interactions based on the underlying game. For example, different forms of the replicator dynamics and depend on the microscopic update rule [93, 98].

The Fokker-Planck equation (1.41) corresponds to a stochastic differential Langevin equation which describes the time evolution of x, rather than its probability density function. By applying the Ito calculus [83, 85] the following stochastic differential equation is found:

$$\dot{x} = A(x,t) + \Gamma(t)B(x,t) \tag{1.42}$$

where $\Gamma(t)$ is multiplicative white noise. Equation (1.42) can be written as

$$dx = A(x,t)dt + B(x,t)dW(t) \tag{1.43}$$

where $W(t)$ is the Wiener process of increment $w(\tau) = W(t+\tau) - W(t)$, with zero mean and autocorrelation $\langle w(\tau_1)w(\tau_2)\rangle = 2\min\{\tau_1,\tau_2\}$. We see that in the limit of $N \to \infty$ the noise term of equations (1.42) and (1.43) vanishes and we recover a replicator equation of section 1.1.5, given by $\dot{x} \propto x(1-x)(f_A(x) - f_B(x))$. Hence, one can derive a stochastic representation of the model which, under appropriate assumptions produce deterministic dynamics. Importantly, by considering the more specific way in which the strategies in the population change, e.g., via imitation, or births and deaths, or with mutational effects, the emerging deterministic equations should suitably reflect the underlying assumptions of the biological problem. A systematic method can be used: one defines specific updating processes which are appropriate for the biological problem. The microscopic interpretation of how the distribution of strategies change in the population thereby specifies the

transition probabilities which enter the master equation. By appropriately re-scaling state space and time, the expansion of the master equation (1.36b) leads to corresponding equations for the time evolution of the probability density of the state variable (equation 1.41), or the time evolution of the state variable itself (equation 1.42). By assuming a large enough system these equations predict exact frequencies of strategies within the population.

1.3 Adaptive dynamics

So far, static game theory has shown whether a mutant strategy can invade a population, deterministic game dynamics has provided a means to predict the frequencies of strategies competing in a population, and stochastic game dynamics have shown the importance of microscopic interactions specifying how strategies spread. Even with the methods developed so far, there remains an important and unanswered question: can an ESS strategy invade another population of strategies? In other words, do ESS strategies attract nearby populations of strategies.

The traditional approach of adaptive dynamics (AD), pioneered by Dieckmann & Law [99], Metz *et al.* [100] and Geritz *et al.* [101] provides methods for predicting the longer term evolution of quantitative traits. The natural place to start when applying to the techniques of AD, is to identify the traits that are subject to change. In this thesis, we stick to the evolutionary game theory convention of strategy rather than trait. The options of what types strategies are appropriate to model has somewhat changed as the modelling breadth of adaptive dynamics has considerably expanded since its conception (see Abrams [102] for a review). The framework has been used to model behaviours such as kleptoparatism [103] and the evolution of cooperation [60, 61].

Whatever strategy under consideration, the techniques of AD are essentially used to understand the fate of small perturbations of a strategy from an ecological perspective. A key feature of AD is that it connects population dynamics and evolutionary dynamics in a way such that the idea of frequency-dependence is incorporated into the framework. Ecological systems are constantly changing, and the environment in which individuals live in changes with such adaptations. As such, an eco-evolutionary feedback describes the changing strategies of individuals in a population.

1.3.1 Fundamentals

The first step using adaptive dynamics is to specify a quantitative phenotypic strategy x in a population of individuals. One then specifies a demographic model that accounts for the effects of biotic and abiotic factors on an individual's fitness. The demographic model can be as complicated as the biological problem. In the simplest form, the demographic model accounts for the competition between a resident and a mutant strategy ("resident-invader" model), and is explicitly described by an ODE model (see equation 1.44). The ODE model can be extended to any number of existing strategies, and thereby to polymorphic resident populations, but for the introduction presented here the focus is on only the basic framework of a resident and invader. Assume that there are ρ strategies present in the population (for the resident-invader model $\rho = 2$). Let x_i be the value of each strategy, and let N_i be the number of individuals playing strategy x_i. The strategies of each population can be represented by a strategy vector $\mathbf{x} = (x_1, x_2, ..., x_\rho)$, and the population sizes

can be expressed by $\mathbf{N} = (N_1, N_2, ..., N_\rho)$. The population dynamics are given by

$$\frac{dN_i}{dt} = N_i W(x_i, \mathbf{x}, \mathbf{N}), i = 1, 2, ..., \rho \tag{1.44}$$

where $W(x_i, \mathbf{x}, \mathbf{N})$ is the average growth rate of strategy x_i which depends on the composition of strategies $\mathbf{x}$ and the relative sizes $\mathbf{N}$, i.e., growth is frequency dependent.

The function W is usually interpreted as a fitness function and for analysis when $\rho = 2$ we write $W(x_i, x_j), \forall (x_i, x_j) \in \{x_1, x_2\} \times \{x_1, x_2\}$. Consider a monomorphic resident population in which only strategy x_1 is present. The growth of the population solely depends on the strategy x_1 itself and the environment. It is assumed that an equilibrium population size is naturally reached. Denote the equilibrium as N_1^*. At the ecological equilibrium characterised by N_1^*, the fitness - or growth rate - of an individual in the population must be equal to zero, otherwise the population would continue to grow indefinitely. Therefore by definition one has that $W(x, x) = 0$.

At the core of adaptive dynamics are two fundamental assumptions. First, competition between the resident and invader occurs on a fast (ecological or sometimes demographic [68]) timescale, whereas mutations (new invaders) are rare, and occur on a much slower (evolutionary) timescale. In this way, the population is nearly always monomorphic and one can reasonably expect the population to be at a dynamical equilibrium by the time the next mutant strategy occurs. Second, the success of the mutant strategy is fully determined by its initial growth rate when rare [99, 101, 104]. When combined, these two assumptions amount to a separation of time scales in the adaptive dynamics framework. The separation of timescales allows one to study monomorphic population evolution: mutations are small enough to be considered perturbations of the resident strategy (the population is approximately monomorphic). On the fast ecological timescale, competition amongst the strategies in the population results in one strategy fixating. By the time a mutant strategy enters the population, the resident population is at a monomorphic ecological equilibrium. In this way, the evolution of a monomorphic population can be described by a sequence of strategy substitution events [105], and each strategy substitution event leads to extinction or fixation of the mutant strategy [105, 106]. One can then predict the longer term changes in strategy within the population.

1.3.2 Time scales

The key idea here is that mutations are assumed to be rare so that the evolving population remains monomorphic. This allows two separate timescales to be examined: the demographic timescale on which competition between the strategies in the population, and fixation occur, and the evolutionary timescale on which a sequence of strategy-substitution drives the adaptations in the strategy space. Depending respectively on whether the mutant strategy fixates or not, the population converges to the mutant strategy or remains unchanged [107]. Thus each competitive event in the demographic model can at most, direct incremental changes in the strategy space. The evolutionary dynamics depend on each of the demographic events: each competition event specifies the directional change if any in the evolutionary timescale, and the sequence of such events gives rise to the evolutionary change. Because after each mutational event on the demographic timescale the population tends

back to a monomorphic dynamical equilibrium, when considering the longer term evolutionary dynamics one can consider that the population or demographic dynamics are constant [101, 106].

1.3.3 Invasion Fitness

The invasion fitness, $W(y,x)$, is defined as the per-capita growth rate of an initially rare mutant strategy y in an environment set by the resident population of strategy x. Since individuals with the mutant strategy y are considered rare, the effect that the mutant strategy exerts on the environment is negligible.

W is generally assumed to be a continuously smooth function of x and y. An approximation of $W(y,x)$ is found by Taylor expanding the mutant strategy about the resident strategy:

$$W(y,x) = W(x,x) + \left.\frac{\partial W(y,x)}{\partial y}\right|_{y=x} (y-x) \tag{1.45}$$

where by definition $W(x,x) = 0$ implies that the mutant y can invade if and only if $W(y,x) > 0$. From equation (1.45) the selection gradient is defined by

$$D(x) = \left.\frac{\partial W(y,x)}{\partial y}\right|_{y=x} \tag{1.46}$$

. and by the definition of the mutant invasion fitness equation (1.45), we see that when the selection gradient is positive (or negative) only mutants with strategies $y > x$ (or $y < x$) can invade in a neighbourhood of x.

1.3.4 Canonical Equation

The canonical equation of adaptive dynamics [99] gives the expected dynamics of the strategy values on an evolutionary timescale as the strategy values are sequentially updated through successive fixations of rare mutant types.

The canonical equation is a deterministic approximation derived via an infinite realisations of a stochastic process. The movement of a population through the strategy space occurs as a Markov chain and as such, the following master equation describes the evolution for the probability, $P(x,t)$, that at time t the resident population is characterised by the strategy x:

$$\frac{d}{dt}P(x,t) = \int [w(x|y)P(y,t) - w(y|x)P(x,t)]dy \tag{1.47}$$

Here $w(i|j)$ are the transition probabilities that the population changes from strategy i to strategy j. w depends on two factors: first the probability of a mutant y occurring in a population of resident x and second, the probability that the mutant does not immediately become extinct. The reader is referred to Dieckmann & Law [99] for a comprehensive ecological derivation of the transition probabilities. Here, the mathematical theory is given.

The mean path

The dynamics of the mean strategy are given by

$$\frac{d}{dt}\langle x\rangle = \int dx x \frac{d}{dt}P(x,t)$$

where using equation (1.47) one finds

$$\begin{aligned}\frac{d}{dt}\langle x\rangle &= \int dx x \int [w(x|y)P(y,t) - w(y|x)P(x,t)]dy \\ &= \int dx \int [xw(x|y)P(y,t) - xw(y|x)P(x,t)]dy \end{aligned} \tag{1.48a}$$

Writing the transition probability as $w(y|x) = w(x;\epsilon)$ with $y = x + \epsilon$, via a change of variables $x \to y - \epsilon$ which is absorbed by the integral limits, one finds that equation (1.48a) can be expressed as

$$\frac{d}{dt}\langle x\rangle = \int dx \int dy(y-x)w(y|x)P(x,t) \tag{1.49}$$

Taking the Van Kampen k^{th} jump moment [108]

$$a_k(x) = \int (y-x)^k w(y|x)dy \tag{1.50}$$

it follows that

$$\frac{d}{dt}\langle x\rangle = \int dx a_1(x)P(x,t) = \langle a_1(x)\rangle(t) \approx a_1(\langle x\rangle(t)) \tag{1.51}$$

If the first jump moment $a_1(x)$ is a linear function of x then $\langle a_1(x)\rangle = a_1(\langle x\rangle)$ [99] meaning

$$\frac{d}{dt}\langle x\rangle(t) = a_1(\langle x\rangle(t)) \tag{1.52}$$

If the first jump moment is a non-linear function of x then one instead should use

$$\langle a_1(x)\rangle \approx a_1(\langle x\rangle) + \frac{\langle (x-\langle x\rangle)^2\rangle}{2} a_1''(\langle x\rangle) + ... \tag{1.53}$$

As long as stochastic deviations from the mean path are relatively small, the approximation of (1.52) holds well. Although from equation (1.53), the evolution of $\langle x\rangle$ depends on its fluctuations as well as $\langle x\rangle$ itself, similar to studying any macroscopic quantity, we ignore such stochastic fluctuations and focus only on the mean path (compare for example to the deterministic dynamics in section 1.2.1). The assumption that the macroscopic quantity itself is a meaningful description of the biological process, when in reality a precise description should also refer to the variance, is generally defended (for example by Dieckmann & Law [99]) by reference to Van Kampen [108] who describes the "macroscopic approximation" as not needing to regard moments higher than the first. Thus, non-linear functions $a_1(x)$ are neglected and the approximation of (1.52) is used to describe the process.

Dieckmann & Law [99] go on to substitute the following explicit transition probability

$$w(y|x) = (y-x)\mu(x)N_x M(x, y-x) \left.\frac{\partial W(y,x)}{\partial y}\right|_{y=x} \tag{1.54}$$

into equation (1.52) to recover

$$\frac{d}{dt}\langle x\rangle = \int (y-x)^2 \mu(x) N_x M(x, y-x) \left.\frac{\partial W(y,x)}{\partial y}\right|_{y=x} dy. \tag{1.55}$$

which describes the evolution of $\langle x\rangle$. By changing the integration variables $y - x = \epsilon$, observing that $M(x, y-x)$ is symmetric about the origin, and omitting the averaging brackets one finds

$$\frac{d}{dt}x = \frac{1}{2}\mu(x)\sigma_\mu^2 N_x \left.\frac{\partial W(y,x)}{\partial y}\right|_{y=x} \tag{1.56}$$

where $\sigma_\mu^2 = \int \epsilon^2 M(y+\epsilon, \epsilon) d\epsilon$. Equation (1.56), known as the so-called "Canonical equation" shows the expected change in x per unit time, as averaged over infinitely many strategy-substitution events.

Deterministic approximation

Equation (1.56) is composed of two factors: the "evolutionary rate of change coefficient" and the "selection gradient" [99]. For the scope of this thesis which only considers scalar strategies, the term $\mu\sigma_\mu^2 N_x$ describing mutational effects can be set to 1 without loss of generality [77]. In this case, the sign of the selection gradient $D(x)$ (equation 1.46) describes the direction of evolutionary change.

1.3.5 Evolutionary Singular Strategies

Solutions $\hat{x}$ for which the selection gradient (equation 1.46) vanishes yield "evolutionary singular strategies" (singular strategies henceforth). The population evolves via successive strategy substitution events, where changes in the strategy x are directed by the sign of equation (1.46), until a neighbourhood is reached where the fitness gradient is zero, i.e., some $\hat{x}$. If there are no such singular strategies then x unilaterally increases or decreases depending on the sign of the selection gradient. If the population reaches a singular strategy $\hat{x}$, the selection gradient offers no further information on future dynamics. Following Geritz *et al.* [101] the higher order derivatives of the invasion fitness function are then used to classify the stability properties of the singular strategy.

ESS-stability

It is useful to compare the properties of a singular strategy to the ESS concept of Maynard Smith, i.e., whether the strategy is invasible. By taking the second order Taylor expansion of equation

(1.45), two conditions determine whether a singular strategy is resistant to invasion:

$$A \equiv \left.\frac{\partial^2 W(y,x)}{\partial y^2}\right|_{y=x=\hat{x}} (y-x)^2 > 0 \iff \hat{x} \text{ can be invaded} \tag{1.57a}$$

$$A \equiv \left.\frac{\partial^2 W(y,x)}{\partial y^2}\right|_{y=x=\hat{x}} (y-x)^2 < 0 \iff \hat{x} \text{ cannot be invaded} \tag{1.57b}$$

When condition (1.57b) holds ($A < 0$), the selection gradient $D(\hat{x})$ equals zero and the invasion fitness has a maxima at $y = \hat{x}$. Therefore, all strategies in some local neighbourhood U of $\hat{x}$, i.e., $x' \in U(\hat{x})$ must yield a negative fitness. Mutant strategies in U cannot invade $\hat{x}$ and the singular strategy is known as an "evolutionary stable strategy" (ESS) [101] or "ESS-stable" [35]. Singular strategies that are ESS-stable are evolutionary end points of the dynamics. Conversely when condition (1.57a) is observed and $A > 0$, the invasion fitness lies at a minima at $\hat{x}$ and all neighbouring strategies yield a positive invasion fitness. In this case $\hat{x}$ can be invaded.

Convergence stability

The ESS-stability concept informs whether a strategy can be invaded but doesn't inform whether strategy is an evolutionary attractor, i.e. whether it can invade a nearby population with a different strategy when initially rare. For a mutant with strategy $y = \hat{x}$ to be able to invade a nearby strategy x the condition $W(\hat{x}, x) > 0$ must be satisfied. Therefore by expanding equation (1.45) around x and evaluating at the point $x = \hat{x}$ it follows that

$$W(\hat{x},x) = W(\hat{x},\hat{x}) + \left.\frac{\partial W(\hat{x},x)}{\partial x}\right|_{\hat{x}=x} (x-\hat{x}) + \left.\frac{\partial^2 W(\hat{x},x)}{\partial x^2}\right|_{\hat{x}=x} \frac{(x-\hat{x})^2}{2}. \tag{1.58}$$

Noting that again the first two terms are zero we have the condition for invasion that

$$B \equiv \left.\frac{\partial^2 W(y,x)}{\partial x^2}\right|_{y=x=\hat{x}} > 0 \iff \hat{x} \text{ can invade} \tag{1.59}$$

Condition (1.59) is independent of the ESS conditions (1.57). One could possibly find a singular point $\hat{x}$ for which any population initially at this value would be stable, but populations arbitrary close to $\hat{x}$ could never reach the point. Even if $\hat{x}$ is uninvasible when resident and can invade other populations when rare, we do not know whether a resident population far from $\hat{x}$ will converge towards it or not.

Convergence stability of a singular strategy $\hat{x}$ requires the additional conditions of $W(y, x) > 0$ when $x < y < \hat{x}$ or $\hat{x} < y < x$, and $W(y, x) < 0$ for $|\hat{x} - y| > |\hat{x} - x|$. Since $D(x)$ changes sign at $\hat{x}$, one requires $D(x)$ to be a decreasing function of x:

$$\frac{d}{dx}D(x) = \frac{d}{dx}\left[\left.\frac{\partial W(y,x)}{\partial y}\right|_{y=x}\right] = \left[\frac{\partial^2 W(y,x)}{\partial x \partial y} + \frac{\partial^2 W(y,x)}{\partial y^2}\right]_{y=x} < 0. \tag{1.60}$$

Since at $W(x, x) = 0$ we see that all of the derivatives of $W(y, x)$ along the line $y = x$ are zero. Therefore

$$\frac{\partial^2 W(y,x)}{\partial x^2} + 2\frac{\partial^2 W(y,x)}{\partial x \partial y} + \frac{\partial^2 W(y,x)}{\partial y^2} = 0 \tag{1.61}$$

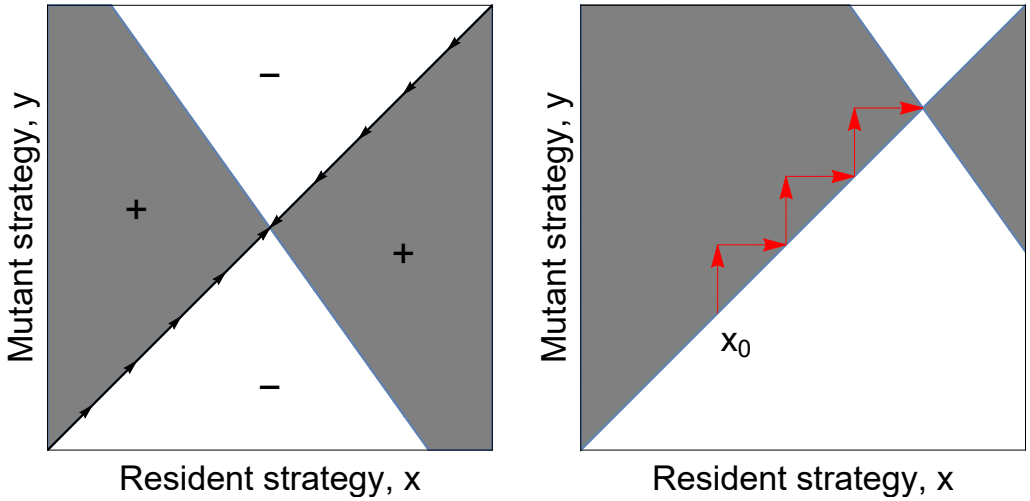
is used to get rid of the mixed partial derivative in equation (1.60). This leaves us with the following condition for a strategy $\hat{x}$ to invade another strategy x:

$$\left[\frac{\partial^2 W(y,x)}{\partial x^2} - \frac{\partial^2 W(y,x)}{\partial y^2}\right]_{y=x=\hat{x}} > 0 \iff \hat{x} \text{ is convergence stable} \tag{1.62}$$

If condition (1.62) holds then the singular strategy is "Convergence Stable" (CS) [109]: a population of a nearby strategy x can be invaded by a mutant strategy y that is nearer the singular strategy $\hat{x}$. Invasion occurs precisely when $W(y,x) > 0$ for $x < y < \hat{x}$ and $x > y > \hat{x}$. Combining the conditions (1.62) and (1.57b) leads to a stability condition sometimes referred to as a "continuously stable strategy" [109] (see also [74]).

Invasion analysis

The evolutionary dynamics can be visualised by the use of pairwise invasibility plots (PIPs). Recall that mutant strategy values y can only successfully invade the resident population adopting strategy value x if $W(y,x) > 0$. In such cases the outcome of a successful invasion results in the mutant replacing the resident. The fitness landscape experienced by the mutant then changes. Pairwise invasibility plots can be used to graphically describe the consequences of such an invasion event. In a PIP, for each resident strategy value x, the value of a mutant strategy y for which $W(y,x) > 0$ is depicted by a (usually shaded) region with a "+" and each strategy pairing such that $W(y,x) < 0$ by a "-" as shown in Fig 1.3. Since by definition $W(x,x) = 0$, along the principal diagonal there is a zero contour, i.e., along the line $y = x$ the mutant's fitness must equal the resident's because both adopt the same strategy value: the mutant's invasion fitness equals zero. Singular strategies are located at the intersection of the boundaries of the positive invasion fitness and the principal diagonal. In PIPs the invasion fitness of rare mutants can be seen by considering vertical lines along which the resident strategy is constant. Fig 1.3a illustrates the direction of evolutionary change in the strategy value by black arrows. We see that for each vertical line lying below the singular strategy, mutants with strategy values greater than the resident's strategy value have a positive invasion fitness. Such a visual description corresponds to the selection gradient $D(x) > 0$ (by equation 1.45) and evolutionary change results in strategies initially below the singular strategy increasing in value. Along any vertical line above the singular strategy, a mutants invasion fitness is only positive when $y < x$. Therefore $D(x) < 0$ and evolutionary change results in the strategy value decreasing. As strategies both above and below converge towards the singular strategy, the PIP illustrates the singular strategy is convergence stable, and since the vertical line through the singular strategy lies in a "-" region, the PIP illustrates that the singular strategy is evolutionary stable. Fig 1.3b visualises how the directional change in a strategy is governed by mutant invasion and subsequent replacement. We see that a sequence of strategy substitution events, as shown by the red arrows, results in the directional change of the strategy. Each mutant strategy is assumed to be sufficiently close to the resident (the vertical red arrows reflect small perturbations) so that equation (1.45) remains valid, i.e., $y \approx x$.



(A) (B)

FIGURE 1.3: How pairwise invasibility plots (PIPs) can be used to visualise the singular strategies and evolution via strategy-substitution events. (A): We see that the singular strategy is both evolutionary stable and convergence stable (see geometric interpretation as described in text). The population evolves (indicated by the black arrows) towards the singular strategy, located at intersection of the principal diagonal and the boundary of the positive mutant invasion fitness. Once the population reaches the singular strategy there is no further evolutionary change in the strategy value. (B): A close up from below the singular strategy. We visualise the strategy-substitution events: starting at a resident population with strategy x_0, each vertical red arrow shows how a mutant with a strategy of greater value than the current resident can invade. The following horizontal arrows describe the mutant replacing the resident (the population is again monomorphic with all adopting the mutant strategy). Each invasion-replacement event constitutes a strategy-substitution.

Fig 1.4 illustrates four examples of PIPs. In each plot we see a different stability of the singular strategy $\hat{x}$, depending on the conditions presented in (1.57b) and (1.62). In Fig 1.4a we see a singular strategy which is evolutionary stable since the invasion fitness is negative both above and below a vertical line intersecting the singular strategy, i.e., mutant strategies $y < \hat{x}$ and $y > \hat{x}$ are unable to invade. In the PIP the vertical line through the singular strategy lies completely in regions marked by "-" . The singular strategy is also convergence stable because nearby monomorphic populations can be invaded by mutants nearer the singular strategy. This ensures the singular strategy is an attractor under small mutational steps. In the PIP this is depicted by the horizontal line through $\hat{x}$ on the mutant axis lying completely inside a "+" region. We see that the singular strategy depicted in Fig 1.4a is therefore a continuously stable strategy and is a possible endpoint of evolutionary change. Fig 1.4b provides an illustration of a PIP where the singular strategy is evolutionary stable but is not convergence stable. Such strategies should be rarely observed in nature because there is no means for the population to converge to them. Monomorphic populations should evolve away from a strategy that is neither evolutionary stable nor convergence stable, as shown in Fig 1.4c. In Fig 1.4d we see a singular strategy that leads to branching. In this case, the population evolves towards the singular strategy, but once realised, the singular strategy can be invaded on either side, i.e., $W(y, \hat{x}) > 0$ for both $y < \hat{x}$ and $y > \hat{x}$. We see this in the PIP as both the horizontal and vertical lines through the singular strategy lie in "+" regions.

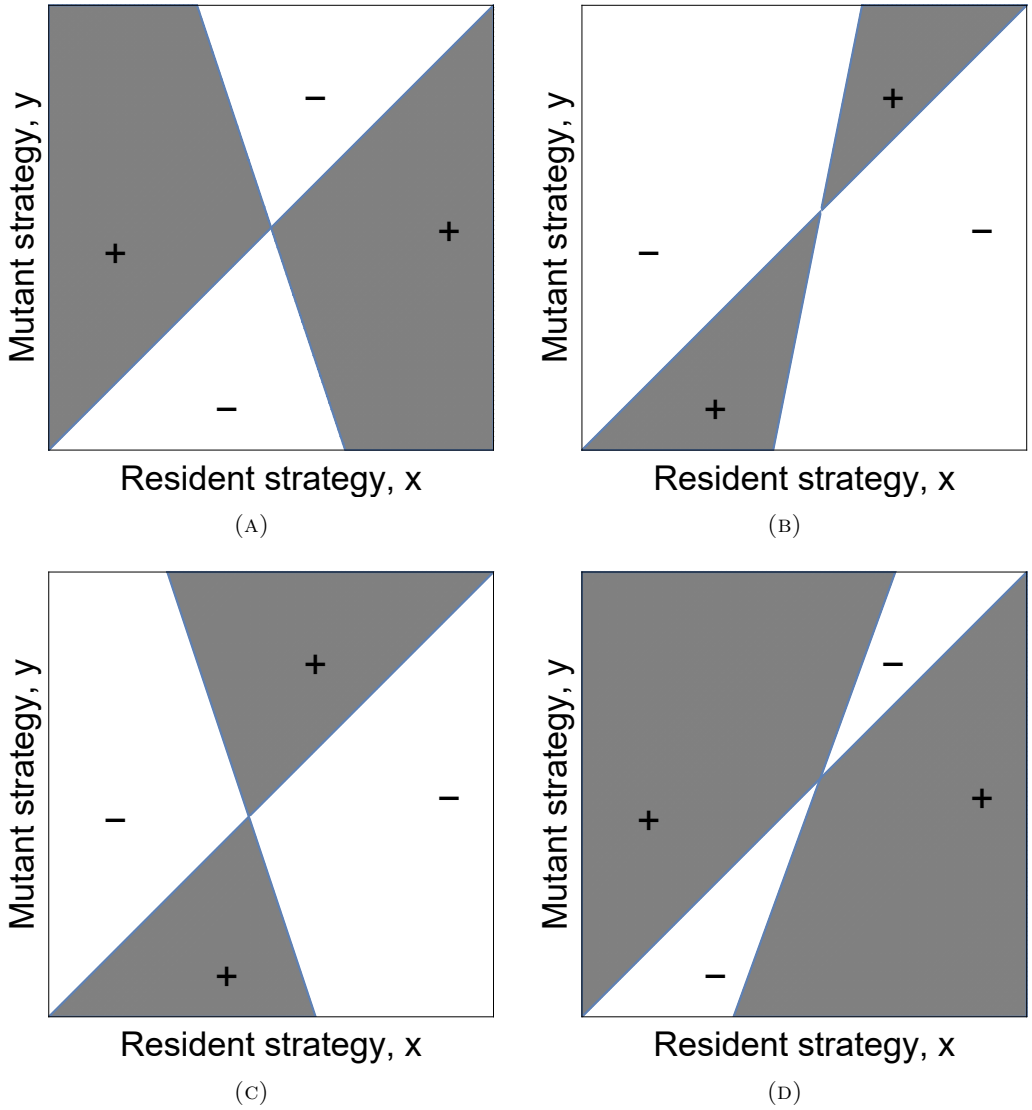


FIGURE 1.4: Pairwise invasibility plots (PIPs) showing the four combinations of evolutionary stability for a singular strategy, as determined by conditions (1.57b) (evolutionary stability) and (1.62) (convergence stability). (A): Evolutionary and convergence stable. Singular strategies can be attained by gradual strategy substitution events, and once reached are resistant to invasion. (B): Evolutionary stable but not convergence stable. (C): Neither evolutionary stable nor convergence stable. (D): Convergence stable but not evolutionary stable. Such a singular strategy can result in a population becoming dimorphic; the singular strategy can be attained by a sequence of strategy substitution events, but then mutant strategies both above and below the singular strategy can invade at the same time.

Summary of adaptive dynamics

Adaptive dynamics draws from models of population genetics [109, 110] and evolutionary game theory [111] using models of frequency (or density) dependent fitness to describe the evolution of strategies. Early incarnations of the theory tend to almost exclusively focus on asexual populations with rare mutations arising from a birth death process [76, 99–101, 110] and a continuous quantitative trait that is passed down through discrete generations [106].

There is a growing recognition for a wider applicability of adaptive dynamics. Of particular interest in this thesis, the assumption of slow evolutionary change in strategy has been relaxed and the theory has been used to model the change in cultural trends [112, 113], cooperative behaviours [57, 60, 81, 114] and kleptoparatisism [103]. For example, in Dobeli [112] p.20, models are presented to explain the diversification of languages and religions. Mutant strategies emerge through errors in cultural transmission rather than through birth death processes. Importantly, the same dynamical possibilities are recovered regardless of the timescale assumptions.

Later in the research of this thesis, we encounter situations where it is more appropriate to consider a continuum of strategy values. For specific examples, consider cooperative investment into some resource, or anti-predation defences. The archetypal games of evolutionary game theory, such as the snowdrift game, can suitably recast into continuous games [57, 60, 81, 112]. In these continuous game analogues, by considering the fate of rare mutant strategies, the framework of adaptive dynamics is used to determine the evolution of strategies that otherwise would not occur.

1.4 Strategies in groups of prey

The goal in this thesis is to predict the behaviours of grouping individuals subject to predation. To do this, the mathematical techniques introduced in sections 1.1-1.3 are utilised. In this section the more specific biological problem is introduced. The description of the biological problem indicates which mathematical framework is appropriate. One can then proceed as follows: derive a description of the biology and specify fitness or at least some quantity proportional to fitness. Having specified the biological problem, one of the mathematical techniques can be applied to predict the changes in behaviours of individuals, which are governed by fitness.

Here we present the main theoretical arguments which explain how by forming groups, individuals are subject to fitness costs and benefits. We consider the general costs and benefits of grouping, and more specifically consider how grouping can provide an individual fitness benefits through reduced predation risk. The most prominent theories which seek to explain how grouping provides protection from predators are the "dilution effect" [31], "encounter dilution effects" [18, 115], "vigilance effects" (or many eyes hypothesis) [49, 116],"confusion effects" [117], and the "selfish herd" hypothesis [118]. The fundamental assumptions and hypotheses of each effect are introduced in this section. An introduction on how the anti-predation effects can operate interdependently is also given. This last issue is of interest in this thesis since the interaction between specific anti-predation effects can select for different anti-predation behaviours in groups, compared to the selective pressures of only one anti-predation effect in isolation. The chapter ends by introducing some relevant previous mathematical models which apply some of the mathematical tools introduced in sections 1.1-1.3 to the explicit problem of anti-predator strategies in prey groups.

1.4.1 The costs and benefits of grouping

Although there are many proposed advantages of living in a group such as improved foraging efficiency [4, 50, 119], information transfer [4, 10], mating [4], proximity to kin [120], there are also costs. Costs involve kleptoparatisism [4], increased competition for resources [121–123] and agnostic competition [4, 124]. Higher levels of pathogen transmission are also associated with larger

groups, potentially making group members more vulnerable to disease [125] and increased parasitic burdens [126]. Groups are nevertheless ubiquitous throughout animal species, suggesting that the balance of these costs and benefits still favours individuals membership.

Perhaps the most favoured benefit, and of most interest in this thesis, is protection from predators [118, 127–129]. As such the anti-predation benefits of grouping have attracted a great deal of theoretical [32, 49, 62, 130, 131] and observational [16, 18, 132–135] research. The research in the subsequent chapters of this thesis is primarily focused on the behavioural strategies that individuals within groups of prey adopt to balance self-maintenance and predation. Evolutionary game theory and other mathematical models can be used to predict the behaviour of prey groups and have had relative success since the selfish herd hypothesis of Hamilton [118]. Yet there is still much unknown about how selection acts on individuals to develop anti-predation mechanisms that in turn encourage certain behavioural (anti-predator) strategies in prey, if and how these strategies compensate or complement one another. It is therefore often hard to explain observed behaviours with theoretical predictions (see for example [32, 62, 130]).

1.4.2 Dilution of risk

One advantage of grouping for an individual is the "dilution" of its personal risk [115, 118, 131]. Dilution of risk encapsulates three distinct concepts in one [4]: the "dilution effect", "encounter dilution" (predator avoidance) effects and the "abatement" effects. Each separate effect is introduced here.

The dilution effect

The dilution effect in principle states that per-capita predation risk is shared equally amongst group members, meaning larger groups lower the risk for any focal group member [115, 136]. To see this, consider a group that consists of "Q" prey individuals. Each group member has a $1/Q$ chance of being a predator target. Assuming that the predator does not preferentially target specific prey, each group member's risk decreases inversely with group size, $1/Q$ [31, 137]. The $1/Q$ rule is appropriate under three additional conditions. First, the predator must only catch one prey at a time. Second, the probability that a predator makes a successful attack must not be influenced by group size (e.g., by confusion effects as discussed shortly). Finally the predator must not attack larger groups at a higher rate than smaller ones [4].

Many studies show that animals in groups are subject to a lower predation risk than solitary animals. Krause & Godin [138] found that a predatory fish attacked larger shoals more than smaller ones, but despite the over-proportionately high attack rates, individuals in larger shoals were still at lower risk than those in smaller shoals. Cresswell & Quinn [25] found that in larger flocks of the wading bird redshank, *Tringa totanus* the increases in individual dilution of risk more than compensated for any increases in attack rate induced by the larger flock size. Perhaps the most notable support for the dilution effect comes from the research of Foster & Treherne [139] who tested the risk dilution hypothesis on sea-skaters, *Halobates robustus*. These grouping insects' dwell on the surface of water and are predated from below, meaning that vigilance for the predator is ineffective, i.e., an increase in group size does not yield any benefits of increased detection. The

attack rate per group was found to be independent of group size, and the number of attacks per individual decreased with group size as would be predicted by the pure dilution effect.

Encounter-dilution (avoidance) effects

Grouping prey can also benefit from avoidance effects [140–142] (also referred to as encounter-dilution [4]). The idea here is that a predator is less likely to detect a single cluster of individual prey items compared to prey that are spread out as scattered individuals. Predator avoidance is likely to procure grouping individuals benefits when the perceptual range of the predator is low in relation to the movement speeds of the predator and prey [4]. In addition to the avoidance effects that affect the predator's ability to detect groups, grouping can also decrease the probability of running into (encountering) a predator. This latter effect can lengthen the time prey are able to spend foraging.

Avoidance effects have been considered geometrically by Vine [140] and Treisman [141] who developed the idea that a predator's visual acuity can be exploited by prey groups forming compact formations. The idea is that the predator scans its environment with low efficiency and considerable time. As such, the greater the angle of deviation required to detect prey, the more likely the predator is to give up the scan. If animals group, the mean of this angle will be greater, and on average the predator will give up the search more often. Vine [140] hypothesised a "selfish" advantage to each group member to combine into a compact formation, similar to the idea of Hamilton [118] as will be discussed shortly. Treisman [141], with a few modifications and extensions to Vine's model, accounted for predators that repeatedly scan their surroundings. The same conclusion that compact grouping lowers the probability of predator detection was found.

Attack abatement

The term $(Q-1)/Q$ adequately describes individual survivorship over one attack, but of equal importance is the overall rate of attacks on the group over some period of time [32, 130, 143]. If groups of Q individuals are more than Q times likely to be detected (and attacked) than a solitary prey, then the advantage gained through dilution is lost [143]. Dilution and avoidance must be considered in concert. The concept of attack abatement was born out of the theoretical synthesis of Turner & Pitcher [115] who argued that the two mechanisms of dilution and avoidance must be considered together. Here, the idea is to separate a focal group member's overall risk into two additive terms: the probability that the group is detected (avoidance effects), and the conditional probability that given the group is detected, the focal individual is targeted (dilution effects). If the following three conditions hold: the probability of group detection increases less than proportionately with group size; given an attack, the probability that the predator makes a successful capture is independent of group size; and the predator can only capture single prey items, there will be an overall reduction in risk for group members

Support attack abatement has been found for example by Wrona & Dixon [144] who used a regression model to separate and compare the effects of encounter and dilution on observational data of predation of a water insect. Their research showed that larger groups were attacked more than small groups, i.e., the opposite of the avoidance effect, but also that within larger groups,

focal members were less likely to be attacked supporting the dilution effect. The authors concluded that the second effect was dominant and that overall, there was an attack abatement effect.

In essence, the attack rate of the predator specifies when grouping is favourable in terms of dilution risk. If a predator attack rate asymptotes with increasing group size [25] then an individual's per capita risk is lowered by a complement of the dilution effect and predator behaviour. This last observation, also found by Jackson *et al.* [142] shows that the encounter-dilution and attack abatement effects are complex, and depend on linearity or nonlinearity of the relationship between predator detection (and subsequent attack rate) and group size. If attack rate increases precisely at the same rate as group size increases, then a group members per-capita risk remains unchanged. Attack rate may be a nonlinear function of group size [25]. Jackson *et al.* [142], for example, found that single individuals were detected less than groups, but once a group reached a size of greater than two, there was little if any increase in group detection with increases in group size. One obvious result from these findings is that for the specific study, there would be strong advantages in increasing group size far above two individuals since dilution effects would increase without incurring a greater number of attacks. However, if the increase in detectability from moving from a singleton to a group of two individuals caused a proportionate increase in attacks, one could expect individuals to either form large groups or remain solitary. The study did not find that groups of two were twice as likely to be detected than singletons.

Observational studies have led to other questions regarding the theoretical assumptions of risk dilution [127, 136, 145, 146]. Critically, none of the three concepts of risk dilution consider possible variation between individuals within a group. For example, age can influence an individual's risk dilution [4] as shown by Sorato *et al.* [147] who found that groups containing young were preferentially attacked relative to groups that did not. Other factors such as differentiated behaviours between group members also complicates the predictions of risk dilution. For example, group members may adopt bold or aggressive responses to predation [2, 24, 28, 148] but such behaviours are not assumed to affect an individual's dilution risk. The effects of spatial position [149–151], variation in conspicuousness between group members [152], or the behaviour of the predator [138, 153] are not considered in risk dilution. Group position has attracted much theory and empirical evidence which we move on to next.

1.4.3 The selfish herd

In a highly commented on paper, Hamilton [118] addressed the issue of spatial positioning within a group, whereby an individual could reduce its per-capita predation risk by moving towards others. This idea is known as the selfish herd hypothesis. In theory, the selfish herd argument is that two group members have relative risks of predation that can be determined by the respective minimum areas around each individual and its nearest neighbour: the individual's so-called domain of danger. Predator attacks are launched from random locations. The predator attacks the nearest group member to it. Thus any group position has an equiprobable chance of being attacked. If the predator launches its attack within a focal group members domain of danger, then that group member is killed. Hence, an individual's predation risk is proportional to individual's domain of danger. To see this more clearly, original idea of Hamilton [118] is described in a little more detail.

Hamilton envisaged a hypothetical lily pond in which frogs occupy the periphery. A snake, disguised from view attacks a frog. The question that Hamilton posed is: how should a focal frog behave to minimise its risk? His answer was that each frog should attempt to not be nearest to the snake, i.e., to minimise its domain of danger (originally measured as half the distance between the focal frog and its neighbours on either side). The focal frog can reduce its domain of danger if and only if there exists a gap between two neighbouring frogs which is narrower than the distance between the focal frog and either of its two adjacent neighbours. Simple illustrations can explain the domain of danger concept Fig 1.5.

Put simply, the selfish herd hypothesis argues that by moving towards a neighbour, an individual is subject to less predation risk, and that aggregation is then a natural consequence of the danger minimising behaviour of each individual [154–156]. One should picture an endless cycle of movement towards the origin of the group as prey continually attempting to place their neighbours between themselves and the peripheral predator. In this case, differential predation risk occurs due to variation in spacing within the group: one individual's increases in safety must come at the expense of other group members [149, 157]

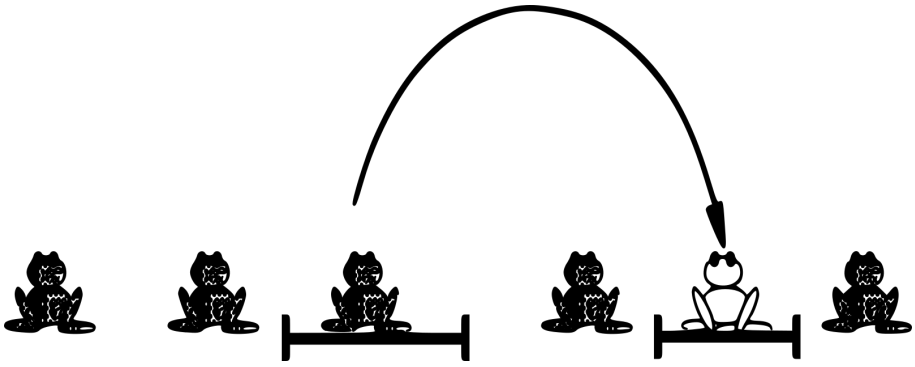


FIGURE 1.5: By moving to the final position shown by the white frog, the third frog from the left reduces its domain of danger. Image re-drawn from Haynes & Moore-Crawford [158]

The selfish herd hypothesis brings up the question of how predators search for, and attack prey. An important assumption made by Hamilton is that the predator appears from a random location and attacks the nearest group member. In the case of marginal predation, i.e., the predator appears from a random location outside the group, this assumption is questionable. Evidence for marginal predation is found in stationary groups (see [4]) where animals on the groups edge are generally surrounded by fewer individuals, yet other studies show that central individuals may be at more risk [159]. An individual's spatial risk thus depends on what type of predator is attacking [4, 151, 156]. Brunton [160] for example, showed that central nests in least tern, *Sterna antillarum* are at relatively more risk than peripheral nests. The discrepancy is due to the relative frequencies of predators with alternative strategies: most nest predation was due to black-crowned herons *Nycticorax nycticorax*, which only attack central nests, and less nest predation was due to American crows *Corvus brachyrhnchos* which targeted edge nests. Importantly, if the authors had only considered attacks from American crows then they would have found support for the selfish herd argument as peripheral nests would have been predated more. This example clearly

illustrates how the efficacy of an anti-predation effect and thereby the behaviour (or strategy) of the prey depends on the behaviour (strategy) of the predator.

Another issue taken to Hamilton's hypothesis is that it's hard to distinguish whether marginal predation occurs because the predator is focusing its efforts on peripheral group members, or because peripheral group members share some specific phenotype which the predator finds attractive (e.g., juveniles, sex, less dominant individuals). Generally, studies that support marginal predation cannot control for these confounding factors [161]. The idea of marginal predation also raises the question addressed by Pulliam [49], of why peripheral individuals should remain part of the group. If each individual acts in this selfish way, then any relative advantage is lost and the collective group's welfare is lower [49, 141]. If predation risk is greater at the periphery of a group, then individuals at the groups edge would be better fleeing and that the group would disband.

Observing the selfish herd in real species systems is difficult, but there is observational evidence supporting the theory [162]. For example, Viscido & Wethey [163] found that in fiddle crabs, *Uca pugilator* subjected to predation, consistent with the selfish herd hypothesis individuals minimised their domains of danger upon attack. Moreover, Kimbell & Morrel [157] provide evidence that the shoaling fish guppies *Poecilia reticulata* make complex aggregation decisions based on the number of and position of neighbours. The authors conclude that when threatened by a predator the fish form more compact and cohesive shoals. More recently the aid of computer simulations have allowed for models of digital evolution based on individual movement rules within groups [155, 156, 164–167]. In such models distinct behavioural strategies can evolve through simulated prey responding to a predator. For example Wood & Ackland [155], using an individual based model with a game theoretic treatment, found that simulated prey formed compact groups, interpreted as Nash equilibrium states, i.e., simulated states where no prey could benefit from altering behaviour. The compact groups evolved by each simulated prey minimising its own domain of danger, as consistent with the selfish herd hypothesis. More recently and in a similar vein, Algar *et al.* [166] showed that simulated prey which explicitly seek to minimise their domains of danger form small clusters, supporting Hamilton's notion that once nearest neighbours are reached, individuals further reduce domains of danger by finding minimal "corners" [118] p.304.

1.4.4 Confusion Effect

The confusion effect succinctly put by Krause & Ruxton [4] p.19 "... describes the reduced attack-to-kill ratio experienced by a predator resulting from an inability to single out and attack individual prey" and can provide antipredation advantages of grouping [117, 168–170]. Groups with multiple moving individuals prove harder to process by the predator due to cognitive and sensory limitations [171, 172]. This results in the predator having difficulties in both selecting and targeting any one group member during an attack [170]. Additionally, because multiple moving targets are harder for the predator to track, changing target during an attack is also more difficult. The confusion effect increases probability that the entire group eludes the predator. Confusion effects can be enhanced through the addition of specialist prey behaviours, for example coordinated motion [173, 174] or swarming behaviour [175]. Therefore, unlike the dilution effect which retains a constant selective pressure with time, the confusion effect additionally selects for advantageous individual strategies within a group. For example Olson *et al.* [175], using a computer simulation model

found that the strategy of swarming by prey is sufficient to exploit the confusion effect. Similar prey reactions have been observed in several species of birds [174] and fish [148, 176]. Picther & Parrish [177] suggested that coordinated movement rules, and in particular evasive manoeuvres, could enhance the confusion effect. Research has considered the adaptive benefit of a phenomena known as "flash-expansion" in which each group member moves away each other [5, 178]. Storms *et al.* [5] observed that starlings frequently adopted flash expansions as a defence mechanism when attacked by a peregrine falcon, at high speed.

Research conducted by Tosh *et al* [172], Ioannou *et al.* [170] and co-authors used artificial models of neural networks to investigate confusion effects. The authors found a U-shaped relationship between group size and confusion, with mid-sized groups benefiting most from confusion effects. This finding was attributed to the contrast between the object (group) and background being at a maximum for intermediate group sizes. A second finding of this research which contrasts with other experimental evidence on the confusion effect (see [179] for example) was that compact or denser groups provided more reliable predator targets. In a similar vein Ruxton *et al.* [171] used an interactive computer programme in which human subjects acting as surrogate predators had to capture virtual prey dispersing on a screen. The authors investigated five predictions relating to the confusion effect and found the following results: (i): the authors found a confusion effect regardless of any specialist prey behaviours, (ii): prey that stood out were found to be captured more than other prey items, (iii): neither group homogeneity nor heterogeneity significantly affected on the probability that a group member was captured. In other words, it made no difference whether all group members looked the same or if there were physical differences in prey items, (iv): solitary individuals were found to be more vulnerable (more likely to be captured) than grouping individuals, and (v): there was no evidence that compact groups enhance the confusion effect, although the findings neither showed that compact groups are more likely to result in a prey capture.

Perhaps the archetypal example of the confusion effect within experimental studies is the fish shoal [4, 11, 179, 180] (although coordination within starling murmuration's come close) [174, 181]. In relation to fish, Landeau & Terborgh [117] showed that bass *Micropterus salmoides* attacking minnows *Hybognathus muchalis* only capture prey 11 percent of the time when the prey shoal consists of 15 individuals. Compare this to 100 percent capture rate on singleton minnow. For a completely different case consider that leopard geckos *Eublepharis macularius* show more "fixations" on multiple prey items, i.e., constant turning of the head, when presented with a group of mealworms *Tenebrio molitor* compared to a solitary prey item [182].

The anti-predator fitness benefits grouping individuals receive through the effects of confusion depend on the type of predator. Jeschke & Tollrian [183] analysed 28 sets of data and found that in all cases tactile predators suffered confusion effects, while visual predators seemed susceptible to confusion effects mainly if their prey was highly agile. Moreover, Cresswell & Quinn [151] argue that there will be stronger selection for flocks of redshank to adopt behaviours which yield confusion benefits for the group, such as reducing within group spacing or coordinated flight, when confronted with peregrine falcons *Falco peregrinus* as opposed to sparrowhawks *Accipiter nisus*. The authors concluded that because the peregrine is a "pursuit" predator, it is respectively more susceptible to confusion effects than the sparrowhawk which is an "ambush" predator. Other notable research of Ioannou *et al.* [179] illustrates the complexity of confusion effects. In this study, the effects of

confusion were found to be sensitive to the stage predator attack. The predator's detection of prey increased with group density. But as the predator drew nearer the prey group, its ability to target one individual decreased. Nearer the prey, fewer individuals within the predators visual range were argued to produce considerable confusion effects. Moreover, Ioannou *et al.* [179] demonstrate a interdependency between confusion and abatement effects as introduced in section 1.4.2: When the predator is at a distance, denser groups are attacked relatively more, but when the predator is close denser groups reduce its targeting success.

A natural extension of the confusion effect is the so-called "oddity effect". Because the confusion effect is due to a predator's inability to switch focus among prey group members, it is considered that when all group members are phenotypically similar the confusion effect is more inhibitive to the predator. Whereas if some group members phenotypically distinctive then the predator finds it easier to focus its attention on the distinctive individuals. This consequently improves the predator's attack success rate. Most of the basis on the oddity effect comes from observational studies: Landeau & Terborgh [117] studied bass fish feeding on minnow shoals and the primary study of Ohguchi [184] investigated sticklebacks predating water fleas. In more detail, Landeau & Terborgh [117] demonstrated that a single odd individual lowered the effect of confusion in small groups. The conclusion of their study was that the effect of oddity in reducing the confusion effect only occurs in smaller groups. Interestingly, the research of Ruxton *et al.* [171] found no support for the assertion that increased vulnerability was confined to small groups. Using an artificial experiment in which humans targeted digital boids on a computer, the authors did find that the oddity effect disappears when the phenotype of the target is common in the group.

1.4.5 Group Vigilance

The "group vigilance effect" (also many eyes hypothesis) predicts that two benefits arise from grouping. First, any focal group member can reduce its personal vigilance without any increase in overall predation risk. Secondly, so long as vigilance is equitably shared amongst the group, all group members receive increased fitness through being able to spend more time foraging. The upshot is that when vigilance effects are considered in isolation (without factors like competition, other anti-predator effects or predator behaviour) there will be selective advantages for grouping, at least up to a certain group size. Over the last half century the group vigilance effect has been the most prominent explanation for grouping, attracting significant amounts of observational [18, 128, 132, 149, 185, 186], and theoretical [31, 32, 46, 63] research. As such, this section will give a more comprehensive review of group vigilance. In the following subsections, some of the confounding factors noted by Elgar [127] are addressed.

Perfect collective detection

This assumption is that if one group member detects the predator, then the whole group is alerted and have an equal probability of escape. This assumption surmounts to an individual's personal risk being independent of its vigilance level. The assumption does not consider that the overall probability for the group to detect a predator may be compromised by the lower personal vigilance rates of its members. If some threshold number of individuals are required to detect the predator,

or individual detection is not transmitted throughout the group, then each individual may not gain increased safety.

Monitoring

Collective detection (as explained above) reflects the notion that "detectors" warn other group members about an attack. As such, group vigilance models generally assume that either prey individuals monitor one another, or that detectors give some form of signal regarding an attack. The idea of monitoring stemmed from the game theoretical concept of "conditional vigilance" [49] (see also section 1.5.1), i.e., that a group member will be vigilant as long as other members remain vigilant. However, there is no clear evidence for monitoring, and the widely reported group size effects [127, 136] show that although individuals respond to conspecifics [187], this does not necessarily suggest that individuals monitor conspecifics [116, 188].

Food density

It is common for patches of richer food densities attract and sustain larger groups [149, 185, 189, 190]. As the quality or density of foraging resources increase, vigilance of group members is argued to decrease. There are two lines of reasoning behind this argument. First, "information theory" suggests that in regimes of poor food density individuals rely more upon social vigilance to glean information of local resources from neighbour group members [124, 191]. The producer scrounger model provides the theoretical basis for this prediction: foragers either produce to locate resources on their own or scrounge to locate resources found by producers. Opportunities for scrounging behaviour decrease with food density, which in turn decreases social vigilance to locate food found by others. Thus vigilance should decrease at a steeper rate than expected with group size alone [191].

Alternatively state dependent predator models [32, 46, 50] predict a correlation between individual vigilance and food density as individuals face a trade-off in survival through a reduction in predation risk, and an increase in food intake. Implicitly, an increase in food density is supposed to result in an increase in food intake rate, which has a knock on effect to an individual's vigilance [46]. With time constraints it can be worthwhile for an individual to increase its foraging rate at the expense of a decrease in its vigilance [32, 192]. This prediction is based on modelling the probability of starvation as a function of vigilance, for different food intake rates [32, 124]. Under time constraints, and when food intake rate is high, a decrease in vigilance has a more pronounced decrease in the probability of starving compared to when food intake is low. For example, if energy reserves are low, then an animal's intake should be high. While the animals energetic budget is not balanced, altering the intake rate (through vigilance) will have a profound effect on expected future fitness [32].

As an example of the effects of food density on vigilance, Barnard [193] found that house sparrows *Passer domesticus* are less vigilant in times of increased food density due to increased foraging rate. Larger groups tended to form in areas of higher seed density, and without scrutiny one could assume a causal relationship between group size and vigilance. In fact "seed density" modulates

both the vigilance rate and group size of the birds [193]. For more examples see Beauchamp [122], Fritz *et al.* [190] or Pays *et al.* [189].

Competition

The relationship between food intake and food density is complicated by competition between foragers. Intra-group competition can manifest itself through "agnostic" competition where individuals perform aggressive displays such as fighting over food [194]. This type of competitive behaviour typically occurs when food density is high and there are no time constraints on foraging [124]. Vigilance is expected to increase with agnostic competition because foragers need to monitor aggressive neighbours and a positive correlation between food density and vigilance is expected.

Alternatively "scramble" competition generally occurs at low food density when individuals compete for limited resources. The effect of this is that foraging becomes more valuable compared to being vigilant. As food density increases, an individual's need to feed decreases, and it can devote more time to vigilance. Examples of scramble competition are summarised by Elgar [127], and observed in Beauchamp [122] and Schutz & Schulze [195]. Also of interest is how scramble competition from grouping alters an individual's functional response, as theoretically explored by Fryxell *et al.* [196].

Posture and morphology

It is often proposed that vigilance and feeding are mutually exclusive activities [18, 127, 191]. If an increase in food density results in an increase in feeding, the obvious outcome is a decrease in vigilance. Many wading birds demonstrate this trade-off between scanning and feeding [124, 195]. If individuals are able to feed head up and maintain some degree of vigilance whilst foraging one could expect a positive correlation between vigilance and food density. As food density increases so does foraging time and time able to scan for predators [189]. A comprehensive review conducted by Beauchamp [124] however concludes that posture plays an insignificant effect in the relationship between group size and vigilance.

Functional response

Functional responses describe the relationship between food density and intake rates. In numerous species food intake increases at a decreasing rate with food density, and eventually plateaus (no further intake is possible with any increase in food density), i.e., a type II functional response. When food density is high, foragers reach this plateau sooner, and intake rate becomes largely independent of food density. Vigilance in this case, should not vary. The role that functional response can have on vigilance is observationally demonstrated in Goss-Custard [197], Kvist & Lindstrom [198].

State

One could predict that vigilance should increase with food density given that any genuine threat of starvation decreases when there is more food [192]. Thus a positive correlation between food density and vigilance would be expected through the indirect effect of an individuals state. To maximise

ones fitness an animal should prioritise safety (through increasing vigilance) when they are satiated and prioritise feeding (reducing vigilance) otherwise [32]. When food density is high an individual should be able to accumulate reserves more quickly and thus reach a state of satiation sooner. Then the individual behaves more vigilantly. Thus as food density increases so should vigilance. For examples of how state can affect vigilance see the research of Bednekoff & Woolfenden [29], Clutton-Brock *et al.* [8] Ebensperger & Hurtado [199] or Ydenberg *et al.* [200].

1.4.6 How anti-predation effects interact to affect prey behaviour

In this thesis, the focus is on how the threat of predation gives rise to the strategic behaviours of individual group members. It is therefore important to identify how each anti-predation effect can select for particular anti-predator behaviours in prey, and how potentially the anti-predations effects interact when producing such selective forces.

Most studies tend to examine one of the three principle anti-predation effects in isolation [49, 188] (but see [32, 175]) because it is hard to control for each independent effect. Safety in numbers in animals for example, is often discussed with respect to detection and dilution operating as separate and discrete effects. Equal dilution of risk among group members depends on the explicit assumption of perfect collective detection: if some group members (the non-detectors) do not have the same information regarding an oncoming predator, then they may be more vulnerable to predator targeting. The approach of this thesis is to consider that detection and dilution are instead intertwined. Collective detection is in reality, always imperfect, and animals that detect predators imperfectly face considerable disadvantages compared to animals that detect directly [31, 132, 145, 201]. In addition to the prey behaviour, one must consider that the prey targeting behaviour of the predator also affects the per-capita differences in risk between detectors and non-detectors. From this perspective, one can understand the behaviours of prey as acting to avoid being a predator target, and if targeted, acting to minimise the probability of capture. As an example, Lima & Zoller [116] find that in house sparrows, *Passer domesticus*, vigilant (detector) birds initiate faster responses when threatened with a predator than non-vigilant (non-detectors). The upshot being vigilant birds were less likely to be predator targets through risk dilution and that there are selective advantages for adopting a vigilant behaviour. Alternatively, consider sentinel behaviour in meerkats *Suricata suricatta* [8]. Vigilant sentinels generate a group wide benefit through increasing safety for all (a collective vigilance effect), yet at the same time receive a private benefit of first response (a personal dilution benefit). In the extreme, one vigilant group member could suffice for entire group safety, and one could expect for a dominant behaviour of individuals being non-vigilant, especially in cases where vigilance incurs foraging costs.

To demonstrate the importance of the magnitude of each anti-predator effect, consider the bird redshank, *Tringa totanus*. Numerous observational studies document how anti-predator behaviours within flocks of redshank are affected by the anti-predation effects of sections 1.4.2-1.4.5 [1, 18, 25, 151, 162, 201–207]. These studies are of interest because they show that the predator hunting behaviour can modulate the efficacy of anti-predator effects. For example Quinn & Cresswell [201] find that when attacked by a sparrowhawk, vigilant redshank initiate earlier responses to the predator and gain a preferable dilution of risk profile, compared to non-vigilant birds. In this case, by adopting a vigilant behaviour a focal individual group member provides (i) fitness benefits for

all group members through increasing the overall probability that the flock detects and "evades" attack, and (ii) receives a personal fitness benefit through its relative smaller probability of being targeted given an attack. In contrast to the findings of Quinn & Cresswell [201], when redshank are attacked by a peregrine, confusion of the predator [151] plays a more important role for prey survivorship. In this case, it is beneficial for individual prey to form larger flocks [151] and to minimise their domains of danger [162]. Although Hamitlon's hypothesis of section 1.4.3 cannot be ruled out, Cresswell & Quinn [162] argue that tightly spaced flocks confuse the predator, and are consequently attacked at a lower rate. To predict the behaviours within the prey group, one must consider the magnitude of sparrowhawk versus peregrine attacks, how each type of attack gives rise to individual survivorship through an interplay of the anti-predation effects, and how certain prey behaviours can operate advantageously or disadvantageously through each anti-predation effect.

In this research, the methodology is to consider the importance of how the anti-predation mechanisms can interact. The interaction of anti-predation effects can lead to different levels of individual predation risk and thereby select for certain behavioural strategies in prey. Importantly, these behaviours would not be expected when considering each anti-predation effect in isolation. Thus, behaviours originating from groups such as collective motion [175], vigilance [18], and herding [118, 166] depend on predation pressures, which in turn are specified by a potential interdependence of the effects introduced in sections 1.4.2-1.4.5.

1.4.7 How can behaviour spread

Behavioural strategies can either be assumed to be hard wired into that individual from birth, or can be labile and likely to change with environmental factors and with the frequencies of other behaviours. In groups of prey, behaviours can change according to the type of predator, and with the behaviours of other group members. Here we introduce the concepts of learning and imitation which are often used to model the spread of behaviour [40, 60, 93]. For the game theoretic analysis of subsequent chapters, the choice of mechanism is not fundamental.

One possibility is that an evolutionary learning process gives individuals some predisposed set of behaviours [21, 208]. In this case it is not the behaviour itself that is under selective pressure to be adaptive, but the psychological mechanisms underlying behavioural flexibility. As such, one would expect animals to evolve a wide ranging set of psychological mechanisms that allow good performance over a range of environments, as documented in experimental [209–211] and theoretical studies [212–216].

Alternatively, associative learning can allow animals adaptive responses to novel environments [217]. On this line, through previous mistakes "the cost of naivety" may determine an individual's strategic choices [218]. As example, processes of trial and error allow for animals to change behaviours in the presence of novel predators [219]. Hence, new behaviours (strategies) can arise by individuals applying previous experience when responding to novel environments. In this sense an individual's behaviour may have the ability to change across a range of environmental contexts [220, 221], and one can assume that individuals exhibit a sufficient degree of flexibility to chose appropriate behaviours for given situations.

In the simplest form behaviours may spread through pure imitation [60, 61, 222] (with no special learning process). For imitative behavioural strategies, individuals must be able to in some way

assess their fitness relative to others: if the behaviour of some observed "role model" appears to be more conducive to survival, an individual chooses the role model's behaviour. On the other hand, imitative behaviours may simply be "contagious" if it only takes the sight of others performing the behaviour to release the behaviour in the focal individual [217]. Here, the relevance to this thesis is that the imitation of fitter strategies argument provides a simple yet plausible theoretical assumption for how strategies spread within a population. To restrict attention to anti-predator behaviours, consider examples of birds imitating the flight responses of others in a flock, upon predator detection [116], individuals adopting a vigilant position after inspecting the vigilance of neighbouring individuals [223], or individuals imitating alarm calls [152, 224].

1.5 The role of mathematical models

So far, mathematical techniques which use fitness as a measure to model the success or change in strategies have been introduced in sections 1.1-1.3. Through sections 1.4.1-1.4.5, the prominent anti-predation effects which modulate individual mortality risk (usually taken as proportional to fitness) within groups of prey and which depend on the behaviours of group members, have been explained. One is now in a position to model the strategy dynamics in groups of prey. By specifying some appropriate anti-predator behaviour (strategy) or set of behaviours within a group, interpreting how behaviour affects an fitness (payoff) through the anti-predation effects, and applying an appropriate mathematical framework, one can predict how behaviours or the frequencies of behaviours emanating from the group change. This procedure is applied in the subsequent research chapters of this thesis.

Often theoretical models assume that strategy interactions are pairwise (two player games), and payoffs a linear combination of constant benefit terms multiplied by frequencies of existing strategies in the population [92]. However, simple two player games cannot capture common synergistic benefits that occur within groups. In groups there are multiple interacting players. The underlying game's payoffs is most likely nonlinear with the number of players (even if they are reducible to pairwise interactions [47]). When games are reducible to two player games evolutionary game theory provides a powerful yet simple framework to explore strategic interactions. To some extent, this comes at the expense of how realistically the biological problem is modelled. More specific to this research, in groups of prey, the payoff that an individual receives from adopting a certain anti-predator strategy may be nonlinear in the number of other individuals playing the anti-predator strategy, for example if collective detection [32, 116], or collective motion [169, 170] are strategies to evade a predator. When strategy interactions cannot be reduced to two player game, evolutionary game theory dynamics become complex (see for example the analytically inaccessible nonlinear public goods game [225]. In this thesis, a key assumption is that a focal group member's strategy (discrete or continuous) affects its fitness through either one or a combination of the anti-predation effects of 1.4.2-1.4.5. Other strategies within the group, at the same time influence the focal group member's survivorship. Fitness is therefore frequency dependent within the group and nonlinear in the strategies within the group.

Previous mathematical models considering groups often use the term "cooperative" (see [32, 59, 63, 226] for example) to describe behaviours which benefit all group members irrespective of

the fitness effects of the performer. Vigilance against predators, for example, is one of the simplest and widely reported examples of the cooperative behaviour within groups [18, 32, 59]. Later in this thesis, we adopt this convention in a specific context where a focal individual's behaviour brings fitness benefits to all other group members. In the modelling sense, whether an individual chooses a cooperative behaviour to intentionally benefit other group members, or simply to benefit itself, is irrelevant. All that matters is the outcome of the game. This final subsection provides a brief introduction to some notable previous mathematical models, their terminology, and their key results.

1.5.1 Pulliam's ESS model

Pulliam [49] presented a model which was subsequently extended by Pulliam *el al.* [130] to predict the ESS vigilance rate in groups. The model directly associates predator survivorship to fitness. Pulliam's idea was that n group members face some time period, T, in which a predator attacks. If no group member detects the predator within this time, the predator attack is successful, and all group members face a $1/n$ chance (i.e., pure risk dilution) of being killed (and a $(n-1)/n$ probability of not being killed). If any group member detects the predator, the entire group escapes with probability $1-b>0$. Thus, group members are less likely to be killed if a group member detects: $(n-b)/n > (n-1)/n$.

Group members are assumed to have random scan rates, λ per unit time, which follow a negative exponential distribution, and each scan lasts for τ seconds. The probability that the predator is not detected by a solitary group member is

$$\frac{e^{-\lambda T}}{\lambda\tau+1}. \tag{1.63}$$

Given that all group members scan independently of one another, the probability that no individual detects the predator is simply the product of all the probabilities, $i=1,...,n$, i.e., that each individual with scan rate λ_i does not detect. The model finds the probability that any particular bird is not killed during a specified time period in which there are αT_{G_i} attacks. Maximisation of this probability leads to an optimal scan rate (strategy) $\hat{\lambda}$. When all group members choose $\hat{\lambda}$, because each individual maximises its probability of survival, the outcome is termed the so-called "cooperative equilibrium". An important result of the model is that the scan rate which characterises the cooperative equilibrium is not an ESS strategy, but prone to deviant group members decreasing their scan rates (and increasing their probability of survival). The deviant scan rate is an ESS since when adopted by all, no group member can increase its survivorship by altering scan rate.

A comparison of observational data to the predicted ESS scan rate indicates that the birds yellow-eyed juncos *Junco phaeonotus*, actually choose strategies that are closer to the cooperative equilibrium. Pulliam *et al.* [130] conclude that the reason for this finding is due to "conditional vigilance" as defined in section 1.4.5. An alternative explanation which partly motivates the subsequent research of this thesis is as follows: scan rates are higher than expected because by personally detecting the predator, a group member is at less risk from the predator and receives a fitness benefit of a higher probability of survival [145].

1.5.2 The model of McNamara & Houston [32]

Using a different approach, McNamara & Houston [32] use a life history model to predict ESS levels of vigilance within groups. The authors investigate group foraging and vigilance behaviours in grouping individuals subject to more than one of the anti-predation effects of sections 1.4.2-1.4.5. The key finding is an ESS level of individual vigilance as a function of individual state and of group size. Here, a brief review of the model is given.

Consider groups of n individuals that are subject to a trade-off between vigilance and feeding. In the model, predation risk is modulated by the many eyes effect and the dilution effect. Individuals choose some foraging investment strategy u, characterised by a state variable x which reflects factors such as satiation. Fitness is defined through "expected future reproductive success" (EFRS), and is an increasing function of x. Feeding increases an individual's state x through increased energetic intake, but also increases the probability that the individual is predated. The probability that the individual is killed depends on the other group members. Thus an ESS reflects the behavioural strategies of all group members.

The following state dependent variables are introduced: $R(x,t)$ defines an individual's EFRS after foraging has ended, $U(x,t)$ the EFRS while engaging in foraging, $V(x,t)$ the increase per unit time in state, $K(x,t)$ the cost for partial termination of foraging, $h(u)$ the associated factor for surviving attack and $\Delta(x,t)$ as a time constraint factor. A focal group members EFRS is affected by the following predation variables: α, the attack rate, $g(u)$ the probability individual fails to detect, A_n the probability of death if no individual detects (a simple dilution effect $A_n = 1/n$), B_n the probability of death when at least one (but not focal) individual detects, C_n the probability of death when the focal detects and $D_v(u)$ the probability of death when the focal forages at rate u and the rest of the group forage at rate v. Finally, the authors use variables γ; the rate of energy gain while feeding, and $q(t)$; the rate of premature termination of foraging after time t, to reflect a focal individual's foraging behaviour.

A focal individual playing strategy u in a group where all other members choose strategy v, experiences a probability of death given by

$$D_v(u) = \underbrace{A_n g(v)^{n-1} g(u)}_{\text{dilution alone}} + \underbrace{B_n\big(1 - g(v)^{n-1}\big) g(u)}_{\text{group but not individual detection}} + \underbrace{C_n\big(1 - g(u)\big)}_{\text{individual detection}} \tag{1.64}$$

Despite being complicated, equation (1.64) simply describes a focal individual's mortality risk when it spends a proportion u of its time feeding, while the rest of the group choose to feed at a rate given by v. Each term of equation (1.64) should be interpreted as follows:

dilution alone: This term gives the probability that a focal group member is killed, given that neither the focal member, nor any other group member detects, i.e., the product of three independent probabilities. In this case, the dilution effect A_n, which equitably shares risk between all group members, is multiplied through $g(v)^{n-1}$; the probability that no other group member detects, given that all other group members use foraging strategy v, and by $g(u)$; the probability that the focal individual also fails to detect, given its foraging strategy u.

group but not individual detection: This term gives the probability that the focal is killed given that it fails to detect, but at least one other group member detects. The product of the probability

that at least one other group member (excluding the focal) detects; $1-g(v)^{n-1}$, the probability that the focal individual fails to detect; $g(u)$, and the probability that an individual dies, given at least one group member detects; B_n.

individual detection: The probability that the focal detects is $1-g(u)$, in which case C_n gives the probability of being killed. Hence, the product $C_n(1-g(u))$ gives the focal individual's mortality risk.

To find an ESS level of vigilance in the group, the model considers a mutant strategy that spends a proportion of time u feeding when the rest of the group spend v. For $\delta > 0$ a small time increment and individuals foraging between times t and $t+\delta$ the ERFS for the mutant, $G_v(u)$, is given as

$$G_v(u) = \big\{1 - \big(q + \alpha D_v(u)\big)\delta\big\}\big[U(x,t) + (V(x,t)\gamma u - K(x,t))\delta\big] + q(t)\delta R(x,t). \tag{1.65}$$

The mutant's fitness (1.65) is explained as follows: the mutant survives to time $t+\delta$ with probability $1-(q+\alpha D_V(u))\delta$, at which point its state is $x+\gamma u\delta$ and its EFRS is $U(x,t)+\big(V(x,t)\gamma u-K(x,t)\big)\delta$. However, with a probability $q(t)\delta R(x,t)$ the mutant prematurely terminates foraging. The best strategy for the mutant is a simple maximisation problem for $G_v(u)$. One can see that this is equivalent to maximising $F_v(u)$ for all $u \in [0,1]$ where

$$F_v(u) = V(x,t)\gamma u - \alpha D_v(u)U(x,t) \tag{1.66}$$

since $V(x,t)\gamma u$ is the increase in EFRS per unit time resulting from feeding, $\alpha D_v(u)$ is the risk of death due to feeding v and $U(x,t)$ is the loss in EFRS as a result of death. $F_v(u)$ is the reward from feeding minus the cost of feeding in terms of predation. An ESS strategy u^* for the whole group (all group members choose u^*) is found in the sense of Maynard Smith & Price [39], i.e., the mutant's best choice is u^* when all other group members choose u^*.

Without going any further into the mathematics, we see how the ESS strategy is derived (for more mathematical details see [32]). The model is important because it shows how the ESS depends on both the effects of dilution and of vigilance, i.e., how the main anti-predation effects interact to give rise to emergent group behaviours, explained in section 1.4.6. Interestingly, the authors briefly propose a way in which the model can capture the effects of confusion. By altering the dilution functions B_n, C_n to be size dependent, and in particular, by using nB_n as a decreasing function of n to reflect a reduced predator attack success rate with group size, somewhat tentatively the model is argued to "represent these [confusion] effects" [32] p.649.

1.6 Outline of the subsequent research

Evolutionary game theory introduced in section 1.1 provides a variety of modelling approaches to study biological problems. In chapter 2 we investigate the changing frequencies of fixed strategies using deterministic dynamics similar to those in section 1.1.5. The strategies in chapter 2 relate to anti-predator behaviours in groups of prey, e.g., cooperative vigilance, and the games played between group members have an intrinsic connection to the multiplayer games of section 1.1.3. To prescribe a measure of fitness (which depends on its strategy) to a group member, we consider

the extensively anti-predation effects introduced in sections 1.4.1-1.4.5, and how they might be interdependent when giving rise to an individual's overall predation risk. We adopt a framework of modelling the population composition as a Markov chain (section 1.2) and define two different transition probabilities. Importantly, we study two different ways in which strategies can spread, in the frequency dependent Moran, and the pairwise comparison of section, both with a behavioural mutation term. This provides differing interpretations of how group members change their behaviours to better performing ones, and we show that in the deterministic limit, we obtain two different dynamic systems.

In chapter 3 we qualitatively assess the theoretical results presented in chapter 2 using data on a specific predator-prey system. Some of the mathematical equivalences introduced in sections 1.1.3 and 1.1.4 are leveraged to draw correspondences from stable population states derived in chapter 2 to individual strategies in the prey species.

Chapter 4 uses the techniques of adaptive dynamics as introduced in section 1.3 and the G-function, as introduced in section 1.1.6. In this chapter we seek to find the change in some continuous anti-predator strategy, as opposed to the change in frequencies of fixed anti-predator strategies. We explore how investment into an anti-predator strategy results in a trade-off between individual risk dilution and other benefits of grouping, explained in section 1.4.6. Different trade-off relationships are considered, and we use adaptive dynamics to categorise evolutionary singular strategies. We compare cases in which there is a continuously stable strategy adopted by all group members, to cases in which a diversification of behaviours is predicted, i.e., a behavioural branching point. Applying the G-function allows us to explore the evolutionary singular strategies using a system of differential equations.

Chapter 2

Finite anti-predator strategy dynamics

Abstract

Three prominent principles are often theorised to explain the anti-predation benefits of grouping: the group vigilance effect, the dilution effect, and the confusion effect. The influence that each mechanism has on a group member's survivorship depends on that group members behaviour, as well as the other behaviours in the group. Most studies consider each anti-predation effect in isolation. However, there is mounting evidence that the mechanisms are intertwined. In this chapter we present a simple game theoretical model which considers how these three anti-predation effects may interact when providing safety benefits for grouping individuals. We show how this interaction can explain the behaviours emerging from groups. Individuals are subject to a binary strategy space; to either "contribute" towards anti-predator activities, or not to contribute. To account for the potential of randomness in decision making, we assume that individuals exhibit a small propensity to explore these options. Our results describe how individual group members contribute towards anti-predator activities. Subsequently, we clarify how anti-predator behaviours change with factors relating to predation risk and environment. Our key findings show that individuals alter their behaviours in response to changes in the relative influence that each of the three anti-predation mechanisms have on survival.

2.1 Introduction

In the introduction, section 1.4.6 explained how the three principal anti-predation effects of group vigilance, confusion and dilution may be complementary in providing fitness enhancing benefits for individuals in groups of prey. Although a broad range of experimental research shows that these three anti-predation effects operate interdependently to give rise to the overall risk experienced by group members [18, 146], most theoretical studies tend to study each anti-predation effect in isolation [49, 188]. In this chapter we explore the interdependence of the effects of vigilance, dilution and confusion. The research presented here is inspired by the following observation made in section

1.4.6: that an individual's per-capita dilution risk is not simply equal to the inverse of group size, but instead, depends on its behaviour as well as the behaviours of other group members [145, 147], which are in turn moderated by the three dominant anti-predation effects. As motivation consider Lima & Bednekoff [227] who find that in flocks of house finch *Carpodacus mexicanus*, the direct (predator) information an individual obtains from personal vigilance is of better quality than indirect information from others. The author's findings are significant in suggesting that vigilant individuals can increase their safety by moving near non-vigilant individuals through a favourable risk dilution: vigilant individuals respond faster and are relatively less likely to be targeted.

In this research chapter we develop a simple and explicit mathematical model to capture how the three principal anti-predation effects can interact, and can give rise to non-trivial compositions of behavioural phenotypes. The model we present explores the strategic options for individuals with a group of prey. Group members can adopt either of two strategies which we term cooperate and defect. We adopt a broad understanding of what it means for an individual to adopt a cooperative behaviour (for a more specific definition see [228]) and use the term for the narrative of the game theoretical model. Individuals that adopt the cooperative behaviour are perceived to contribute towards the whole group escaping the predator. Thus, we use the term cooperate to reflect a type of public goods production [59], in the sense that some group members contribute to an increase in the safety of all group members. We propose that the individuals contributing towards predator evasion, however, are not necessarily doing so to directly increase the safety of others. By contributing towards overall group safety, an individual may simply be behaving to reduce its own predation risk, which concomitantly improves the safety of others, e.g., vigilant is generally assumed to benefit a focal individual (early response) and the whole group (alerting others to threat). We do not consider indirect benefits of cooperation such as kin selection or reciprocated altruism [229–231]. Individuals that adopt the behaviour we term defect make no contribution to predator defence. To illustrate, actively contributing to detecting a predator can render one's experience of the dilution effect more favourable [25, 31, 116, 132, 195, 201]. Whereas, alarm calling can magnify an individual's chance of being explicitly targeted, due to increased conspicuousness [172, 232–234] or due to preferential predator selection [232, 235]. In both cases, the relative per-capita predation risk an individual contributing to overall group safety – i.e., the cooperate behaviour – faces depends on (i) its experience of the dilution effect given that some group member is attacked, i.e., its risk of being the predator target and (ii) the probability that no group member is attacked, i.e., the whole group evades due to confusion effects or group vigilance. A delicate interplay between the anti-predation effects therefore influences each group member's behavioural choices.

By combining the interaction of the effects of vigilance, confusion and dilution into a simple theoretical model we aim to explain the behavioural choices grouping individuals make in response to predation. To our knowledge, no previous theoretical model has investigated the emergent selection of anti-predation strategies in a population of groups of prey, that are a consequence these effects. Elucidating this will explain the selection of anti-predator behaviours for a wide range of group living animals that are subjected to predation.

2.2 Model

We consider groups of prey in which each group member must balance self-maintenance with avoiding predation [236, 237]. By self-maintenance we refer to foraging and other activities that have future consequences for an individual's reproduction. Within a group, an individual can choose from the strategies that we term to cooperate (C) or defect (D), as described above. The choice of strategy specifies the individual's predation risk and general grouping benefits. We model a focal group member's payoff as the product of their general benefits from grouping and the probability that they are not targeted by the predator upon an attack. The probability that the focal group member is not targeted by the predator is composed by the sum of the probability that the entire group escapes the predator (no group member is targeted) which we term "evasion", and the conditional probability that given a group member is targeted, it is not the focal individual.

An individual's general benefits from grouping are denoted by γ_i. Such benefits depend on the individual's respective choice of strategy $i = C, D$. We assume that γ_i are independent of predation risk and will to some extent be species specific. The general grouping benefits, γ_C, experienced by individuals playing cooperate, differ from general grouping benefits, γ_D, experienced by those playing the defect strategy. We allow for both of $\gamma_D \lesseqgtr \gamma_C$. Henceforth for mathematical analysis we often simply refer to the benefit differential defined as $\delta := \gamma_C - \gamma_D$.

We apply a function of $g(j, n)$ to describe the probability that the entire group evades the predator. Group evasion is assumed to depend on the (integer) number of C strategists in the group; j, and group size; n. In the subsequent analysis, when appropriate we replace the extensive variable j with the intensive variable x; the frequency of individuals playing strategy C (see equations (2.23) of section 2.3.2). In the case of $g(x, n) = 1$ no individuals are targeted and every group experiences zero risk of predation.

We apply $g(x, n)$ to capture the effects of confusion through an n-dependency, and vigilance through an x-dependency. Both of n and x may operate synergistically to give rise to overall predator evasion. The exact functional form that $g(x, n)$ assumes is contingent on the predator-prey system. A monotonic increasing function in both n and x with an upper bound of 1 (and a lower bound of 0) is appropriate. For analytical progress we suggest the following functions that we believe capture a broad range of predation cases:

$$g(x, n) = p_0 + \frac{p_1 n}{k_1 + n} + \frac{p_2 x}{k_2 + x} \tag{2.1}$$

where we consider the frequency of C strategies and absolute numbers in the group to respectively modulate the vigilance and confusion effects in an additive way, or

$$g(x, n) = p_0 + \frac{p_1 n}{k_1 + n} \times \left(1 - \frac{p_2 x}{k_2 + x}\right) \tag{2.2}$$

where the interaction between group size and the frequency of C strategies is taken to be multiplicative. Other functions are of course possible, but not explored here for simplicity of analysis.

Our modelling of risk dilution is motivated by the dilution function of Dehn [31] and the results of Lima & Bednekoff [145]. An underlying assumption is that when all group members choose the same strategy, each has an equal probability of not being targeted given by $1 - 1/n$, i.e., a

pure dilution of risk. We introduce a parameter $\epsilon \in (-1, 1)$ to reflect the observation that certain strategic behaviours may be more (or less) prone to targeting by a predator: if $\epsilon < 0$ the individuals playing strategy C receive a fitness benefit of being relatively less likely to be targeted compared to those individuals playing D. If $\epsilon > 0$ the individuals playing strategy C face a fitness penalty of being relatively more likely to be targeted. As such, we propose that the respective dilutions of risk experienced by a randomly chosen group member which plays either strategy C or D, in a group of size n with j players which choose C (and $n - j$ choose D) are given by

$$\text{Prob}\big(\text{ not targeted} \mid \text{choose to play } C\big) = \mathcal{D}_C(j, n, \epsilon) = \frac{(j-1)(1+\epsilon) + n - j}{j(1+\epsilon) + n - j} \tag{2.3a}$$

$$\text{Prob}\big(\text{ not targeted} \mid \text{choose to play } D\big) = \mathcal{D}_D(j, n, \epsilon) = \frac{j(1+\epsilon) + n - j - 1}{j(1+\epsilon) + n - j}. \tag{2.3b}$$

Equations (2.3) reflect the observation that the relative dilution of risk a group member experiences is dependent on its strategic choice, and on the composition of other strategies within the group. Here, if $\epsilon = 0$ neither strategy is more prone to predator targeting and $\mathcal{D}_C(j, n, 0) = \mathcal{D}_D(j, n, 0) = 1 - 1/n$. We also have that $\mathcal{D}_C(n, n, \epsilon) = \mathcal{D}_D(0, n, \epsilon) = 1 - 1/n$: if all group members play either C or D, then all face equal risk.

Payoff is derived by applying the above assumptions regarding grouping benefits, the probability of group evasion (equations 2.1 and 2.2) and an unequal risk of dilution (equations 2.3). A focal individual in a group with $n - 1$ opponents, receives payoff π_C when choosing to play strategy C, and π_D when choosing to play strategy D. Here

$$\pi_C(j) = \gamma_C\Big\{g(j+1, n) + \big(1 - g(j+1, n)\big)\mathcal{D}_C(j+1, n, \epsilon)\Big\} \tag{2.4a}$$

$$\pi_D(j) = \gamma_D\Big\{g(j, n) + \big(1 - g(j, n)\big)\mathcal{D}_D(j, n, \epsilon)\Big\} \tag{2.4b}$$

given that the focal individual faces j opponents in the group that play strategy C, and $n - 1 - j$ opponents in the group that play strategy D. The $j + 1$ term in the evasion function $g(j + 1)$, and risk dilution function $\mathcal{D}_C(j + 1, n, \epsilon)$ in equation (2.4a) reflects the focal individual contributing to an additional C strategy in the group (given j opponents in the group already play C): when the focal individual's choice in strategy is considered, the total number of C strategies in the group becomes $j + 1$. When the focal individual chooses strategy D, there remains j opponents playing strategy C in the group.

The sign structures of ϵ and δ, explained in Fig 2.1, can subject an individual to conflicting pressures in its choice of strategy. We identify two regimes of Fig 2.1 that are of interest (and produce non-trivial dynamics) which reflect a trade-off: $\delta, \epsilon < 0$ and $\delta, \epsilon > 0$. These cases are the focus of this research chapter.

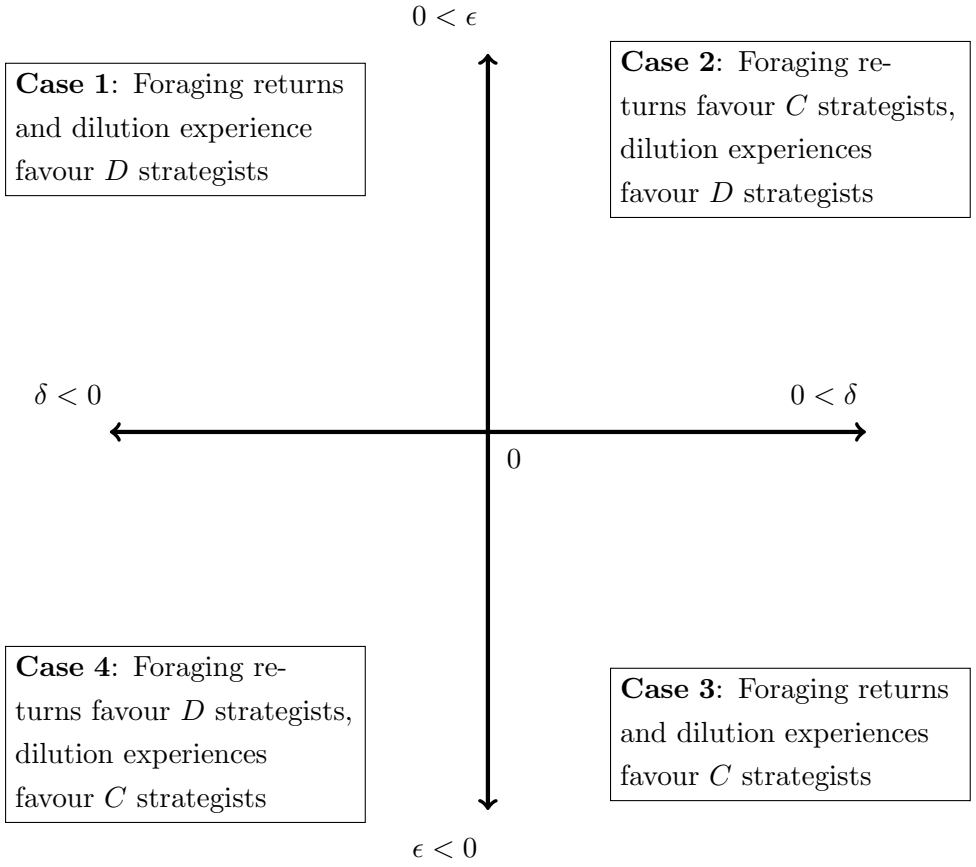


FIGURE 2.1: The model trade-off: When $\epsilon < 0$ (resp $\epsilon > 0$) those individuals playing strategy D (resp C) are more likely to be targeted by the predator. When $\delta < 0$ ($\delta > 0$) those individuals playing strategy D (C) receive greater general grouping benefits.

2.2.1 Dynamics

We now consider a well mixed population of N individuals, the size of which does not change with time. Since the population is well mixed, each individual is able to interact with any other individual. The state of the population can be fully specified by the number of individuals playing strategy C, given by $i \in \{0, ..., N\}$, and is defined through a discrete state, discrete time Markov process [238–240]. Within the population groups are randomly formed. Groups are subject to bouts of foraging and predation which yield individual payoffs. We assume that each group occupies a site and that group size is fixed by the prevailing environmental conditions, such as feeding site quality, population health and perceived predation risk. This assumption is an extremal interpretation of the ideal free distribution hypothesis in which resources are equally spread between a number of sites [241, 242].

The concept of how the strategic distribution of the population changes is straightforward; payoffs are determined from strategic interactions within each group, and imitation of the better performing strategies occurs anywhere (a well mixed population). Imitation is not necessarily pure: with a certain probability the strategies with higher payoffs are selectively imitated, and with another probability we allow for individuals to randomly "explore" their strategic options [40,

Parameter	Symbol	Value	Description
Parameters used in the strategy update process (section 2.3.2)			
Population size	N	$N >> 1$	Infinitely large, or large but finite
Group size	n	variable	How many individuals in the group
Population frequency of C	x	$x \in [0, 1]$	State variable in the replicator dynamics
Population fitness	$f_C(x), f_D(x)$	variable	Respective population fitness of C and D strategists
Exploration rate	μ	variable	The rate at which any individual randomly samples between strategies C and D
Parameters used for payoff functions (section 2.2)			
no. of C opponents in group	j	integer $\in [0, n-1]$	Non-predation benefits of strategy $i = C, D$
General benefits	γ_i	variable	Non-predation benefits of strategy $i = C, D$
Evasion function	$g(j, n)$	variable	Probability that no group member is killed
Benefits differential	δ	variable	Difference in general benefits between C and D strategists
Unequal dilution term	ϵ	variable	When $\epsilon < 0$ the D strategists are preferentially targeted; when $\epsilon > 0$ the C strategies are preferentially targeted.
Evasion constant	p_0	variable	Probability of group evasion, independent of group size, n, and the number j of C strategies
Evasion n asymptote	p_1	variable $\in [0, 1]$	Defines the probability of evasion depending on group size
Evasion j asymptote	p_2	variable $\in [0, 1]$	Defines the probability of evasion when all group members play C
Evasion n sensitivity	k_1	variable >0	Measures the sensitivity of evasion to group size
Evasion j sensitivity	k_2	variable $\in [0, 1]$	Measures the sensitivity of evasion to group composition

TABLE 2.1: Parameters relating to fitness and the dynamics

243]. Exploration does not introduce a novel strategy (i.e, a different strategy than the two present), but instead lowers the probability that one individual correctly imitates another. For the timescale on which the imitation occurs, we assume predation pressures and general group attributes to be fixed. Each individual therefore has a payoff determined by their strategic response to the prevailing environmental and predation pressures.

We proceed as follows: in section 2.3.1 we present the results of deterministic dynamics when $\mu = 0$ and $N \to \infty$ which give the frequency of individuals choosing strategy C. In section 2.3.2 we show how these results are justified by presenting a stochastic representation of the model. A master equation governs the stochastic dynamics of the probability that the population is in a given state. We present two potential transition probabilities for the imitation process. Each transition probability is based on different assumptions of how individuals update their strategic choices. In particular, we show that the choice of a behavioural frequency dependent Moran process (bm) [94] or a pairwise comparison process (pc) [90], can in the deterministic limit lead to different dynamic systems when $\mu > 0$. In section 2.3.3 we investigate these qualitative differences.

Here, we state the dynamic equations that the rest of this chapter analyses. Using either of two behavioural updating processes denoted by (bm) and (pc), we find that two subtly different deterministic equations emerge from the master equation expansion (see section 2.3.2):

$$\dot{x} = \frac{(1-x)x(f_C(x) - f_D(x))}{\langle f \rangle} + \mu \frac{(1-x)f_D(x) - x f_C(x)}{\langle f \rangle} \tag{2.5a}$$

$$\dot{x} = \frac{x(1-x)\big(f_C(x) - f_D(x)\big)}{\Delta f} + \mu \frac{(\frac{1}{2} - x)\Delta f + x(1-x)\big(f_D(x) - f_C(x)\big)}{\Delta f}. \tag{2.5b}$$

where f_C and f_D are the respective fitness of C strategists and D strategists, $\langle f \rangle$ is the average fitness in the population, Δf is the maximum difference in fitness between strategies C and D, and

$$f_C(x) = \pi_C(x) \ , \ f_D(x) = \pi_D(x). \tag{2.6}$$

Equations (2.5a) and (2.5b) characterise the results presented in the rest of this chapter. In the case of $\mu = 0$ equation (2.5a) becomes the adjusted replicator equation [35], and when $\mu > 0$ equation (2.5a) becomes the replicator-mutator equation [43]. Up to a rescaling of time, when $\mu = 0$ equation (2.5b) recovers the standard replicator dynamics. When $\mu > 0$ equation (2.5b) does not recover the replicator-mutator equation, but rather a general mutation-selection equation. A a full derivation of the dynamics is given in section 2.3.2.

2.3 Results

For mathematical analysis we present the results for a model with a reduced evasion function which only depends only on the fraction of C strategists: $g(x) = p_0 + \frac{p_1 x}{k+x}$. We do this partly not to overwhelm the notation, and partly to exhibit the interesting dynamics which can arise from a simple concept. Including group size effects in $g(x)$ through an n-dependent additive (see equation 2.1) term does not qualitatively alter the results as shown in Appendix A.3.2, but makes mathematical analysis harder to interpret.

2.3.1 Theoretical results for $\mu = 0$

When $\mu = 0$ we can apply either of equations (2.5a) or (2.5b) as the denominator terms are absorbed into a dynamical rescaling of time, to find

$$\dot{x} = x(1-x)\left\{\delta - \left(1 - p_0 - \frac{p_1 x}{k+x}\right)\left(\frac{\delta + \epsilon\gamma_C}{n(1+\epsilon x)}\right)\right\} =: J(x) \tag{2.7}$$

The roots, x^*, to equation (2.7) represent different compositions of strategies in the population. We refer to these solutions as population states. We find four different possible states which we denote x_1, x_2, x_3 and x_4. As such $x_1 = 0$ and $x_2 = 1$ respectively represent states consisting only of individuals playing strategy D, or only of individuals playing strategy C. The states x_3 and x_4 describe regimes where both strategies are present. We find that

$$\begin{aligned} x_3 =& \frac{1}{2\delta\epsilon n}\big(\delta(1-n-\epsilon nk-p_0-p_1)+\epsilon\gamma_C(1-p_0-p_1)+\Sigma\big) \\ x_4 =& \frac{1}{2\delta\epsilon n}\big(\delta(1-n-\epsilon nk-p_0-p_1)+\epsilon\gamma_C(1-p_0-p_1)-\Sigma\big) \end{aligned} \tag{2.8a}$$

where

$$\Sigma = \sqrt{4\delta\epsilon nk\big(\delta(1-p_0-n)+\epsilon\gamma_C(1-p_0)\big)+\big(\delta(n+\epsilon nk+p_0+p_1-1)+\epsilon\gamma_C(p_0+p_1-1)\big)^2}. \tag{2.9}$$

Equation (2.7) leaves the simplex Δ^2 invariant, and (asymptotic) stability of a state x^* is determined by the single eigenvalue of the (scalar) Jacobian matrix, i.e., simply $\partial_x J(x)|_{x=x^*}$. One finds that x^* is asymptotically stable if $J'(x^*) < 0$.

Group size effects

As group size is an important factor when considering the anti-predator strategies groups, we use this parameter as a natural bifurcation parameter. As such, stability is dependent on three critical values of group size n_1, n_2 and n_3:

$$\begin{aligned} n_1 &= \frac{(1-p_0)(\delta+\epsilon\gamma_C)}{\delta} \\ n_2 &= \frac{(\delta+\epsilon\gamma_C)\big((1+k)(1-p_0)-p_1\big)}{\delta(1+\epsilon)(1+k)} \\ n_3 &= \frac{(\delta+\epsilon\gamma_C)}{\delta(\epsilon k-1)^2}\left\{(1-p_0+\epsilon k(p_0-p_1-1)-p_1)+2\sqrt{\epsilon k\big((1-\epsilon k)(p_0-1)+p_1\big)}\right\} \end{aligned} \tag{2.10a}$$

In the scenario of $\epsilon, \delta < 0$ as described by case 4 of Fig 2.1 and illustrated in Fig 2.2a, the model produces regimes of dominance, co-existence and bi-stability. The pure state $x_1 = 0$ is stable if group size is larger than n_1 since

$$\begin{aligned} \lambda_1 = J'(x_1) = \frac{\delta(n-1+p_0)+\epsilon\gamma_c(p_0-1)}{n} < 0 \iff \\ n > \frac{\delta(1-p_0)+\epsilon\gamma_C(1-p_0)}{\delta} = \frac{(1-p_0)(\delta+\epsilon\gamma_C)}{\delta} = n_1 \end{aligned} \tag{2.11a}$$

The pure state $x_2 = 1$ is only stable if group size is smaller than n_2:

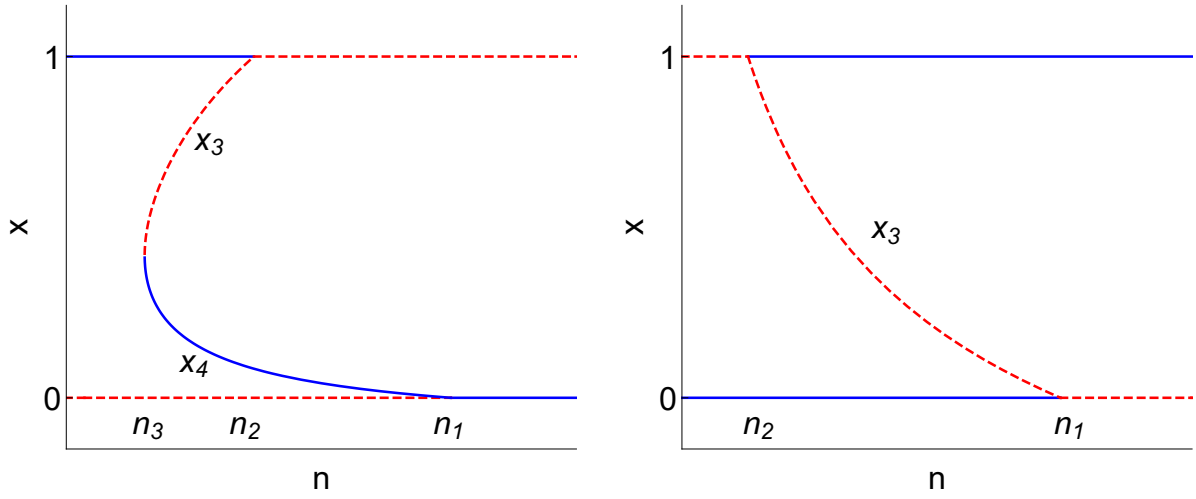
$$\begin{aligned} \lambda_2 = J'(x_2) = \frac{(\delta + \epsilon\gamma_C)\big((1+k)(1-p_0) - p_1\big) - n\delta(1+\epsilon)(1+k)}{n(1+\epsilon)(1+k)} < 0 \qquad &\iff \\ n\delta(1+\epsilon)(1+k) > (\delta + \epsilon\gamma_C)\big((1+k)(1-p_0) - p_1\big) \qquad &\iff \\ n < \frac{(\delta + \epsilon\gamma_C)\big((1+k)(1-p_0) - p_1\big)}{\delta(1+\epsilon)(1+k)} = n_2 \qquad &(2.12a) \end{aligned}$$

Mixed strategy population states x_3 and x_4 concurrently occur for $n \in (n_3, \min\{n_1, n_2\})$. At n_3 the system undergoes a fold bifurcation, such that the interior solutions leave the complex plane and become real. For $n < n_3$ both interior solutions are imaginary. For a range of group sizes between n_3 and $\min\{n_1, n_2\}$ all four solutions are real, however at most two are stable and x_3, x_4 are only physical according to parameter conditions outlined in Appendix A.3. When physical the mixed strategy state x_3 is always unstable, whereas x_4 is always stable when physical. Both fixed points leave the physical region via transcritical bifurcations at $(0, n_1)$ and $(1, n_2)$. If $n_1 < n_2$ then for $n \in [n_3, n_1)$ both x_2 and x_4 stable independent of one another, and for $n \in [n_1, n_2)$ we find that x_2 and x_1 are stable independent of one another. In both cases the stable fixed points are separated in phase space by the unstable x_3. If $n_1 > n_2$ then x_2 and x_4 are stable independent of one another for $n \in [n_3, n_2)$ and when $n \in [n_2, n_1)$ we find that only x_4 is stable in our system (not shown diagrammatically). While the interior solution x_4 is stable and physical we find $\partial_n x_4 < 0$: the mixed population state always decreases in the frequency of individuals playing C as group size increases; Fig 2.2a.

In all cases where there exist two stable fixed points, respective basin of attraction size and initial conditions determine the dynamics, i.e., which side of the unstable interior x_3 the initial state lies. At n_1 the system undergoes a transcritical bifurcation: for $n > n_1$ (resp $n < n_1$) the state x_1 is stable (resp unstable) and the interior solution $x_4 \notin (0, 1)$ (resp $x_4 \in (0, 1)$) is unstable (stable). At n_2 a second transcritical bifurcation occurs. For $n > n_2$ (resp $n < n_2$) the solution x_2 is unstable (stable) and the mixed state $x_3 \notin (0, 1)$ (resp $x_3 \in (0, 1)$) is stable (resp unstable).

In the case of $\epsilon, \delta > 0$ as described by case 2 of Fig 2.1 and illustrated in Fig 2.2b, the critical group sizes are such that n_2 smaller than n_1. The state x_1 is stable for $n < n_1$, and x_2 is stable for $n > n_2$. The only possible interior solution is given by x_3. We find that $x_3 \in (0, 1)$ if and only if $n_2 < n < n_1$. However x_3 is always unstable. The other mixed state x_4 never exists between zero and one, and we pursue it no further. For a range of group sizes between n_2 and n_1 we find that both of x_2 and x_1 are stable, and separated in the phase space by the unstable state x_3. Transcritical bifurcations at $(0, n_1)$ and $(1, n_2)$ define this region of bi-stability (see Fig 2.2b). Basins of attraction and initial conditions again fully specify the resulting dynamics. In the regime of bi-stability, by increasing group size we find that the basin of attraction of the state x_2 increases. Thus for some value of $n \in (n_2, n_1)$ the strategy of playing C becomes risk dominant. The group size n_1 can be considered as a threshold group size: groups of size larger than n_1 propagate individuals choosing to play the strategy C in the population.

A full analysis of all theoretical results can be found in Appendix A.3.



(A) $\delta, \epsilon < 0$ (B) $\delta, \epsilon > 0$

FIGURE 2.2: Theoretical representation of our system. Solid blue curves depict stable population states, and red dashed curves show unstable states. (A): The three critical group sizes (in text) specify regions of bi-stability, co-existence and dominance of strategies in the population. (B): The critical group sizes specify regions of bi-stability and dominance. Note that unless states otherwise all subsequent figures adopt the same colour scheme, i.e., blue-stable fixed point, and red-unstable.

Analysing the anti-predation effects

The importance of each anti-predation effect can shown by evaluating the results upon independently removing each of the effects. If we do not consider a difference in the dilution experience between those individuals playing strategy C and those individuals playing strategy D (i.e., $\epsilon = 0$) then the results become trivial and the strategy with greatest grouping benefits dominates; for $\delta > 0$ individuals will always choose to play C.

The effects of vigilance and confusion are effectively removed by setting $p_1 = 0$. In this case we find that there exists only one interior solution:

$$x_3^{(p_1=0)} = \frac{\gamma_C(1-p_0)\epsilon + \delta(1-n-p_0)}{\epsilon\delta n} \tag{2.13}$$

which is always unstable. One sees that $x_3 \in (0,1)$ depending on two critical group sizes:

$$n_1^{p_1=0} = \frac{(1-p_0)(\delta+\epsilon\gamma_C)}{(1+\epsilon)\delta} \tag{2.14a}$$

$$n_2^{p_1=0} = \frac{(1-p_0)(\delta+\epsilon\gamma_C)}{\delta}. \tag{2.14b}$$

While $\delta, \epsilon > 0$ we have that $x_3^{(p_1=0)} \in (0,1) \iff n_1^{p_1=0} < n < n_2^{p_1=0}$ whereas if $\delta, \epsilon < 0$ the reverse (i.e., $n_2^{p_1=0} < n < n_1^{p_1=0}$). When $x_3^{(p_1=0)}$ exists it acts to separate the basins of attraction of stable states x_1 and x_2. The population will always be monomorphic with the proportion of C strategists being equal to zero or one.

Varying other parameters

By considering how n_1 and n_2 change with other models parameters we illustrate the biological features that our model captures, and how they influence the population frequency of strategy C.

In the trade-off regime of case 4 in Fig 2.1 ($\delta, \epsilon < 0$), the value n_1 is the minimum group size for which a population in which all individuals play strategy D is viable (i.e., x_1 is stable), and n_2 the maximum group size for which a state of all C strategists is viable (i.e., x_2 stable). By increasing the general grouping benefits to those individuals playing strategy D, we find that both n_1 and n_2 decrease – i.e., – both $\partial_{\gamma_D} n_1 < 0$ and $\partial_{\gamma_D} n_2 < 0$. The converse applies when increasing grouping benefits afforded to those individuals choosing strategy C ($\partial_{\gamma_C} n_1 > 0$ and $\partial_{\gamma_C} n_2 > 0$).

Both n_1 and n_2 depend on the parameter p_0; the probability of evasion independent of the frequency of C strategists. Additionally n_2 depends on k; the sensitivity of evasion to the frequency of individuals playing strategy C, and p_1; the parameter that specifies the asymptote of $g(x, n)$ if all individuals play strategy C. We observe that $\partial_{p_0} n_1 < 0$ and $\partial_{p_0} n_2 < 0$. This is explained as follows: as p_0 increases there is a lower chance of any group member being targeted, independent of the frequency of C strategists. This has the effect to relatively lower the influence of fitness benefits through risk dilution for each group member. Because those individuals playing strategy C receive an additional fitness benefit through risk dilution, and the effect of risk dilution on fitness decreases, the relative fitness of the C strategists decreases relative to those individuals playing strategy D. We also find that $\partial_{p_1} n_2 < 0$ and $\partial_k n_2 > 0$. The result of a decrease in the maximum group size to maintain a population of C strategies with an increase in p_1 illustrates that for no change in the number of individuals playing strategy C, by increasing the absolute probability of group evasion (increasing p_1) the dilution effect becomes less important in an individual's survival. In this case the fitness of those individuals playing D increases relative to those playing C (since $\delta < 0$) and n_2 decreases. This result ($\partial_{p_1} n_2 < 0$) reflects cases in which the effects of group vigilance or confusion are enhanced without any change in the strategic behaviours of group members, and thereby increase the probability that the group as a whole evades attack for no change in x. The result of $\partial_k n_2 > 0$ is due to an increase in k lowering the probability of evasion for any group of given size and frequency of individuals playing strategy C. In order to maintain a high probability of evasion — i.e., to maintain the level of safety pre increase in k – a greater frequency of those individuals playing strategy C is required. Games within larger groups (of size n_2) than previously permitted (before the increase in k) will then also result in a population state in which all individuals play C. The increase in group size for which the pure population state of x_2 is stable reflects individuals choosing the strategy C in larger groups in order to maintain the anti-predation benefit of evasion, i.e., the value of n_2 increases.

For the case of $\delta, \epsilon > 0$ we find that varying the parameters produce diametrically opposed effects on the critical group sizes compared to $\delta, \epsilon < 0$. Parameter changes that increase (decrease) n_i and encourage (discourage) a certain behaviour in the regime of $\delta, \epsilon < 0$, now decrease (increase) n_i and discourage (encourage) that behaviour in the regime of $\delta, \epsilon > 0$, due to the opposite logic of above. A full analysis of all theoretical results can be found in Appendix A.3.

2.3.2 Derivation of the deterministic dynamics

To recover equation (2.7) we provide the following stochastic descriptions of the model. Changes in the population state to occur through a one-step process, governed by transition probabilities $T^{\pm}, \Upsilon^{\pm}$ from state i to $i \pm 1$. The exploration rate, μ, enters the mathematics precisely the same as a mutation [96] and can be conceptualised as recurrent behavioural mutations.

Transition probabilities

We first consider that transitions are determined by a behavioural frequency dependent Moran process [93]. In this case, the population can increase by an individual that adopts the strategy C by one of two ways. With a probability $1-\mu$ an individual who plays strategy C is chosen to be imitated with a probability proportional to fitness: $(i/N)(f_C/\langle f\rangle)$ where $\langle f\rangle = (if_C + (N-i)f_D)/N$ is the average fitness in the population. A randomly selected imitator is a D strategist (with probability $(N-i)/N$) and adopts the strategy C. Alternatively, with a probability μ an individual playing strategy D is chosen to be imitated (with probability proportional to fitness: $((N-i)/N)f_D/\langle f\rangle)$, and the randomly chosen imitator plays strategy D (with probability $(N-i)/N$) . However, rather than adopting the strategy of the individual chosen to be imitated (i.e., the D strategist), the imitator explores its strategic choices and adopts the strategy C. Hence we arrive at the following transition probabilities for the behavioural Moran (bm) process:

$$T^{+} = (1-\mu)\frac{i}{N}\frac{N-i}{N}\frac{f_C}{\langle f\rangle} + \mu\frac{N-i}{N}\frac{N-i}{N}\frac{f_D}{\langle f\rangle} \tag{2.15a}$$

$$T^{-} = (1-\mu)\frac{i}{N}\frac{N-i}{N}\frac{f_D}{\langle f\rangle} + \mu\frac{i}{N}\frac{i}{N}\frac{f_C}{\langle f\rangle} \tag{2.15b}$$

By definition, if there is no change in the number of individuals playing the strategy C, we have $T^0 = 1 - T^+ - T^-$.

We also consider the pairwise comparison process [40]. A focal individual is chosen to compare its fitness with a randomly selected role model, and adopts the role model's strategy with a probability proportional to the difference in fitness. With a probability $(N-i)/N$ (resp i/N) the focal individual is a D (resp C) strategist. With a probability $1-\mu$ pure imitation occurs. In this case, the role model is a C (resp D) strategist with probability i/N (resp $(N-i)/N$). Switching to the role models strategy occurs with probability $1/2+(1/2)(f_C-f_D)/\Delta f$ (resp $1/2+(1/2)(f_D-f_C)/\Delta f$ if the focal is a C strategy and role model is a D strategy) where Δf is the maximum difference in fitness. With a probability of μ the focal individual randomly explores either of the two available strategies with equal probability. Hence, the transition probabilities for the pairwise update process (pc) are

$$\Upsilon^{+}(i) = (1-\mu)\frac{i}{N}\frac{N-i}{N}\left(\frac{1}{2} + \frac{1}{2}\frac{f_C - f_D}{\Delta f}\right) + \frac{\mu}{2}\frac{N-i}{N} \tag{2.16a}$$

$$\Upsilon^{-}(i) = (1-\mu)\frac{i}{N}\frac{N-i}{N}\left(\frac{1}{2} + \frac{1}{2}\frac{f_D - f_C}{\Delta f}\right) + \frac{\mu}{2}\frac{i}{N} \tag{2.16b}$$

Either of the transition probabilities defined by equations (2.16) or (2.15) have the following properties. If $\mu = 0$ then when $i = N$ one has $T^+(N) = T^-(N) = \Upsilon^+(N) = \Upsilon^-(N) = 0$ and if

$i = 0$ then $T^+(0) = T^-(0) = \Upsilon^+(0) = \Upsilon^-(0) = 0$. That is, the boundaries of the Markov chain are absorbing in the limit of no exploration, but when $\mu > 0$ the boundaries are reflecting.

In the transition equations (2.15) and (2.16), individual fitness is derived from the payoffs obtained from within a group via the linear mapping [93, 244]:

$$f_C = 1 - \omega + \omega \Pi_C(i) \tag{2.17a}$$

$$f_D = 1 - \omega + \omega \Pi_D(i) \tag{2.17b}$$

where $\omega \in [0, 1]$ quantifies the contribution of payoff from the game to fitness [92, 93]. We assume that $\omega \to 1$ and interactions in the game specify f_C and f_D. Equations (4.4a) also use $\Pi_C(i) = \sum_{j=0}^{n-1} H_C(i,j)\pi_C(j)$ and $\Pi_D(i) = \sum_{j=0}^{n-1} H_D(i,j)\pi_D(j)$ which map payoffs from games played within each group, to average population payoffs for each strategy. The term $H_C(i,j)$ gives the probability that precisely j out of the other $n-1$ players in the group of a player using strategy C are also C strategists (and $n-1-j$ are D strategists), given that there are i individuals of strategy C in the population. Likewise, $H_D(i,j)$ is the probability that an individual playing strategy D faces precisely j opponents in the group who are of strategy C (and $n-1-j$ of strategy D), in a population with an integer i of C strategies.

Towards the deterministic dynamics

Using either of the transition probabilities defined by equations (2.15) or (2.16), the stochastic process can be described by a master equation [84]:

$$\begin{aligned} P^{\tau+1}(i) =& P^\tau(i-1)T^+(i-1) + P^\tau(i+1)T^-(i+1) + \\ & P^\tau(i)\big(1 - T^+(i) - T^-(i)\big) \end{aligned} \tag{2.18a}$$

where $P^\tau(i)$ is the probability that the state of the system is given by i at time τ. For large N, the state space is re-scaled $x = i/N$, and time re-scaled $t = \tau/N$, and the probability density function $p(x,t) = NP^\tau(i)$ is introduced. By either a truncated Kramers-Moyal expansion (see Appendix A.1) about x and t, or a van Kampen expansion (see Appendix A.2), the diffusion approximation of the imitation process is found to have the form of a Fokker-Planck equation:

$$\frac{\partial}{\partial t} p(x,t) = -\frac{\partial}{\partial x}\big[A(x,t)p(x,t)\big] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\big[B^2(x,t)p(x,t)\big] \tag{2.19}$$

with drift and diffusion terms $A(x,t)$ and $B^2(x,t)$ respectively given by

$$A(x,t) = T^+(x) - T^-(x) \tag{2.20a}$$

$$B^2(x,t) = \frac{1}{N}(T^+(x) + T^-(x)) \tag{2.20b}$$

or replacing $T^\pm$ with $\Upsilon^\pm$ depending on the update process. By applying the Ito calculus [83, 245], a stochastic differential equation describing the evolution of the state x (rather than its probability density) can be used:

$$\dot{x} = A(x,t) + \Gamma(t)\sqrt{B^2(x,t)} \tag{2.21}$$

where $\Gamma(t)$ is multiplicative white noise. Equation (2.21) is the Langevin equation corresponding to the Fokker-Planck equation (2.19) (with no noise induced drift). By assuming an infinite population: $\lim_{N\to\infty} B^2(x,t) = 0$, the equation (2.21) becomes deterministic, and the dynamics of x change in continuous time on $\Delta^2 = [0,1]$.

Thus, for $N \to \infty$, equation (2.21) provides a deterministic representation of the system: $\dot{x} = A(x,t)$, and through specific choices of transition probabilities (2.15) or (2.16), equation (2.21) recovers one of two deterministic equations given by:

$$\dot{x} = A_{(bm)}(x,t) = T^+(x) - T^-(x) \tag{2.22a}$$

$$\dot{x} = A_{(pc)}(x,t) = \Upsilon^+(x) - \Upsilon^-(x) \tag{2.22b}$$

which are precisely equations (2.5a) and (2.5b).

In the limiting case of $\mu = 0$, either of equations (2.5a) or (2.5b) correspond to the replicator dynamics (2.7) because the dynamical rescaling of time can be absorbed into the one dimensional simplex, and thus, the fixed points of the system are unchanged. However, when $\mu > 0$ equations (2.5a) or (2.5b) produce quantitatively different fixed points depending on the value of μ. This is because both the selection and mutation dynamics are rescaled in time, i.e., the $1/\langle f\rangle$ and $1/\Delta f$ terms can no longer be absorbed into the simplex.

The last step is to give an analytical expression of fitness used in equations (2.22). Given a population of infinite size in which there is a frequency x playing strategy C and frequency $1-x$ playing strategy D, the expected payoffs Π_C and Π_D can be written as

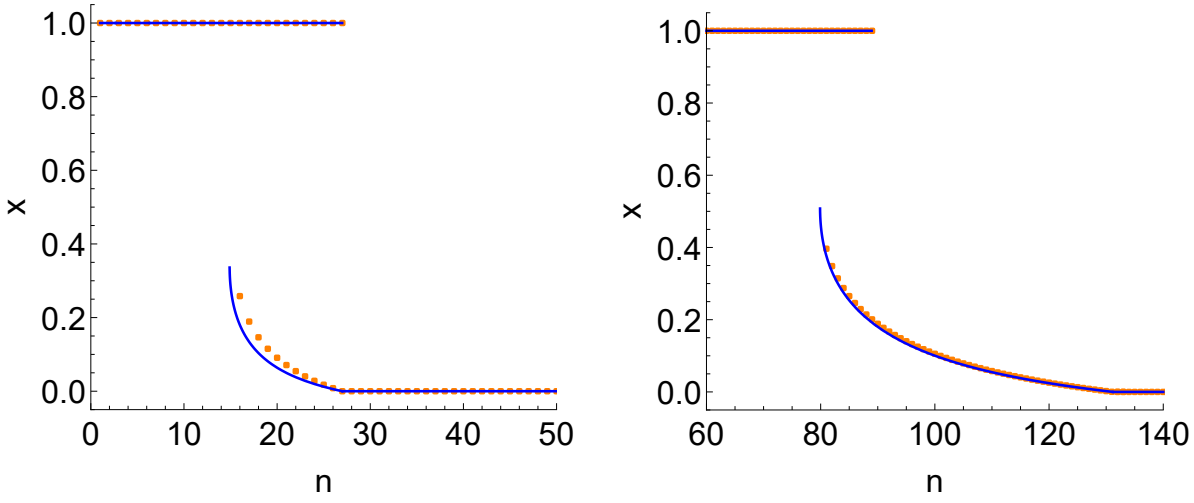
$$\Pi_C(x) = \sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{n-1-j} \pi_C(j/n) \approx \pi_C(x) \tag{2.23a}$$

$$\Pi_D(x) = \sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{n-1-j} \pi_D(j/n) \approx \pi_D(x). \tag{2.23b}$$

Equations (2.23) make use of the following approximation. Given any continuous function f, the Bernstein polynomial is defined by $B_n\big(f\big)(x) = \sum_{a=0}^{n} f(a/b) b_{a,b}(x)$ where is $b_{a,b}(x) = \binom{b}{a} x^a (1-x)^{b-a}$ is the Bernstein coefficient. The Bernstein polynomial converges uniformly to its coefficient in $[0,1]$, i.e., $\lim_{n\to\infty} B_n\big(f\big)(x) = f(x)$. We emphasise that this approximation is used for analytical progress; otherwise the replicator equation is a polynomial of degree $n-1$, and only amenable to numerical methods. In particular, the speed of the convergence is relatively unknown (see Archetti [56] for a discussion). Fig 2.3 illustrates the analytical approximation for finite n we make here.

TABLE 2.2: Simulation parameters

Parameter values used in Fig 2.4a (*) and Fig 2.4b (**)			
Parameter	**Symbol**	**Value**	**Meaning**
Population size	N	1000	large but finite population
Number of subgroups	m	variable	varies according to n_0
subgroup size	n_0	variable	strategy behaviour changes with subgroup size
Number of fitness events	M	100	determines individual fitness levels
Tournament size	K	$4^*, 1^{**}$	number of pairwise competitions between individuals
Number of generations	E	$800^*, 1000^{**}$	required iterations to obtain final state
Simulated final state	$\langle X_N \rangle$	variable	frequency of C strategists



(A) (B)

FIGURE 2.3: Comparisons between the Bernstein approximation and numerical solutions. Blue curves show the analytical approximations using equations (2.23) and orange points give the exact numerical solutions to the polynomial replicator equation (explained in text). (A): The approximation is less accurate for smaller group sizes. (B): For larger group sizes the approximation improves, and in the limit $n \to \infty$, one sees $\pi_C(j/n) \to \pi_C(x)$. Parameter values, (A): $\gamma_C = 1.9, \gamma_D = 1.92, \epsilon = -0.7, p_0 = 0.3, p_1 = 0.6, k = 0.1$ and (B): $\gamma_C = 1.96, \gamma_D = 1.964, \epsilon = -0.35, p_0 = 0.16, p_1 = 0.53, k = 0.14$.

Simulation

Due to the approximations made in equations (2.23), and in particular, our modelling of x in the dilution effect, we are aware that a potential outlet for stochastic noise to alter our results arises. For this, we adopt a fractional probability of attack after the initial group evasion function has failed. We do not take into account potential changes in group size due to successful attacks – instead targeted individuals receive a fitness penalty. As such, we would expect this to result in them changing strategy over time or alternatively for them to be more likely replaced in the larger

scale population dynamics occurring on longer timescale. This serves to maintain the population at the equilibrium value N. We adopt a continuum hypothesis assumption on the dilution function to approximate the change of an individual being targeted. For n sufficiently large this will be a valid approximation, but as n decreases it will lead to increasing stochastic effects in the population. We explicitly check our results in this region by direct stochastic simulation.

By direct stochastic simulation we test the robustness of our model for small group sizes n. We randomly select m subgroups (denoted by $\tilde{n}_i$ for each $i = 1, ..., m$) of size n from a larger population of size N such that $N = \tilde{n}_1 + \tilde{n}_2 + ... + \tilde{n}_m : |\tilde{n}_i| = n \ \ \forall i \in \{1, ..., m\}$. The initial composition of the population $\mathbf{X}_{IC_{pop}}$ is determined by assigning each individual some probability of being a C strategist. We simulate each of the m subgroups and evaluate each individual's fitness due to its interaction with the subgroup. Individuals are then returned to the population.

Step 1: Fitness Evaluation. We randomly construct a subgroup of size n from the population. The fitness of each individual in the subgroup is determined by both their strategy and of the other individuals in the subgroup. The fitness levels of individuals are additively calculated by on average "M" fitness determining events within different subgroup compositions.

Step 2: Selection. We then initiate a round of tournament selection. A new population is created by selecting n random pairs from the population subjected to fitness evaluation. A member of the new population is created by selecting the strategy with the highest fitness of the selected pair. The pairs are then replaced into the original population. This process is repeated until with have a new population of size N and then the old population is discarded. This gives us a composition state $\mathbf{X}_{pop} = (q_1, q_2, ..., q_N)$ for the population where $q_i = 0, 1$ is the strategy of individual i ($q_i = 0$ when individual i plays D and $q_i = 1$ if individual i plays C).

"E" realisations of Step 1 to Step 2 is a "generation" and gives us a final composition for the population of size N; $\mathbf{X}_N = (q_1, q_2, ..., q_N)$. The mean state in the population is then given by $\langle X_N \rangle = \frac{1}{N} \sum_{i=1}^{N} q_i$. We take this population average as representative of the replicator dynamics. We control the initial population compositions $\mathbf{X}_{IC_{pop}}$ to obtain the regions of attractions for the simulated states.

The parameters we use for the simulation are given in Table 2.2, and the results of the simulation are shown in Fig 2.4.

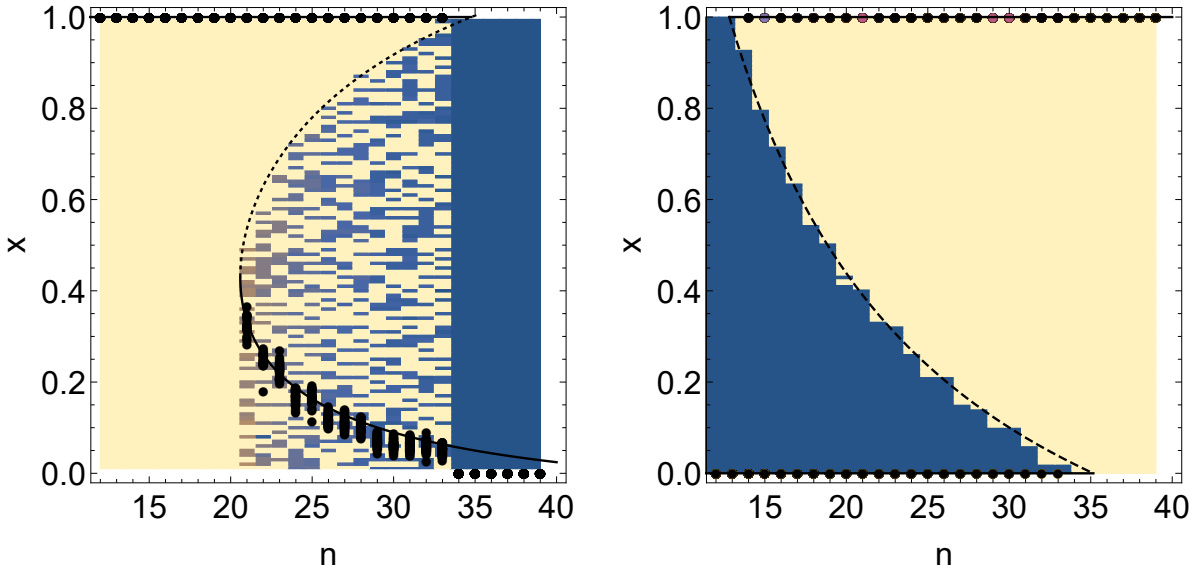


FIGURE 2.4: Simulations: the position in the $n - x$ plane gives an initial population state $\mathbf{X}_{IC_{pop}}$ given group size n. The colour at each position gives the final state $\langle X_N \rangle$; on a scale where blue corresponds to zero and cream to one. Each final state corresponds to a simulated final frequency of C strategies shown by one of the black dots (along any vertical cross section of each image). Solid black lines show the theoretical stable solutions of the adjusted replicator dynamics and dashed black lines the unstable solutions. Adjacent regions where two colours meet show simulated regions of instability and should be compared to the theoretical unstable fixed points. Initial states differ by increments of 0.01 and group size by increments of 1. Parameter values: (A) $\gamma_C = 1.9, \gamma_D = 1.92, \epsilon = -0.7, p_0 = 0.3, p_1 = 0.6, k = 0.1$ and (B) $\gamma_C = 1.92, \gamma_D = 1.9, \epsilon = 0.6, p_0 = 0.4, p_1 = 0.4, k = 0.6$

2.3.3 Results for $\mu \neq 0$

In this section we analyse how including nonzero rates of exploration qualitatively affects the results presented in section 2.3.1. Given that the dynamics with $\mu > 0$ are completely specified by equations (2.22) we use $A_{(bm)}(x, \mu)$ to denote the drift term in the replicator-mutator equation (2.5a) and $A_{(pc)}(x, \mu)$ for equation (2.5b).

We find the following roots of equation (2.5a): $x_1^{bm}(\mu), x_2^{bm}(\mu), x_3^{bm}(\mu), x_4^{bm}(\mu)$, and the following roots of equation (2.5b): $x_1^{pc}(\mu), x_2^{pc}(\mu), x_3^{pc}(\mu), x_4^{pc}(\mu)$. In general, these fixed points are analytically unwieldy but clearly recover the solutions of equation (2.7) in the limit of $\mu = 0$:

$$x_1^{bm}(0) = x_1^{pc}(0) = x_1, \qquad x_2^{bm}(0) = x_2^{pc}(0) = x_2 \tag{2.24a}$$

$$x_3^{bm}(0) = x_3^{pc}(0) = x_3, \qquad x_4^{bm}(0) = x_4^{pc}(0) = x_4 \tag{2.24b}$$

When $\mu > 0$ there is always a non-zero probability that individuals randomly explore their strategic choices. In this case, the states of x_1, x_2 are no longer valid fixed points of equations (2.5a) and (2.5b). Increasing μ pushes the stable population states away from the boundaries of $x = 0, 1$, so that each respective strategy becomes less dominant. In the limit of $\mu = 1/2$, with a probability of one half, individuals adopt strategies at random.

Bifurcation analysis for μ

Although the population states are analytically inaccessible, we proceed to provide a qualitative description of the behaviour of the solutions as μ changes, i.e., a bifurcation analysis. In the case of $\mu = 1/2$, equation (2.5a) gives a trivial fixed point of $x = 1/2$ (which is well known [246]). This fixed point is also valid for equation (2.5b). It is easy to see that $x = 1/2$ is a root of either of equations (2.5a) or (2.5b) when $\mu = 1/2$:

$$\begin{aligned} A_{(bm)}(x, 1/2) = 0 &\iff (1-2x)\big(xf_C(x) - xf_D(x) + f_D(x)\big) = 0 \\ A_{(pc)}(x, 1/2) = 0 &\iff \Delta f(1-2x) = 0 \end{aligned}$$

Since the dynamics are trivial for $\mu \geq 1/2$ we only consider the local stability of fixed points as μ varies between zero and one half. We give a detailed analysis only for the dynamics of (2.5a) (as the dynamics of equation (2.5b) are qualitatively similar). For notational ease we write $A(x, \mu) := A_{(bm)}(x, \mu)$ and drop the superscript bm in the fixed point so that $x_i^{bm}(\mu) = x_i(\mu)$, $i = 1, ..4$.

If $x(\mu)$ is a solution to equation (2.5a) then by the implicit function theorem [72],

$$\frac{dx(\mu)}{d\mu} = -\frac{A_\mu(x, \mu)}{A_x(x, \mu)} \tag{2.26}$$

where we use $A_v(x, \mu)$ for the partial derivative of A with respect to v. Recall that $A(x, 0)$ is simply the (adjusted) replicator equation, and $A_x(x, 0)$ its Jacobian. Using the fixed points $x_1(0) = x_1 = 0$ and $x_2(0) = x_2 = 1$ we find that

$$A_\mu(0, 0) = 1 \ , \ A_\mu(1, 0) = -1 \tag{2.27}$$

As stability of the adjusted replicator dynamics (equivalently, equation 2.7) must be preserved up to a non-hyperbolic value of μ^B for which $A_x(x, \mu^B) = 0$, i.e., a bifurcation value, equation (2.26) is singular and of constant sign for $\mu < \mu^B$. Thus, the initial direction of travel in the $\mu - x$ plane for $x_1(\mu)$ and $x_2(\mu)$ can be entirely determined by their stability at $\mu = 0$:

$$\frac{dx_1(\mu)}{d\mu}\Big|_{\mu=0} = -\frac{1}{A_x(0, 0)} < 0 \iff J(0) > 0 \tag{2.28a}$$

$$\frac{dx_1(\mu)}{d\mu}\Big|_{\mu=0} = -\frac{1}{A_x(0, 0)} > 0 \iff J(0) < 0 \tag{2.28b}$$

and

$$\frac{dx_2(\mu)}{d\mu}\Big|_{\mu=0} = -\frac{-1}{A_x(1, 0)} > 0 \iff J(1) > 0 \tag{2.29a}$$

$$\frac{dx_2(\mu)}{d\mu}\Big|_{\mu=0} = -\frac{-1}{A_x(1, 0)} < 0 \iff J(1) < 0 \tag{2.29b}$$

where $J(x)$ is defined in equation (2.7). To find the qualitative behaviour of $x_3(\mu)$ and $x_4(\mu)$ with increasing μ we again use equation (2.26). To find the sign of $A_\mu(x, \mu)$ at the either solution, we first observe that

$$A_\mu(x, \mu) = \frac{(1-x)f_D(x) - xf_C(x)}{(1-x)f_D(x) + xf_C(x)} = 0 \tag{2.30}$$

has precisely one unique root, $\hat{x}$, in the interval $(0,1)$ which is independent of μ. Therefore, the sign of $A_\mu(x,\mu)$ in the $\mu - x$ plane is simply partitioned by a horizontal line specified by $\hat{x}$ (see Fig 2.5b). To show that $\hat{x} \in (0,1)$, we use the intermediate value theorem with $I = [0,1]$, $A_\mu(0,\mu) > 0$ and $A_\mu(1,\mu) < 0$; there exists some $c \in (0,1)$ such that $A_\mu(c,\mu) = 0$.

To see that $\hat{x} = c$ is unique in $(0,1)$ we apply Budan's theorem [247]. Direct calculations show that the roots of equation (2.30) are those of the following univariate polynomial:

$$
\begin{aligned}
&c_1(x) = a_3x^3 + a_2x^2 + a_1x + a_0, \text{ where}\\
&a_0 = \gamma_D k - \gamma_D kn - \gamma_D kp_0\\
&a_1 = \gamma_C k\big(n - 1 + \epsilon(p_0 - 1) + p_0\big) - \gamma_D\big(k + n - 1 + (\epsilon - 1)kn + p_0 - kp_0 + p_1\big)\\
&a_2 = \gamma_D\big(n - 1 + \epsilon(k-1)n + p_0 + p_1\big) + \gamma_C\big(n - 1 + p_0 + p_1 + \epsilon(kn - 1 + p_0 + p_1)\big)\\
&a_3 = \epsilon n(\gamma_C + \gamma_D).
\end{aligned}
$$

Note that $a_0 < 0$. Let $R_{(0,1]}$ be the number of real roots of $c(x)$ in the half open interval $(0,1]$ and v_k be the number of sign changes in the sequence of the coefficients, $a_i, i = 0,1,2,3$, of $c(x+k)$. Then $v_0 - v_1 - R_{(0,1]}$ is a non-negative even integer. If $\delta, \epsilon < 0$ then $a_0 < 0, a_2 > 0$ and $a_3 < 0$ and therefore $\max\{v_0\} = 2$. Denoting $b_i, i = 0,1,2,3$ for the coefficients of $c(x+1)$, we find that $b_0 > 0$ and $b_3 < 0$. Therefore, $\min\{v_1\} = 1$. Combining $\max\{v_0\} = 2$ and $\min\{v_1\} = 1$ implies that there is at most one root $\hat{x} \in (0,1]$ which solves $c(x) = 0$. We already know, however, there exists at least one root. Hence $R_{(0,1]} = 1$ and $\hat{x}$ is unique. Alternatively, in the case of $\delta, \epsilon > 0$, we find that $a_3 > 0$ and $a_0 < 0$. Here, we use that $c''(x)$ changes sign at most once for $0 < x < 1$, combined with $c(0) < 0, c(1) > 0$, implies that $c(x)$ crosses the x-axis only once in $(0,1)$. Then there can only be one root, $\hat{x} \in (0,1)$.

Having established the uniqueness of $\hat{x}$, we use $A_\mu(0,0) > 0$ and $A_\mu(1,0) < 0$ such that if $x < \hat{x}$ then $A_\mu(x,\mu) > 0$ and if $x > \hat{x}$ then $A_\mu(x,\mu) < 0$.

If we consider the case of $\delta, \epsilon > 0$ as described by case 4 of Fig 2.1, then $A_x(0,0), A_x(1,0) < 0$ and $A_x(x_3,0) > 0$ where

$$
A_x(x,0) \equiv (1-2x)f_C(x) - (1-2x)f_D(x) + (1-x)x(f_C'(x) - f_D'(x)). \tag{2.32}
$$

If $x_3 \in (1/2,1)$, we can define $F_1(x) := f_C(x) - f_D(x)$ which is concave on the simplex ($F_1''(x) < 0, \forall x \in [0,1]$) with $F_1(0) < 0, F_1(1) > 0$ and $F_1(x_3) = 0$. By Jensen's Inequality, $\forall q \in [0,1]$ the following must hold:

$$
\begin{aligned}
&F_1(qx_3 + (1-q)1) \geq qF_1(x_3) + (1-q)F_1(1) \iff\\
&f_C(qx_3 + (1-q)) - f_D(qx_3 + (1-q)) \geq (1-q)\big(f_C(1) - f_D(1)\big) \geq 0
\end{aligned} \tag{2.33a}
$$

where we use (2.32) such that $A_x(1,0) = f_D(1) - f_C(1) < 0$. Any $x' \in [x_3,1)$ can be represented by $x' = qx_3 + (1-q)1$ for some $q \in (0,1)$. Given that $f_C(x') \geq f_D(x')$ and $x' \geq x_3 > 1/2$, equation (2.30) must be negative. Now consider $x' \in (0,x_3)$. Because at x_3, equation (2.30) is negative and at $x = 0$ equation (2.30) equals one, by the intermediate value theorem the solution $\hat{x} \in (0,x_3)$.

The solution must satisfy

$$\hat{x} = \frac{1}{1 + \frac{f_C(\hat{x})}{f_D(\hat{x})}}. \tag{2.34}$$

However $\hat{x}$ cannot be greater than or equal to one half since if the ratio $f_C(\hat{x})/f_D(\hat{x}) = \theta < 1$, then equation (2.34) would imply that $1/2 \leq \hat{x} = 1/(1+\theta) < 1/2$ which provides a contradiction. Thus $\hat{x} \in (0, 1/2)$ and for all $x > \hat{x}$ equation (2.30) is negative and for all $x < \hat{x}$ equation (2.30) is positive.

If $x_3 \in (0, 1/2)$ we define $F_2(x) := f_D(x) - f_C(x)$ with $F_2(0) > 0, F_2(1) < 0, F_2(x_3) = 0$. Since $F_2''(x) > 0$ and x_3 is unique, we see that $f_D(x'') > f_C(x'')$ for all $x'' \in (0, x_3)$. In this case equation (2.30) is positive. If $x'' \in [x_3, 1/2)$, then $F_2(x'') < 0$, but we use equation (2.34) and note that $f_C(x'')/f_D(x'') < 1$. Thus, $\hat{x} \notin [x_3, 1/2)$ and equation (2.30) is positive. It follows that $\hat{x} > 1/2$ when $x_3 \in (0, 1/2)$. Note that if $x_3 = 1/2$ then $\hat{x} = 1/2$.

In the case of $\delta, \epsilon < 0$ (case 4 of Fig 2.1) with x_2 stable and either of x_4 or x_1 stable separated in phase space by x_3 (see Fig 2.2a) the payoff functions must have properties such that $f_C(1) > f_D(1)$, $f_C(x_3) = f_D(x_3)$ and $f_C(x_4) = f_D(x_4)$. If x_1 is stable then $f_C(0) < f_D(0)$ and x_4 is not present in the system. If x_4 is physical, it is stable and $f_C(0) > f_D(0)$. While x_4 lies in the interval $(0, 1)$ we define the function $F_3(x) := f_D(x) - f_C(x)$ which is strictly convex for $x \in [0, 1]$. By the same argument as above, we use that

$$F_3(qx_4 + (1-q)x_3) \leq qF_3(x_4) + (1-q)F_3(x_3) = 0 \tag{2.35}$$

Therefore $f_D(x') \leq f_C(x')$ for all $x' \in (x_4, x_3)$. If $x_4, x_3 > 1/2$ this implies that equation (2.30) is negative at x' and thus $\hat{x} < 1/2$. If $x_4, x_3 < 1/2$ then as $f_C(x')/f_D(x') < 1$ for $x' > x_3$ we see that $\hat{x}$ must be greater than one half. If $x_4 < 1/2 < x_3$ then $\hat{x} \in (x_4, 1/2)$.

Since the curve of $A_\mu(x, \mu) = 0$ is independent of μ, once a solution branch $x_i(\mu)$ satisfying equation (2.26) reaches the (constant) line defined by $\hat{x}$ it remains there invariant with μ. From this observation one can state the following: if the unstable interior solution of the replicator dynamics, x_3, has precisely the same value as $\hat{x}$, then $dx_3(\mu)/d\mu = 0$ (as $A_\mu(x_3, \mu) = 0$ in equation (2.26) then $x_3 = \hat{x} = 1/2$ (see Fig 2.5a).

We now move on to analysing $A_x(x, \mu)$. First we note that $A_x(x, \mu) = 0$ provides threshold values of μ for which the stability of the solutions obtained from the replicator dynamics can change. We find that the only root of $A_x(x, \mu) = 0$ is

$$\hat{\mu}(x) = \frac{p_0 + p_1 x + p_2 x^2 + p_3 x^3 + p_4 x^4 + p_5 x^5 + p_6 x^6}{q_0 + q_1 x + q_2 x^2 + q_3 x^3 + q_4 x^4} \tag{2.36}$$

where each p_i, q_i are expressions of model parameters, and singularities of $\hat{\mu}(x)$ lie out of $x \in [0, 1]$.

The intersections of the curve of $\hat{\mu}(x)$ and the solution branches given by $x_i(\mu)$ produce a subset of co-dimension 2. The non-hyperbolic fixed points of the system within this space are generally of co-dimension 1, unless $x_i(\mu)$ intersects $\hat{\mu}(x)$ at its minima in the $x - \mu$ plane, in which case the singularity is more degenerate (since $A_{xx}(x, \mu) = 0$) and is of co-dimension 2 [72]. For any parameter configuration this subset provides the space of possible local bifurcations as illustrated in Fig 2.5 by blue and red solution branches, $x_i(\mu)$, in Fig 2.5a and a black curve of $A_x(x, \mu) = 0$

depicted in Fig 2.5b. Fold bifurcations occur at the intersections of solutions $x_i(\mu)$ and any point of the curve $A_x(x,\mu)=0$ so long as it does not coincide with the intersection of $A_\mu(x,\mu)=0$ or occur at the minima of $A_x(x,\mu)=0$ in the $x-\mu$ plane. Transcritical bifurcations occur when solution curves $x_i(\mu)$ coincide with the intersection of the curves $A_x(x,\mu)=0$ and $A_\mu(x,\mu)=0$ so long as this does not take place at the minima of $A_x(x,\mu)=0$ in the $x-\mu$ plane (i.e $A_{xx}(x,\mu)\neq 0$). Pitchfork bifurcations do not occur with our choice of parameter bounds ($|\epsilon|>1$ is a required condition).

Fig 2.5 illustrates how the sign of equation (2.26) determines the direction of travel of $x_i(\mu)$ with μ, by depicting grey regions as those where $dx_3(\mu)/d\mu<0$ and white for which $dx_3(\mu)/d\mu>0$ (see Fig 2.5a). Fig 2.5b provides illustrations of the curves defined by $\hat{x}$ and $\hat{\mu}$, i.e., respective zero contours of $A_\mu(x,\mu)$ and $A_x(x,\mu)$ which determine the sign of equation (2.26) and specify the grey and white regions. All possible generic bifurcations are shown in Fig 2.5a and the curve of $A_x(x,\mu)=0$ – necessary for a bifurcation – is shown in Fig 2.5b as a black cubic in the $x-\mu$ plane. For all subsequent figures this line is omitted.

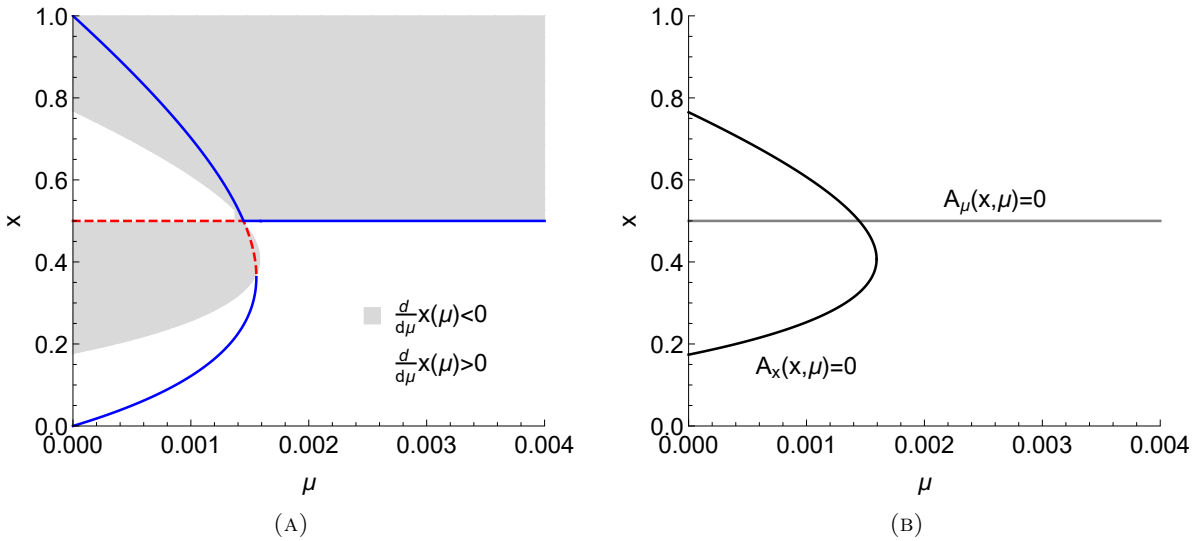


FIGURE 2.5: (A): Illustration of bifurcations. Grey regions show $dx(\mu)/d\mu<0$ and white show $dx(\mu)/d\mu>0$ as determined by equation (2.26) in text. In the grey regions, solution branches are decreasing with μ and in the white, increasing with μ. The intersection of the grey and white show points where equation (2.26) is undefined (bifurcations) or equals zero (no change in direction). Examples of a transcritical bifurcation (at the intersection of solution curves and $A_\mu(x,\mu)=0$ and $A_x(x,\mu)=0$ shown in (B)) and a fold bifurcation (where solutions cross $A_x(x,\mu)=0$) are shown. Notice that a condition for the transcritical bifurcation is that interior solution $x_3(\mu)$ is equal to $\hat{x}$ and is invariant with μ. (B): Solutions curves of $A_x=0$ and $A_\mu=0$ from equation (2.26). The root of $A_\mu(x,\mu)=0$; $\hat{x}\approx 1/2$ is shown by the grey line. The concave black curve in the $x-\mu$ plane is defined by $\hat{\mu}(x)$; the root of $A_x(x,\mu)=0$. All bifurcations must occur along $A_x(x,\mu)=0$.

If $\delta,\epsilon>0$ the only potential interior solution of the replicator dynamics is x_3, which is unstable. Then one has that $A_x(0,0), A_x(1,0)<0$ and $A_x(x_3,0)>0$. The respective stabilities of $x_1(\mu), x_2(\mu)$ and $x_3(\mu)$ are preserved up to a value of μ where a respective solution branch intersects $\hat{\mu}(x)$. Combining the information into equation (2.26) there exist four regions where $dx(\mu)/d\mu$ changes

sign. From equations (2.28) and (2.29) we know that $dx_1(\mu)/d\mu > 0$ and $dx_2(\mu)/d\mu < 0$. The position of the $x_3(\mu)$ at $\mu = 0$ determines whether $x_3(\mu)$ increases or decreases with μ: if $x_3(0) = x_3 > \hat{x}$ then x_3 lies in a region where $A_\mu(x, \mu) < 0$ and $A_x(x, \mu) > 0$ and by equation (2.26) we see that $dx_3(\mu)/d\mu > 0$. In this case a bifurcation point occurs at μ^B when $x_3(\mu)$ collides with $x_2(\mu)$ and the solutions become complex (not illustrated). Otherwise if $x_3(0) = x_3 < \hat{x}$, we see that $dx_3(\mu)/d\mu < 0$ and $x_3(\mu)$ bifurcates out of phase space when it collides with $x_1(\mu)$ as illustrated in Fig 2.6b and Fig 2.6d.

If $\delta, \epsilon < 0$ and two interior solutions $x_3 \in (0, 1)$ and $x_4 \in (0, 1)$ with $x_4 < x_3$, then from section 2.3.1 x_1 and x_3 are unstable, with x_2 and x_4 stable. In this case $A_x(0, 0) > 0$, $A_x(x_3, 0) > 0$ and $A_x(x_4, 0), A_x(1, 0) < 0$, by equation (2.28), $x_1(\mu)$ doesn't enter phase space and by equation (2.29), $x_2(\mu)$ is initially decreasing. The curve of $\hat{\mu}(x)$ is cubic in the $x - \mu$ plane and must intersect x at $\mu = 0$ between x_1 and x_4, between x_4 and x_3 and between x_3 and x_2 (see Fig 2.6a or Fig 2.6c) in order to preserve the stabilities of the replicator dynamics. Putting together equations (2.28), (2.29) and the information about $\hat{\mu}$ and $\hat{x}$ determines the possible sign changes of $dx(\mu)/d\mu$. Since $A_x(x_4, \mu) < 0$ we find that if $x_4(\mu) < \hat{x}$ at $\mu = 0$ the solution $x_4(\mu)$ always increases Fig 2.6a. In this case, if $x_3(0) < \hat{x}$ then (since $A_x(x_3, 0) > 0$) $x_3(\mu)$ is initially decreasing, and collides with $x_4(\mu)$ (not plotted). If $\hat{x} < x_3(0)$ then $x_3(\mu)$ is initially increasing and collides with $x_2(\mu)$ Fig 2.6a and Fig 2.6c.

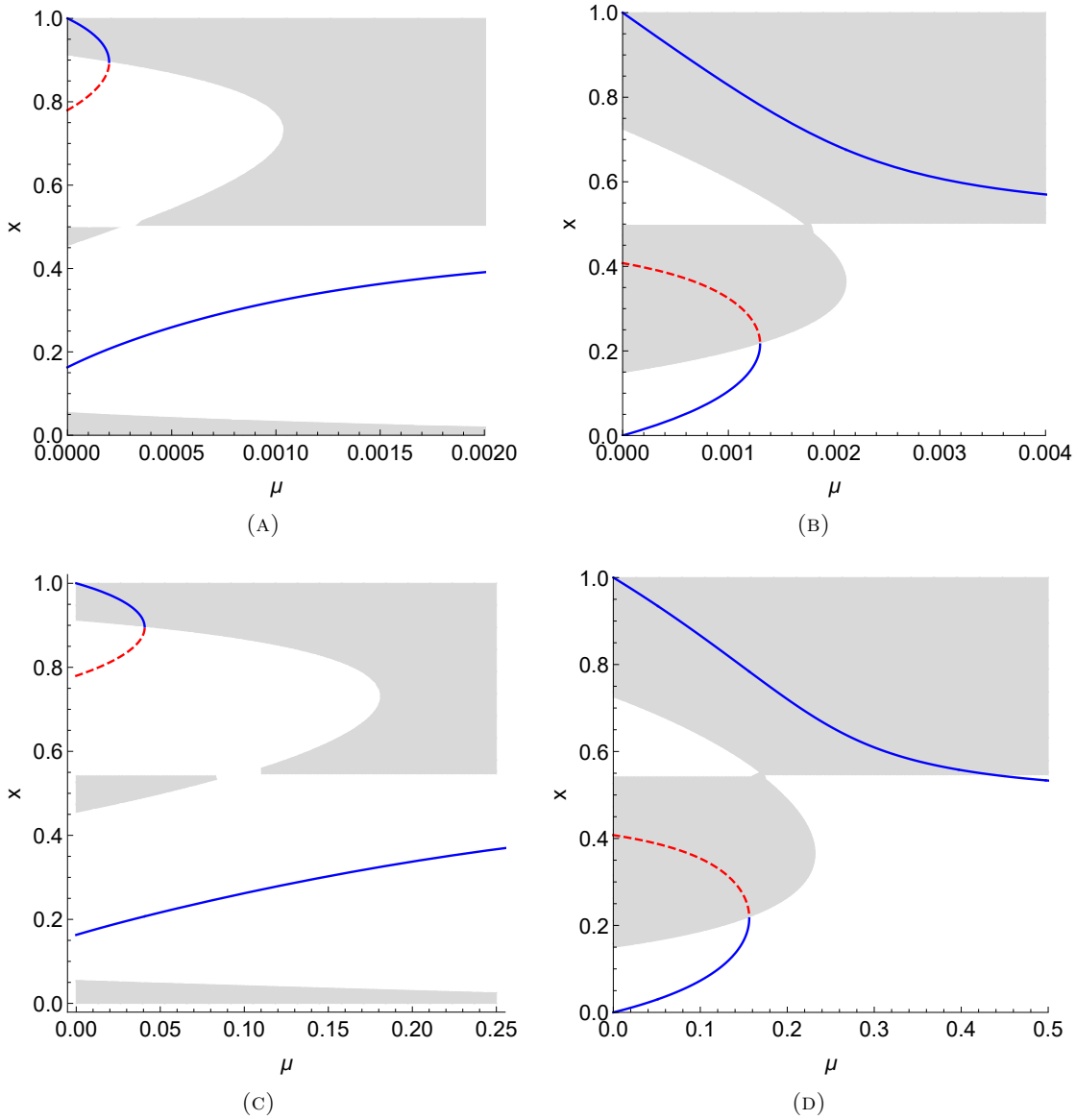


FIGURE 2.6: (A),(B): Behavioural Moran imitation and (C),(D): pairwise update. (A): $(\delta, \epsilon < 0)$. The replicator dynamics are bi-stable in $x_4 \approx 0.17$ and $x_2 = 1$. The solution branch $x_1(\mu)$ never appears in phase space. The solution $x_2(\mu)$ decreases with μ and collides with $x_3(\mu)$ in a fold bifurcation at which both solutions turn complex. The solution $x_4(\mu)$ increases with μ and equals $1/2$ when $\mu = 1/2$ (B): $(\delta, \epsilon > 0)$. The replicator dynamics are bi-stable in $x_1 = 0$ and $x_2 = 1$. The solutions $x_1(\mu)$ and $x_2(\mu)$ respectively increase and decrease with μ and a collision between $x_3(\mu)$ and $x_1(\mu)$ resulting in the fixed points leaving phase space. (C),(D): Qualitatively similar pattern, but for larger μ values. Parameter values: $\gamma_C = 1.9, \gamma_D = 1.92, \epsilon = -07, p_0 = 0.3, p_1 = 0.6, k = 0.1, n = 25$ (Left) and $\gamma_C = 1.92, \gamma_D = 1.9, \epsilon = 0.6, p_0 = 0.4, p_1 = 0.4, k = 0.6, n = 25$ (Right)

Group size bifurcations

When $\mu > 0$ the critical group sizes analysed in section 2.3.1 are analytically inaccessible. The system given by equation (2.5a) can be viewed as a perturbation of equation (2.7). By including the perturbation term (the coefficient of μ) of equation (2.5a) the orbit structure near the bifurcation points defined by n_1, n_2, n_3 is affected by the addition of terms of order less than $\mathcal{O}(x^3)$. We find

that the transcritical bifurcations at $(n_1, 0)$ and $(n_2, 1)$ are structurally unstable: for $\mu \neq 0$ but $\mu << 1$, each transcritical bifurcation is replaced by either of a pair of fold bifurcations occur at critical values of $n_1^{f_1}, n_1^{f_2}$ and $n_2^{f}, n_2^{f_2}$, or by a pair of continuity curves (no bifurcations), one physical and one not. Note that only one fold bifurcation occurs for $x \in [0, 1]$ local to each of n_1 and n_2. Perturbation of the fold bifurcation at (n_3, x_4) does not introduce any qualitatively new dynamical behaviour.

Case 4 of the trade-off Fig 2.1 (i.e., $\delta, \epsilon < 0$) is shown in Fig 2.7 and Fig 2.8. Fig 2.7a illustrates that if $\mu = 0$ the critical group sizes are the same as the results of section 2.3.1, and can be inferred from comparison to Fig 2.2a. For small exploration rates, e.g., $\mu = 0.0001$ shown in Fig 2.7b the perturbation is small, and replicator-mutator dynamics closely resemble the adjusted replicator equation (Fig 2.7a), with the exception that the states $x = 0, 1$ are now repelling. As such the transcritical bifurcations occurring at $(n_1, 0) \approx (47, 0)$ and $(n_2, 1) \approx (33, 1)$ are respectively replaced by a local fold bifurcation and a continuity curve. The fold bifurcation of the replicator dynamics remains. With increasing μ the fold bifurcation diverges farther from the transcritical bifurcation and at a critical value $\bar{\mu}$ the replicator-mutator dynamics lose their invariance with the fixed points of the replicator equation Fig 2.7c. Thus, increasing μ decreases the maximum group size $n_2(\mu)$ for which $x_2(\mu)$ is stable, and eliminates the value of $n_1(\mu)$ since the $x = 0$ state is never reached. This acts to reduce the region of group sizes in phase space where the population state is bi-stable, and simultaneously, push both of $x_1(\mu)$ and $x_2(\mu)$ towards $1/2$.

The qualitative picture is the same for the pairwise process (Fig 2.8) however invariance with the replicator dynamics is maintained for significantly larger exploration rates.

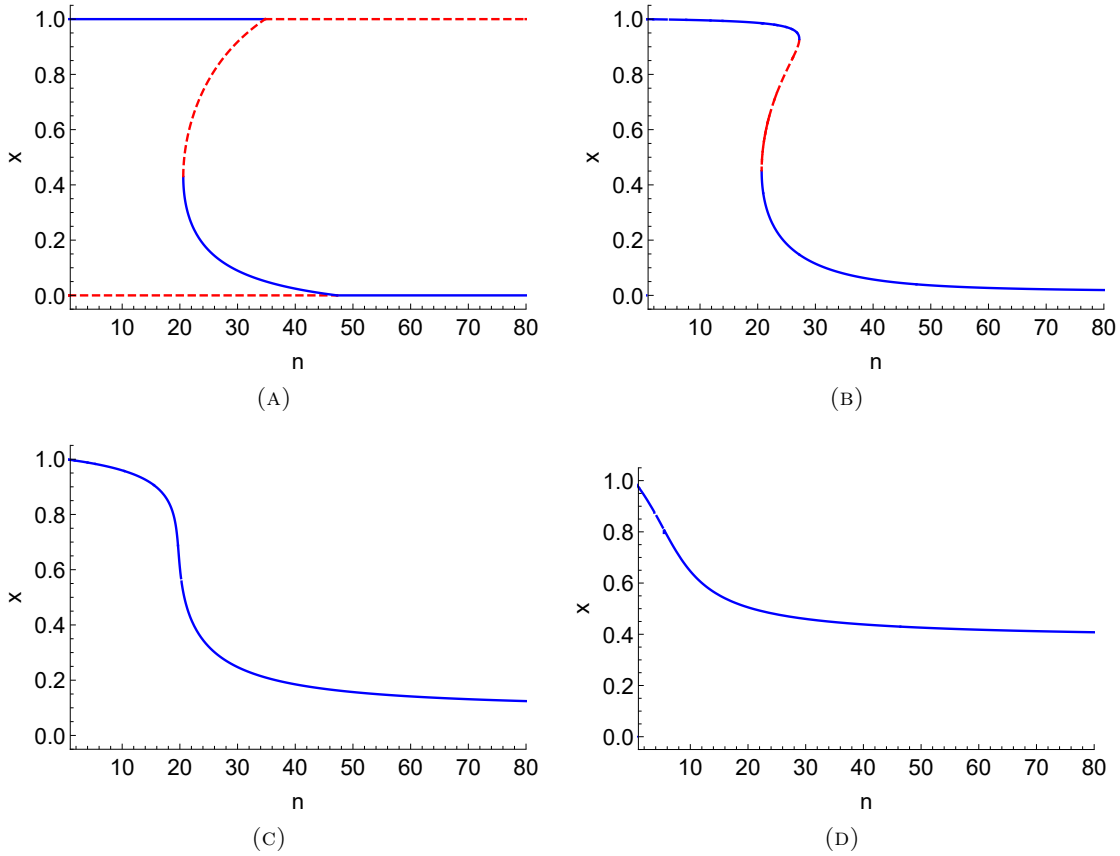


FIGURE 2.7: The effect of exploration on group size bifurcations for the behavioural Moran process: $\delta, \epsilon < 0$. (A): The replicator dynamics, as analysed in section 2.3.1. (B): For $\mu = 0.0001$ the replicator-mutator solutions largely resemble the replicator dynamics. Transcritical bifurcations are replaced by local fold bifurcations that diverge father from the bifurcation point with increasing μ. (C) and (D): $\mu = 0.001$ and $\mu = 0.01$. Solutions to the replicator-mutator equation are no longer invariant to the replicator dynamics. Parameter values: $\gamma_C = 1.9, \gamma_D = 1.92, \epsilon = -07, p_0 = 0.3, p_1 = 0.6, k = 0.1$

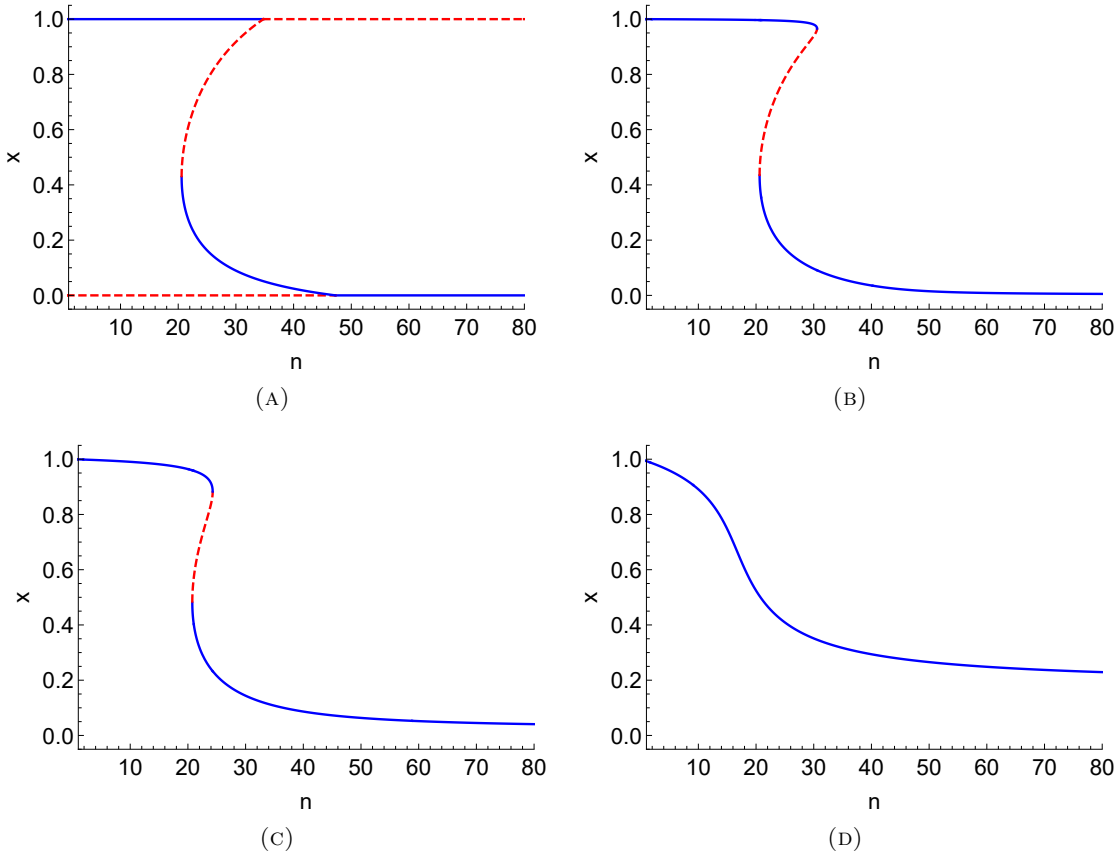


FIGURE 2.8: The effect of exploration on group size bifurcations for the pairwise comparison process: $\delta, \epsilon < 0$. (A): The replicator dynamics, i.e., $\mu = 0$. (B) and (C): $\mu = 0.0001$ and $\mu = 0.0001$ the resulting dynamics remain invariant with the replicator equation. Transciritcal bifurcations are locally replaced by a fold bifurcation and continuity curve, and the fold bifurcation remains. (D) For a large enough $\mu = 0.01$ group size bifurcations no longer occur. Parameter values: $\gamma_C = 1.9, \gamma_D = 1.92, \epsilon = -0.7, p_0 = 0.3, p_1 = 0.6, k = 0.1$

When $\delta, \epsilon > 0$ the transcritical bifurcations occurring at $(n_2, 1) \approx (14, 1)$ and $(n_1, 0) \approx (35, 0)$ in Fig 2.9a are replaced by fold bifuractions that diverge with increasing μ (see Fig 2.9b,Fig 2.9c and Fig 2.9d). Hence, we find that increasing μ decreases the maximum group size $n_1(\mu)$ for which $x_1(\mu)$ is stable, and increases the minimum group size $n_2(\mu)$ for which $x_2(\mu)$ is stable. At the same time, increasing μ pushes both of $x_1(\mu)$ and $x_2(\mu)$ towards $1/2$: the regions of bi-stability become smaller , and within the regions of bi-stability, the basin of attraction of the risk dominant strategy decreases.

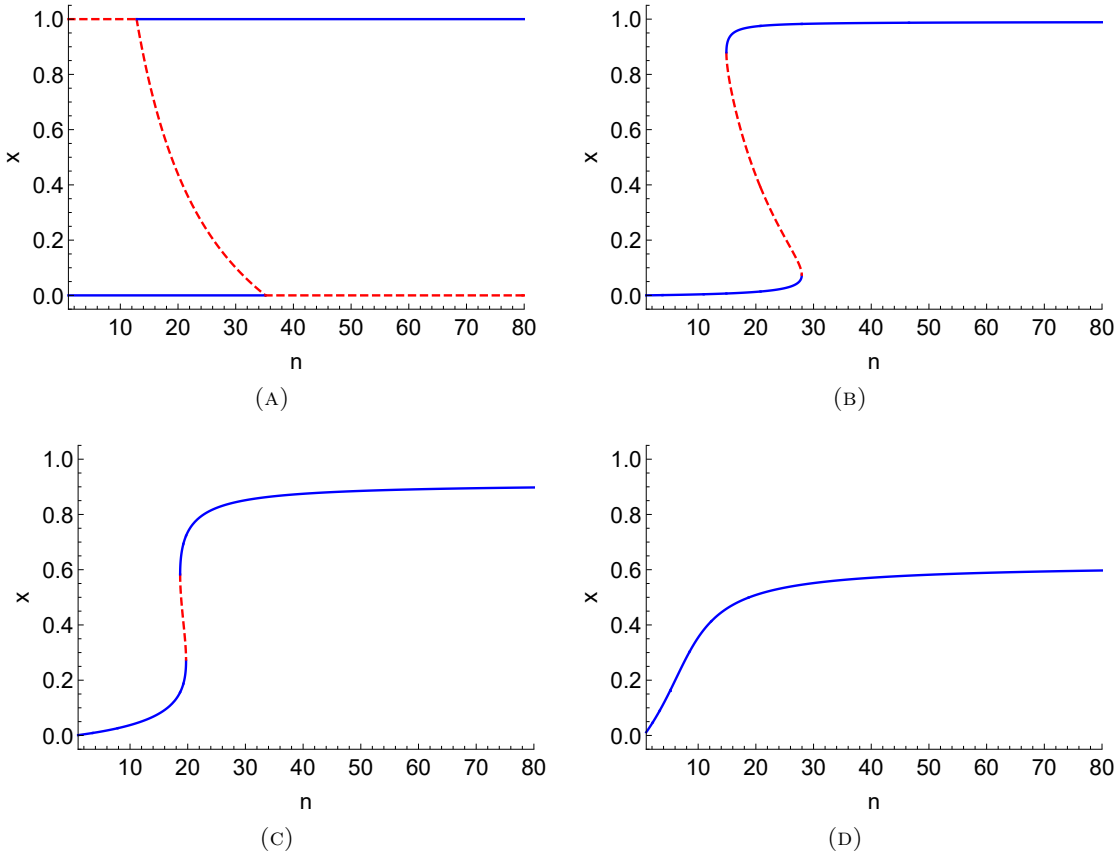


FIGURE 2.9: The effect of exploration on group size bifurcations for the behavioural Moran process: $\delta, \epsilon > 0$ (A): The adjusted replicator equation: the critical group sizes are shown by the transcritical bifurcations at roughly $n_1 \approx 35$ and $n_2 \approx 18$. (B) and (C): $\mu = 0.0001$ and $\mu = 0.001$. The replicator-mutator equation retains invariance with the replicator dynamics. Fold bifurcations diverge with $n_1 \approx 30, n_2 \approx 16$ in (B) and $n_1 \approx 20, n_2 \approx 19$ in (C). (D): When $\mu = 0.01$ invariance is lost, and stability does not change with n. Parameter values: $\gamma_C = 1.92, \gamma_D = 1.9, \epsilon = 0.6, p_0 = 0.4, p_1 = 0.4, k = 0.6$

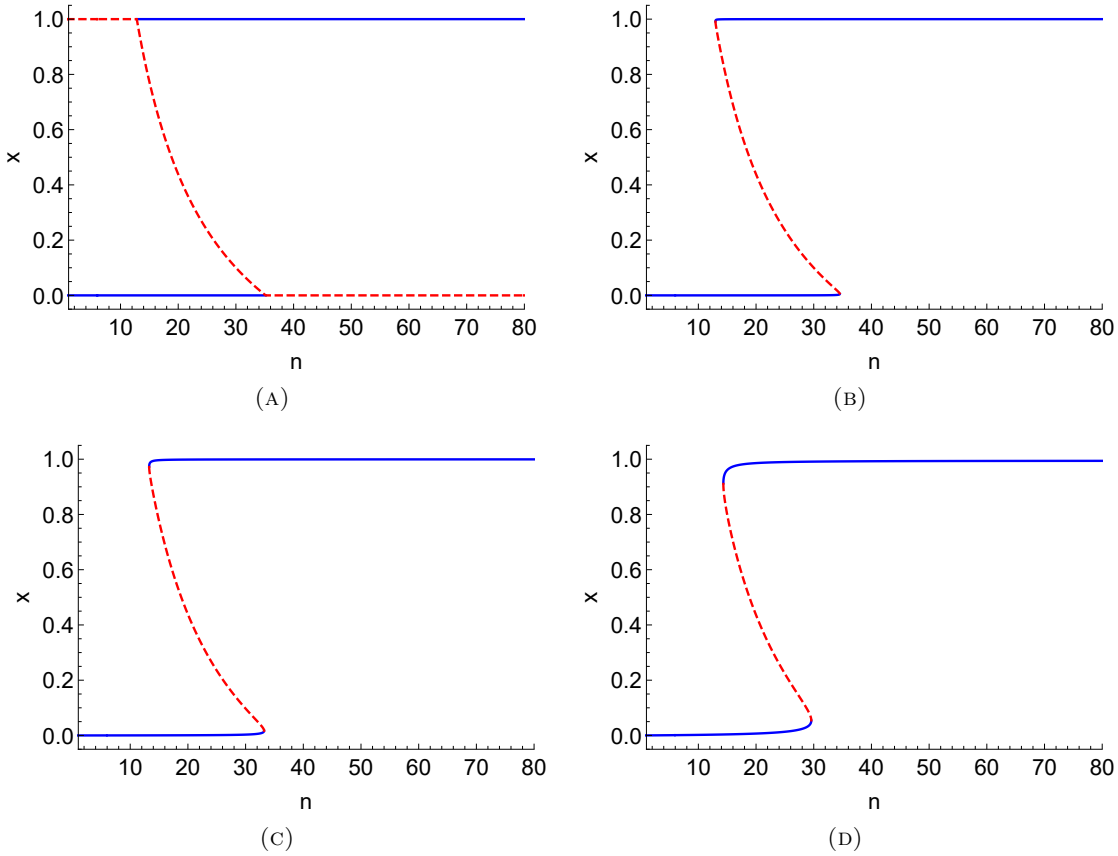


FIGURE 2.10: The effect of exploration on group size bifurcations for the pairwise comparison process: $\delta, \epsilon > 0$ (A),(B),(C) and (D): The solutions of the selection-mutation equation largely resemble the adjusted replicator equation. In (A),(B),(C) and (D) $\mu = 0.001, 0.01$ and 0.1. Even in (D) where $\mu = 0.1$, the dynamics are bi-stable like the replicator equation. Param values: $\gamma_C = 1.92, \gamma_D = 1.9, \epsilon = 0.6, p_0 = 0.4, p_1 = 0.4, k = 0.6$

The effects of the parameter ϵ

As shown in section 2.3.3, when the unstable interior solution of the replicator dynamics x_3, equals one half, equation (2.30) is satisfied and $x_3(\mu) = \hat{x} = 1/2$ is invariant with μ. Moreover, the solution at $\mu = 1/2, x = 1/2$ is always stable for the dynamics. Thus, there exists some value μ_{tb}, such that for $\mu < \mu_{tb}$ the solution $x_3(\mu) = 1/2$ is unstable, and for $\mu > \mu_{tb}$ the solution branch is stable. The change in stability of $x_3(\mu)$ occurs through a codimension 1 transcritical bifurcation where $x_3(\mu)$ and $A_x(x, \mu) = 0$ intersect. A unique value of ϵ corresponding to the bifurcation is found by solving $A_\mu(x, \mu) = 0$ and denoted by ϵ^*, i.e., ϵ^* ensures that $x_3(0) = x_3(\mu) = 1/2$. The other solution branch that collides with, and transfers stability to $x_3(\mu)$ depends on the sign cases of δ, ϵ (see Table 2.1.) In either case, the transcritical bifurcation is structurally unstable: when $\epsilon \neq \epsilon^*$ fold bifurcations and continuity curves occur. We use a scalar parameter unfolding of the system with ϵ as an unfolding parameter to capture qualitatively different dynamical behaviour.

We first consider the structural trade-off regime of $\delta, \epsilon > 0$, as illustrated by case 2 of Fig 2.1. Fig 2.11a (guided with arrows) and Fig 2.11c (same qualitative behaviour) compare phase portraits for $\epsilon = \epsilon^*$ and $\epsilon \lneqq \epsilon^*$. In this regime, $x_3 = x_3(0) = 1/2$ is unstable, and $x_1 = x_1(0) = 0$

and $x_2 = x_2(0) = 1$ are stable. From analysis of section 2.3.3 we know that $x_1(\mu)$ initially increases with μ and $x_2(\mu)$ initially decreases with μ. In the case of $\epsilon = \epsilon^*$ the solution $x_3(\mu)$ is unstable for $\mu < \mu_{tb}$. Stability is transferred to $x_3(\mu) = 1/2$ at μ_{tb} through a transcritical bifurcation in which $x_3(\mu)$ collides with $x_2(\mu)$. At a slightly larger value of μ_f, the solutions $x_2(\mu)$ and $x_1(\mu)$ collide and disappear through a fold bifurcation. For a small range of $\mu \in (\mu_{tb}, \mu_f)$ the dynamics are bi-stable with both $x_3(\mu)$ and $x_1(\mu)$ stable and separated in phase space by $x_2(\mu)$. For $\mu < \mu_{tb}$ the dynamics are bi-stable in $x_2(\mu)$ and $x_1(\mu)$ separated by $x_3(\mu)$. The case of $\epsilon > \epsilon^*$ results in a fold bifurcation at μ_f where $x_2(\mu)$ and $x_3(\mu)$ collide (in the $x > 1/2$ region of phase space), and a continuity curve $x_1(\mu)$ which remains stable. The dynamics are bi-stable in $x_1(\mu)$ and $x_2(\mu)$ separated by $x_3(\mu)$ for $\mu < \mu_f$. When $\epsilon < \epsilon^*$ the fold bifurcations occurs where $x_1(\mu)$ and $x_3(\mu)$ collide, and $x_2(\mu)$ remains stable. The dynamics are bi-stable in $x_1(\mu)$ and $x_2(\mu)$ separated by $x_3(\mu)$ for $\mu < \mu_f$.

The results for case 4 of the trade-off; $\delta, \epsilon < 0$, are shown in Fig 2.11b and Fig 2.11d. When $\epsilon = \epsilon^*$ a transcritical bifurcation occurs at μ_{tb} where stability is transferred from $x_2(\mu)$ to $x_3(\mu) = 1/2$. At a larger $\mu_f > \mu_{tb}$ a fold bifurcation occurring in the upper-half plane removes the $x_1(\mu)$ and $x_2(\mu)$ from the system. For a range of $\mu_{tb} < \mu < \mu_f$ dynamics are bi-stable with $x_2(\mu)$ and $x_3(\mu)$ separated in phase space by $x_1(\mu)$. For $\mu < \mu_{tb}$ the solutions $x_1(\mu)$ and $x_2(\mu)$ are stable and separated by $x_3(\mu)$. For $\mu > \mu_f$ an equal frequency of each strategy is the only stable state. For the case of $\epsilon < \epsilon^*$ a pair of fold bifurcations replace the transcritical bifurcation: at μ_{f_1} the solutions $x_1(\mu)$ and $x_3(\mu)$ leave phase space, and at μ_{f_2} the solutions $x_1(\mu)$ and $x_3(\mu)$ bifurcate back into phase space. A further fold bifurcation occurs at μ_{f_3} at which $x_2(\mu)$ and $x_3(\mu)$ collide and exit phase space. For $\epsilon > \epsilon^*$ a fold bifurcation between $x_2(\mu)$ and $x_3(\mu)$ occurs, and $x_1(\mu)$ remains stable.

For both of the above cases we see that decreasing ϵ is favourable to the players that adopt strategy C: the solution branch $x_2(\mu)$ is stable for greater regions of μ meaning that the population state containing the highest frequency of C strategists is stable for greater exploration rates. This finding is seen by comparing $\epsilon^* < \epsilon$ to $\epsilon < \epsilon^*$ in either of Fig 2.11a or Fig 2.11b. Concurrent with decreasing ϵ, the region of μ where $x_1(\mu)$ (the population supporting the greatest proportion of D strategies) is stable contracts. Increasing ϵ lowers the fitness of the C strategy and the region of μ where $x_2(\mu)$ is stable contracts as ϵ increases. The region of μ for which $x_1(\mu)$ is stable expands.

Notice that all the while $x_3(0) \in (0, 1)$ bifurcations are inevitable because the initially unstable $x_3(\mu)$ must either turn stable or leave phase space for $\mu \to 1/2$.

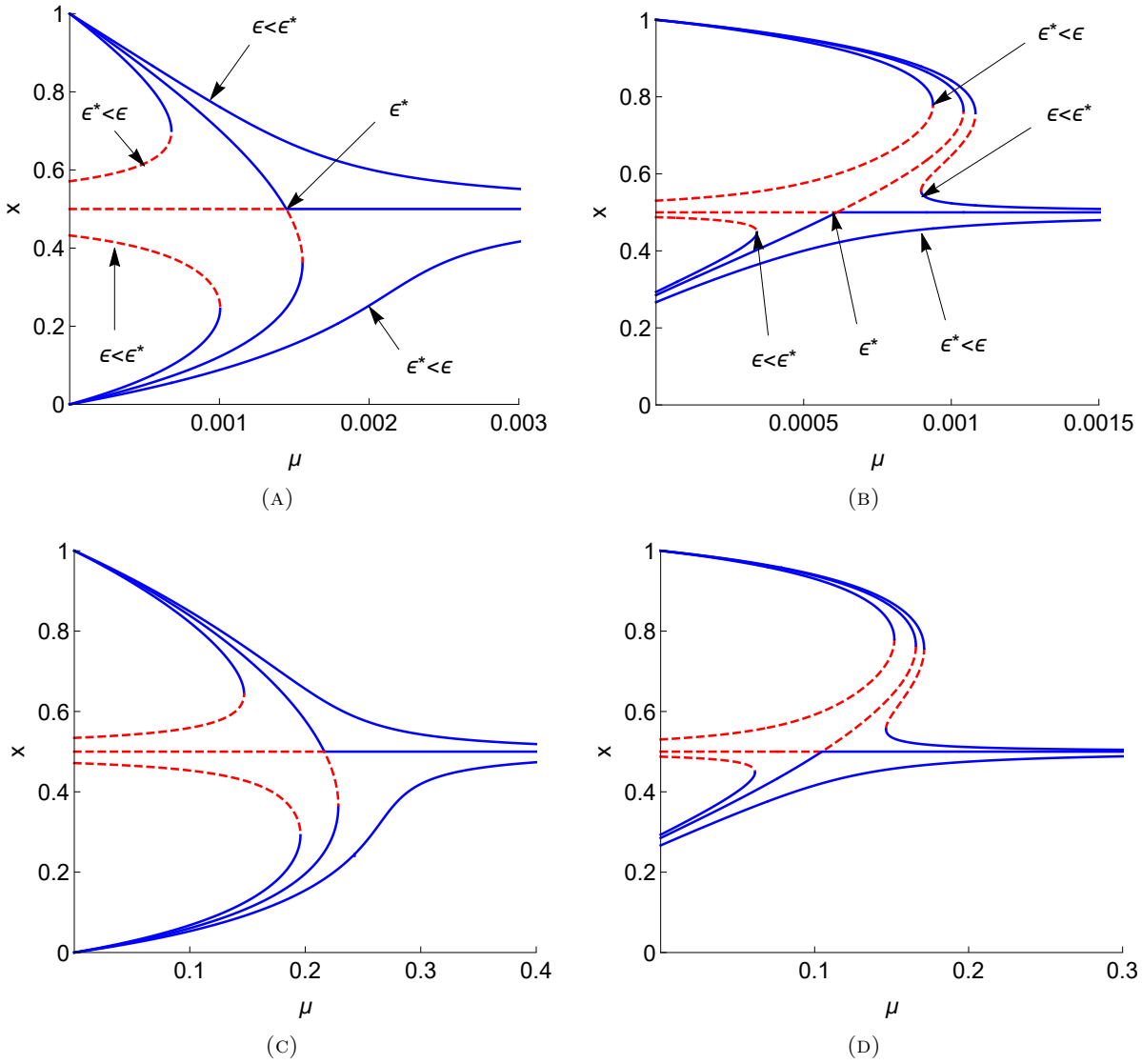


FIGURE 2.11: Unfolding the transcritical bifurcation. (A),(C) when $\delta, \epsilon > 0$ and (B),(D) for $\delta, \epsilon < 0$. In all plots the transcritical bifurcation occurs in the parameter regime of $\epsilon = \epsilon^*$. (A): The transcritical bifurcation occurs at $\mu \approx 0.0015$ in which stability is transferred from $x_2(\mu)$ to $x_3(\mu) = 1/2$. For $\epsilon \gtrless \epsilon^*$ the phase space is split between the stable solution branches of $x_1(\mu)$ and $x_2(\mu)$ separated by the unstable $x_3(\mu) \neq 1/2$ with fold bifurcations and continuity curves replacing the transcritical bifurcation. (B): The transcritical bifurcation occurs at $\mu \approx 0.0007$. For $\epsilon < \epsilon^*$ two fold bifurcations, and for $\epsilon > \epsilon^*$ fold bifurcation and continuity curve, replace the transcriticalbifurcation. Parameter values (A),(C): $\gamma_C = 1.92, \gamma_D = 1.9, p_0 = 0.4, p_1 = 0.4, k = 0.6$, (B),(D): $\gamma_C = 1.9, \gamma_D = 1.92, p_0 = 0.3, p_1 = 0.6, k = 0.1$.

Extension: two time scale oscillatory behaviour for dynamic μ

Section 2.3.3 illustrates the qualitative differences in the resulting dynamics of the frequency of C strategists that arise from the master equation expansion for $\mu \neq 0$, compared to the standard replicator equation (2.7) that is recovered in the limit of $\mu = 0$. These results are important because in animal behaviours, the propensity for an individual to explore its strategic choices may be relatively high [21]. In this section we are motivated to consider the longer term dynamics on a timescale where $\mu \neq 0$ and $\dot{\mu} \neq 0$. We consider the possibility that the rate of strategic

exploration varies according to the dominance of strategies within the population. The idea is that in the absence of a dominant strategy within the population, individuals are less likely to sample alternative strategies and μ should be small. Alternatively, if some strategy is dominant in the population, then one expects each individual in the population to be more willing to explore its alternative options, and μ to be high.

As illustrated in Fig 2.11 for a range of μ, two relatively dominant and stable solutions can co-exist independent of one another, separated by an unstable mixed equilibrium. Here, we utilise such regimes of bi-stability to enable a fast switching behaviour in the underlying frequency of x, to capture the idea of non-constant strategic exploration. We introduce a feedback control for the parameter μ. The closed loop fast-slow dynamics are given by

$$\dot{\mu} = -K\mu(1-\mu)(x-\beta_1)(\beta_2-x) =: g_1(\mu,x) \tag{2.37a}$$

$$\dot{x} = \mu\frac{(1-x)f_D(x) - xf_C(x)}{\langle f\rangle} + \frac{(1-x)x(f_C(x)-f_D(x))}{\langle f\rangle} =: g_2(\mu,x) \tag{2.37b}$$

where $\beta_1 \in (1/2,1)$ and $\beta_2 \in (0,1/2)$ and $K << 1$ represents the separation between the timescales of $\dot{x}$ and $\dot{\mu}$. As $K \to 0$ the exploration rate becomes constant (i.e., does not change with time), and the dynamics are those of equation (2.5a) in section 2.3.3. The dynamics of (2.37) imply that in the absence of either strategy being dominant, the rate of exploration decreases so that the population moves away from $x = 1/2$, whereas, if either of the strategies dominate then μ increases and the population moves towards an equal abundance of each strategy.

The coupled dynamics are invariant on $\Lambda := \{(x,\mu) : x \in [0,1] \text{ and } \mu \in [0,1]\}$, with boundary bd$\Lambda$ and interior intΛ. A solution given by $\mathbf{x}^* = (\mu^*, x^*) = (1/2, 1/2)$ is always valid. Additionally, there exist boundary solutions defined by $\mu = 0$ and $\hat{x}_i(0)$. By the implicit function theorem, any solution to the $\dot{x}$-dynamics can be locally represented as a graph, $\tilde{\mu}(x)$. The lines defined by $x = \beta_1$ and $x = \beta_2$ belong to the nullclines of $\dot{\mu}$ and intersect with the solution curves $\dot{x} = 0$ in the $\mu - x$ plane to define the interior fixed points, i.e., $\mathbf{x}^{\text{int}} \in \text{int}\Lambda$. Motivated by the bifurcation analysis of section 2.3.3 we consider the cases in which either of $x = \beta_1, \beta_2$ intersect $\dot{x} = 0$ at a degenerate fixed point of the x-dynamics. We denote these non-hyperbolic equilibria by $\mathbf{x}^H$. Fig 2.12 illustrates these cases of interest. Here we seek to examine the local dynamics about $\mathbf{x}^H$ in more detail. We focus on the case of $\delta, \epsilon < 0$, but the case of $\delta, \epsilon > 0$ produces similar results.

Consider the case of either fold bifurcation of $\dot{x}$ in the upper half of the plane shown in Fig 2.12 and suppose that $x = \beta_1$ of intersects the critical manifold of $\dot{x} = 0$ at this point. Refer to the x-coordinate as $\beta_1^H =: x_H$. The coordinate for μ can be found through $\mu_H = \tilde{\mu}(x_H)$ so that the interior solution of equation (2.37) is given by $\mathbf{x}^H = (\mu_H, x_H)$. Straightforward analysis shows that at $\mathbf{x}^H$, the vector field $V = (\dot{\mu}, \dot{x})$ lies on a zero-contour, i.e., $\text{div}V|_{\mathbf{x}^H} = 0$. To see this, we use

$$\text{div}V = \frac{\partial}{\partial\mu}g_1(\mu,x) + \frac{\partial}{\partial x}g_2(\mu,x). \tag{2.38}$$

It is clear from

$$\frac{\partial}{\partial\mu}g_1(\mu,x) = K\mu(x-\beta_1)(\beta_2-x) - K(1-\mu)(x-\beta_1)(\beta_2-x) \tag{2.39}$$

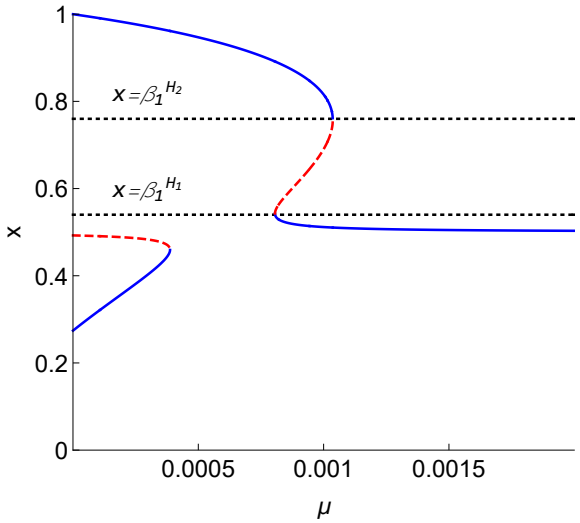


FIGURE 2.12: Illustration of $\mathbf{x}^H$. (A): One of the nullclines of $\dot{\mu}$ shown by $x = \beta_1$ (black dotted line) intersects the critical manifold of $\dot{x} = 0$ at degenerate fixed point. Two cases are shown by $x = \beta_1^{H_1}$ and $x = \beta_1^{H_2}$

that for $x \in (\beta_2, \beta_1)$, $\partial_\mu g_1(\mu, x) < 0$, and for $x \in (\beta_1, 1)$ or $x \in (0, \beta_2)$, $\partial_\mu g_1(\mu, x) > 0$. Moreover, when $x = \beta_1, \beta_2$ then $\partial_\mu g_1(\mu, x) = 0$. The $\partial_x g_2(\mu, x)$ term is precisely $A_x(x, \mu)$ from section 2.3.3. By assumption of the fold bifurcation we know that

$$A_x(x_H, \mu_H) \equiv \partial_x g_2(\mu, x)|_{x=x_H=\beta_1^H} = 0. \tag{2.40}$$

Combining the results of (2.39) and (2.40), we see that at $x = \beta_1^H, \mu = \mu_H$ the divergence is equal to zero. Moreover, we can always find a point, $x' = x_i(\mu)$, along $\dot{x} = 0$ such that $x' \neq x_H = \beta_1^H$ and $\partial_x g_2(\mu, x)|_{x=x'} > 0$. If $x' > x_H$ then $\partial_\mu g_1(\mu, x) > 0$ at x' by equation (2.39) and the divergence is positive at $(\hat{\mu}(x'), x')$ and negative at $(\hat{\mu}(x''), x'')$ for $x'' < x_H$. Thus, the divergence of V changes sign at $\mathbf{x}^H$.

The Jacobian matrix of the system is

$$\mathbf{J} = \begin{pmatrix} \frac{\partial}{\partial \mu} g_1(\mu, x) & \frac{\partial}{\partial x} g_1(\mu, x) \\ \frac{\partial}{\partial \mu} g_2(\mu, x) & \frac{\partial}{\partial x} g_2(\mu, x). \end{pmatrix} \tag{2.41}$$

with eigenvalues $\lambda_\pm = (\text{tr}\mathbf{J} \pm \sqrt{(\text{tr}\mathbf{J})^2 - 4\text{det}\mathbf{J}})/2$. At the interior fixed points of interest; $\mathbf{x}^H$, by assumption, $\partial_x g_2(\mu, x) = 0$ and equation (2.39) shows that $\partial_\mu g_1(\mu, x) = 0$. At $\mathbf{x}^H$, one finds that $\mathbf{J}$ only has off diagonal entries, and eigenvalues at $\mathbf{x}^H$ have the form $\lambda_\pm = \pm\sqrt{-\text{det}\mathbf{J}}$. Moreover, for $x_H = \beta_1^H \in (1/2, 1)$ one finds that

$$\frac{\partial}{\partial x} g_1(\mu, x) = -K(1 - \mu)\mu(\beta_1 + \beta_2 - 2x) \tag{2.42}$$

is positive. Alternatively if $x_H = \beta_2^H \in (0, 1/2)$ then equation (2.42) is negative. All that remains to determine the sign of $\frac{\partial}{\partial \mu} g_2(\mu, x)$ evaluated at $x_H = \beta_1^H$. Using the analysis of section 2.3.3 and in particular equations (2.30) and (2.34), we know that while $x_H < \hat{x}$ the partial derivative is positive, and while $x_H > \hat{x}$ the opposite holds.

Combining the all of the above conditions, we find that when $x_H < 1/2$, $\partial_x g_1(\mu, x)|_{x_H} < 0$ and $\partial_\mu g_2(\mu, x)|_{x_H} > 0$, and if $x_H > 1/2$ then $\partial_x g_1(\mu, x)|_{x_H} > 0$ and $\partial_\mu g_2(\mu, x)|_{x_H} < 0$. In both cases, at $\mathbf{x}^H$, the Jacobian becomes

$$\mathbf{J}|_{\mathbf{x}^H} = \begin{pmatrix} 0 & J_{12} \\ J_{21} & 0 \end{pmatrix} \tag{2.43}$$

where $J_{12} := \partial_x g_1(\mu, x)|_{x_H} > 0$ and $J_{21} := \partial_\mu g_2(\mu, x)|_{x_H} < 0$, such that $\det\mathbf{J}|_{\mathbf{x}^H} > 0$ at $\mathbf{x}^H$, and $\mathrm{tr}\mathbf{J}|_{\mathbf{x}^H} = 0$. The eigenvalues of $\mathbf{x}^H$ are complex conjugate given by $\lambda_\pm = \pm i\sqrt{J_{12}J_{21}}$.

Thus, we can state the following: when the solution branch, $x(\mu)$, of the x-dynamics undergo a one-dimensional fold bifurcation, then the eigenvalues of (2.37) are purely imaginary. That interior fixed points of (2.37), can be expressed through $x = \beta_1$ and $\mu = \tilde{\mu}(x) = \tilde{\mu}(\beta_1)$, and the eigenvalues $\lambda_\pm$ are continuous with β_1. Moreover, at the interior solutions $\mathbf{x}^H$, the trace of the Jacobian vanishes ($\mathrm{tr}\mathbf{J} = 0$) and the determinant of the Jacobian is positive ($\det\mathbf{J} > 0$). These conditions are necessary and sufficient for the existence of a Hopf bifurcation.

Importantly, in section 2.3.3 (when μ is fixed) we identify multiple conditions on the stability and physicality of the solutions of the replicator dynamics, which result in inevitable fold bifurcations when $\mu > 0$ is used as a bifurcation parameter: $0 < \mu < 1/2$. This is an important observation because it means that in situations in which there is a trade-off between cooperating and defecting (i.e., $\delta, \epsilon < 0$ or $\delta, \epsilon > 0$) and the propensity of individuals to explore their strategic choices changes with the dominant frequencies in the population, then it is possible that oscillatory dynamics are inevitable. The oscillations are the result of a Hopf bifurcation which we now examine more closely.

Hopf Bifurcation

Hopf bifurcations can be supercritical or subcritical. In the supercritical case, stable limit cycles bifurcate into phase space, whereas in the subcritical case, the bifurcation gives rise to unstable limit cycles. The periodic orbits are restricted to some vicinity of the bifurcation point, and generally lose structural integrity the farther they move from this point. The regimes of supercritical and subcritical bifurcations are distinguished by the first Lyapunov coefficient, l_1, [72]. l_1 captures nonlinear effects at the bifurcation point: although the eigenvalue stability criterion is neutral, up to second order $\mathbf{x}^H$ may still be attracting ($l_1 < 0$). In this event the bifurcation is supercritical and a stable limit cycle appears when $\mathbf{x}$ becomes stable ($\beta_1 < \beta_1^H$). If $\mathbf{x}^H$ is repelling (to second order) then $l_1 > 0$ and an unstable limit cycle appears for $\beta_1 > \beta_1^H$.

Using $x = \beta_1$ as the bifurcation parameter we see that the bifurcation is subcritical if $l_1(\beta_1^H) > 0$ and supercritical if $l_1(\beta_1^H) < 0$, where $l_1(\beta_1)$ is found following Kuznetsov [248]:

$$l_1(\beta_1) = -\frac{1}{4\omega^3}\left\{K(\mu-1)\mu_H\Big((\beta_2-\beta_1)L_1L_2 + L_3\big((\beta_1-\beta_2)L_4 + 2L_2\big)\Big)\right\} \tag{2.44}$$

where $\omega = \sqrt{\det\mathbf{J}} = \sqrt{J_{12}J_{21}} > 0$ and

$$L_1 = \frac{\partial^3}{\partial x^3}g_2(\mu, x)\Big|_{\mu=\mu_H, x=\beta_1}, L_2 = \frac{\partial}{\partial \mu}g_2(\mu, x)\Big|_{\mu=\mu_H, x=\beta_1}, L_3 = \frac{\partial^2}{\partial x^2}g_2(\mu, x)\Big|_{\mu=\mu_H, x=\beta_1} \tag{2.45a}$$

$$L_4 = \frac{\partial^2}{\partial x \partial \mu}g_2(\mu, x)\Big|_{\mu=\mu_H, x=\beta_1} \tag{2.45b}$$

It is possible to provide an analytical expression for l_1, but due to its complexity we do not. The signs of coefficients L_i in equation (2.44) could also be deduced by studying the convexity of the fitness functions $f_C(x)$ and $f_D(x)$ along with the location of $\mathbf{x}^H$ in phase space. In all numerical investigations we find that $l_1 < 0$: all bifurcations result in a stable limit cycle introduced to phase space. Fig 2.13 shows the Hopf bifurcation for the example of an unstable solution branch of the $\dot{x}$-dynamics enclosed by two fold bifurcations (shown in Fig 2.12). For $\beta_1 < \beta_1^{H_1}$ the interior fixed point is a stable focus (first panel). At $\beta_1^{H_1}$ a stable limit cycle bifurcates into phase space, and exists for $\beta_1 \in (\beta_1^{H_1}, \beta_1^{H_2})$ (panels 2-5). At $\beta_1^{H_2}$ the limit cycle disappears and for $\beta_1 > \beta_1^{H_2}$ the interior fixed point is a stable focus.

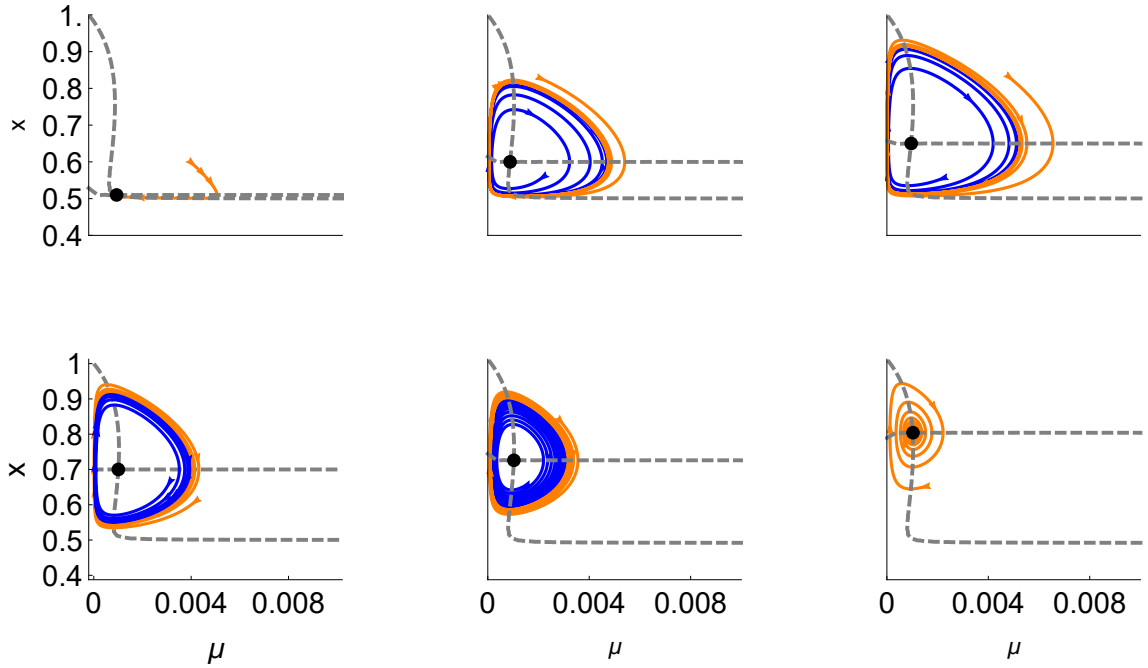
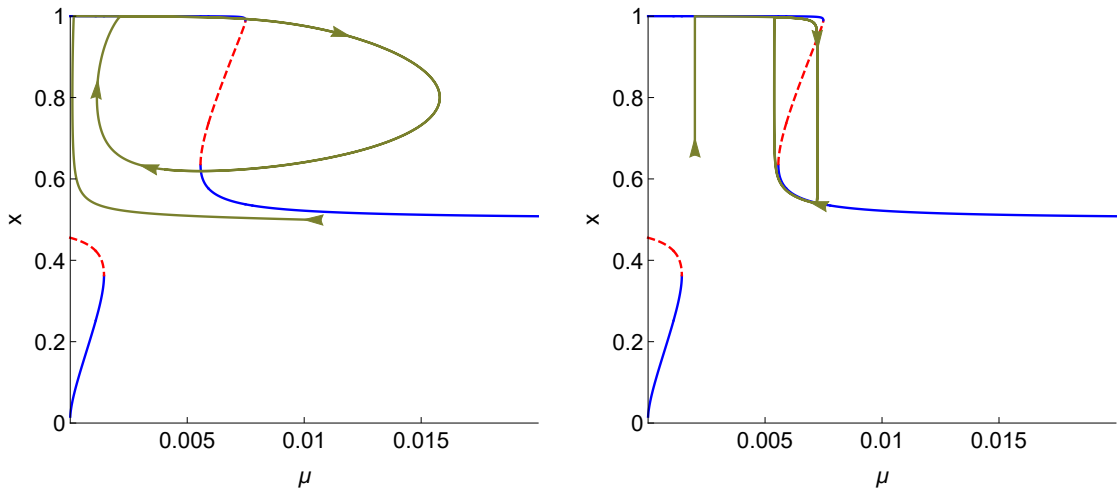


FIGURE 2.13: The formation and disappearance of the limit cycles. Each panel depicts the nullcines of $\dot{x} = 0$ and $\dot{\mu} = 0$ by the grey dashed lines which intersect at the interior fixed point. Each respective value of $\beta_1 = 0.51, 0.6, 0.65, 0.7, 0.725, 0.8$ determine the fixed point. In the cases where limit cycles exist two solution trajectories are shown, orange with initial conditions outside the closed orbit, and blue inside. Fold bifurcations in the $\dot{x}$-dynamics occur at roughly $x = 0.52$ and $x = 0.78$, i.e., $\beta_1^{H_1} \approx 0.52$ and $\beta_1^{H_2} \approx 0.78$. When $\beta_1 \in (1, 0.78)$ or $\beta_1 \in (0.5, 0.52)$ the fixed point is a stable focus and when $x \in [0.51, 0.78]$ the solution is unstable and surrounded by a stable limit cycle.

The oscillations correspond to periodic switches between the relatively mixed equilibrium and a dominant solution. The nature of oscillatory behaviour depends on the timescale separation parameter K. In the limit of $K \to 0$ the limit cycles cling to the critical manifold of $\dot{x} = 0$ because x changes much faster than μ. Fig 2.14 provides a geometric illustration of this. Along the dominant x branch, μ is increasing and the limit cycle approaches the upper fold singularity. There, jump downwards to the relatively mixed frequency solution, where μ is decreasing. Thus, the cycle approaches the vertical line defined by $\mu = 0$. Here, the cycle jumps back to the dominant branch of x, and the process starts again. As K increases this process becomes less defined, and the cycles diverge further from the stable solution branches of $\dot{x}$.



(A) $K = 0.2$ and $\delta, \epsilon < 0$ (B) $K \to 0$ and $\delta, \epsilon < 0$

FIGURE 2.14: The timescale separation. (A): For $K = 0.2$ the changes in x occur on a timescale slightly faster than the change in μ. The limit cycle (green) has large amplitude. (B): For $K = 0.0001$ the exploration rate μ changes on a much slower timescale and the limit cycle clings to the stable solutions of $\dot{x}$ up to singularities.

2.4 Discussion

We have found that the variety of behaviours exhibited by grouping prey in nature can only be reproduced when vigilance, confusion and dilution are all modelled together. We find that the potential for transitions from population states characterised by all individuals choosing the cooperative anti-predator behaviour, or all choosing to defect, to a mixture of each behaviours occur only when benefits from grouping are traded-off against all three of vigilance, confusion and dilution effects. Because both costs and benefits are dependent on the ecological and environmental contexts, our model is able to explain the manner in which grouping individuals change their behaviour in response to external changes. Only when there is a trade-off between anti-predatory behaviour and grouping benefits do we find ecologically relevant cases. This predicts two distinct behavioural phenotypes for grouping animals: the first where contributing to overall group evasion leads to greater group benefits but greater risk of predation ($\epsilon, \delta > 0$), and the second where contributing to group safety leads to decreased predator targeting at the expense of diminished grouping benefits ($\epsilon, \delta < 0$) as illustrated in Fig 2.1. In both cases we predict unique population attractors for each configuration which are dependent on model parameters.

General conclusions for $\delta, \epsilon < 0$

In the absence of behavioural exploration our model predicts three potential population states. Namely either of the two pure population states; all play the strategy C, or all play D, and a mixed population state that describes a co-existence of cooperative and defector behavioural phenotypes. Here we draw conclusions in the case of perfect behavioural decision making ($\mu = 0$).

For the case of $\delta, \epsilon < 0$ we identify a negative group size effect: the frequency of the cooperative behavioural type decreases with increasing group size. This result is in line with most theoretical

[32] and observational studies [18, 127, 136, 191] suggesting that vigilance should decrease with group size. The group size effect we find can be characterised by both an increase in the basin of attraction of the stable population state sustaining only D strategists, and a decrease in the frequency of the C strategy at equilibrium, with an increase in group size (Fig 2.2a). This regime is characterised by three critical group sizes. When group size exceeds a critical value, n_2, choosing the defect strategy is the only viable behaviour for a group member: all group members have a sufficiently high chance of survival, and there is little incentive to contribute to predator evasion. Essentially, the incentive to play the strategy we term cooperate is decreasing with the number of other individuals playing cooperate in the group, and with group size. With increasing group size, the fitness of those individuals playing the strategy which we term defect increases relative to the fitness of those individuals choosing to play the cooperate strategy.

For an intermediate range of group sizes, we find that there is a possibility for two stable population states independent of each other: either of a polymorphic or pure population state of only D strategies, and a population state of only C strategies (as shown in Fig 2.2a). The bi-stable regime is separated in phase space by an unstable polymorphic population state and the population dynamical attractor depends on the initial fraction of individuals playing the strategy C. Polymorphic populations or a population of only C strategists occur when group size lies in a range given by $n \in (n_3, n_1)$, and stable population states of either all C strategies or all D strategies occur if $n \in [n_1, n_2)$, where n_1 is not necessarily less than n_2. Notice that the values of the critical group sizes characterise the properties of the population dynamical attractors. As such, it is possible that the only the polymorphic population state is stable, i.e., if $n_1 > n_2$. The interpretation of a polymorphic population in which a fraction, x_4, play pure strategy C and $1-x_4$ play D is equally valid for a mixed strategy mixed equilibrium in which each individual plays the strategy C with probability x_4 and plays D with probability $1 - x_4$.

When groups are smaller than the critical size n_1 the only stable population state is characterised by all individuals choosing to play the cooperate strategy. The anti-predation fitness benefits of preferable risk dilution ($\epsilon < 0$) that individuals playing the C strategy receive increase with decreasing group size: advantages of not being picked are more pronounced when there are less to chose from. With decreases in group size, the fitness an individual choosing to cooperate increases relative to an individual choosing the D strategy.

The change in the population frequency of C strategists with group size is modulated by the effects of confusion, dilution and vigilance. All group members experience the same safety of group evasion, and as grouping benefits are independent of group size. Thus, as group size changes, unequal dilution of risk is the only factor that causes a nonlinear rate of change for the fitness difference between the individuals playing C and those playing D. The smaller the group the greater the magnitude of effect $\epsilon < 0$: as group size decreases the individuals playing strategy D become relatively more vulnerable through increased predator targeting. Therefore, with decreasing group size, the relative fitness differential between the two strategies increases in favour of playing C. Through their respective effects on overall predator evasion, confusion and vigilance effects specify the rate at which the fitness differential changes with group size. The more probable the group as a whole escapes, the less influence dilution of risk has on individual fitness. In this sense the probability of group evasion controls the absolute fitness differential for any group size. At

the individual level, the effects of vigilance and confusion thereby contribute to which strategy is chosen.

General conclusions for $\delta, \epsilon > 0$

Contrastingly, for the case of $\delta, \epsilon > 0$ as described in Fig 2.2b, we have shown a positive group size effect (an increase in the basin of attraction of the stable population state sustaining only C strategists) on the behavioural phenotypes which contribute to anti-predator evasion in prey groups. In particular, we show that when group size exceeds a critical group size n_1, the only stable population state is for all individuals to play the strategy which we term cooperate. With increases in group size, the dilution effect offers increases in safety for all group members. The relative fitness advantages that individuals playing strategy D receive (through being less likely targets) diminish with increasing group size. For very large groups, the differences in the dilution of risk between individuals playing each strategy become negligible. As such, the relative magnitudes of fitness components change with increasing group size in a way that general grouping benefits become more important and group size effects less so. Since those individuals who contribute to predator defences (C strategists) receive greater general benefits than those who play D, there will exist some threshold group size for which it is advantageous to play strategy C. Note that this finding does not suggest that the strategy which contributes towards predator evasion is chosen because it provides safety benefits for all group members. It could be the upshot of by playing C an individual is afforded more favourable general grouping benefits.

For small groups of size $n < n_2$ (where $n_2 < n_1$) the only stable population state is for no individuals to play strategy C. This finding is explained by the opposite argument of that above: the magnitude of the effect of ϵ on overall risk profile is decreasing with group size. Thus, in smaller groups, the difference in risk dilution experienced by an individual who contributes to group evasion (a C strategist), compared to the risk dilution experienced by an individual who does not (a D strategist), is greater. The greater the difference in dilution risk between the strategies the greater the difference in fitness, since other grouping benefits are independent of group size.

For a range of intermediate group sizes, $n \in (n_2, n_1)$ we find that the population state is described by a region of bi-stability, with both population states characterised by either all individuals playing strategy C or D locally stable. When group size is within this range, the final population state is governed by the initial frequency of the C strategy. In these scenarios, the preferable behaviour depends on the expected number of C strategies in the group. For example, favourable risk dilution may provide those individuals playing defect respectively higher fitness only when there is some threshold number of individuals who contribute to predator evasion, i.e., a certain number of C strategists within the group. For $n_2 < n < n_1$ the stable population state reached depends on the expected initial composition of each group.

At the individual level, the behavioural transition from defector to cooperative strategies with group size, can be explained by the interplay between general benefits of grouping and how the effects of confusion, dilution and vigilance moderate predation risk. If the group cannot evade attack (through vigilance or group size effects), then risk dilution has a large effect on an individuals safety. In this case, by adopting the strategy that reduces personal targeting risk, a group member gains significant increases in safety and thereby fitness. If it is extremely likely that a predators

attack fails, then dilution effects are negligible, and the risks of personal targeting do not have such influence on fitness. Under our modelling of dilution, the survival probabilities for each strategy increase and (eventually) converge with increasing group size. But the general grouping benefits are constant and consistently favour those individuals playing the cooperate strategy over those playing defect. It is therefore no surprise that there exists a threshold group size (i.e., $n^* \in (n_2, n_1)$ for which the cooperate strategy becomes risk dominant, and a further group size n_1 for which the only viable behaviour in the population is to play strategy C. Hence, for $\delta, \epsilon > 0$ we find that the cooperate strategy is subject to an Allee type effect.

The mathematical system helps elucidate how certain grouping animal species become extinct or suffer depleted populations. As we vary environmental parameters such as habitat availability or other external parameters, the changes in stability in the dynamical system directly correspond to changes in behaviour at the population level. Thus effectively explaining the effects of external factors on decision making. If the viability of certain behaviours are incompatible with the animal species specific requirements, then this may have severe consequences. For example, consider that a population consisting of groups which is in an ecologically stable behavioural state. Following external environmental changes the stable configuration may abruptly switch and if the species is not sufficiently behaviourally plastic it will become ecologically unviable.

The importance of strategy exploration

Our results indicate that when strategy exploration is small but non-negligible ($\mu \neq 0$), stochasticity at the level of an individual's decision making augments the resulting dynamics. However, the core qualitative dynamical features of the case of no exploration are retained. For example, while individuals possess some small propensity to randomly sample between the two strategic behaviours ($\mu << 1$), equations (2.5a) and (2.5b) produce qualitatively similar results for the composition of the population, when compared with the case of $\mu = 0$ (equation (2.7)). Moreover, when exploration rates are very small, the pure population states are never valid. Instead the population is always characterised by a polymorphic equilibrium state, i.e., a mixture of cooperative and defector individuals. In the limit of large exploration rates, i.e., when $\mu \geq 1/2$, we find equal abundances of each strategy. These results are valid for a polymorphic population state characterised by $x \in (0, 1)$, and a mixed strategy monomorphic population state (where each individual plays strategy C with probability x, and plays the D strategy with probability $1 - x$)

We show that it is important to consider the possibility of imperfect behavioural imitation, especially when modelling animal behaviours. Here we explain why. First, evolutionary game theory in the form of the replicator dynamics implicitly treats group members like rational agents. Yet rather than always adopting the optimal strategy, animals are never fully rational (and in fact neither are humans [40]). Even when individuals possess the behaviours of others' in their own repertoire, most animal species face some cognitive limitations in their ability to adaptively assess and imitate behaviours that would yield a greater fitness than their own [249]. There are likely to be mistakes in the imitation process. The irrationality of, and constraints imposed on an individual's decision making can be captured by including an exploration rate.

Secondly, animals often forage in stochastic environments [250]. Unpredictability of an individual's surroundings may render an appropriate responses to balance state and predation pressures

impossible. As one example of this consider the behavioural responses of grouping individuals when there are visual obstructions [16]. In this case the ability to imitate a better performing conspecific or make preferable decisions may be severely curtailed. Taken to the extreme, if individuals cannot obtain information from others regarding how to appropriately respond to a predator then the assumption that behaviours spread through imitation is flawed, and one would expect purely random strategic decision making. Both of the above points are be captured by $\mu > 0$ in the dynamical system, and indicate the importance of exploration in the probability of adopting any given behaviour.

General connection with other social type dilemmas

One could observe that without including the dilution effect into the payoff structures, the model reflects a type of (nonlinear) public goods game [54, 55, 251]. The resemblance occurs because only one of the strategy types contributes towards the probability that the whole group escapes (yet both strategy types benefit from group escape). Hence our choice of "cooperator" for the behavioural phenotype that contributes towards the predator evasion. What distinguishes our model from general forms of the public goods game (analysed by Archetti & Scheuring [251]) is the way that dilution of risk alters payoff: modelling the effects of detection and dilution interdependently prevents a tragedy of the commons [252] type scenario because in the absence of cooperators, the probability of survival (and fitness) is non-zero due to all defectors facing an equal probability of not being targeted. This fitness safeguard of dilution by numbers thereby changes the payoff structure that is found in standard social dilemmas. If dilution effects are set to zero (i.e., the square brackets in equation 2.4) the model recreates the structure of the public goods game in certain parameter regimes.

Our results also show that because dilution of risk is not equitably shared amongst group members, cooperation need no longer be costly relative to defection. It is possible for the best outcome of the game when all group members adopt the same strategy, is for all cooperate. When all group members adopt the same strategy, cooperation is preferable if cooperators receive greater grouping benefits, i.e., when $\delta > 0$, or if $\delta < 0$, but the increased safety through group evasion outweighs the benefits differential. The model therefore captures the snowdrift-type games [57] analogous to vigilance behaviour, whereby if no group members cooperate the entire group is worse off but when all group members cooperate, each individual has incentives to defect. It would be interesting to consider a threshold number of cooperators required to procure group evasion fitness benefits, as similar to the volunteers dilemma [59].

Since cooperation can be beneficial or costly in terms on an individual's risk dilution, we propose that a cooperators investment into group wide predator evasion is not necessarily performed to directly increase the fitness or others. Adopting the cooperative strategy may simply be a self-serving behaviour that makes individuals less likely to be a target ($\epsilon < 0$), but at the same time improves the fitness of others' through group evasion. A detailed review of the terminology, e.g., "negative pseudo-reciprocity" [228, 253] is beyond the scope here, and we use cooperation to encompass a range of payoff trade-offs. For example, vigilance is a cooperative behaviour [32] and it has been shown that personal vigilance is more advantageous than relying on the vigilance of others. It is quite possible that non-vigilant individuals face a fitness penalties from failing to

detect the cues of their vigilant counterparts [62, 116], suffering late detection of a predator and a delayed response [254, 255]. We encompass scenarios where both cooperation may be positive for both actor and recipient (+,+) and where it benefits the recipient but is costly to the actor. The latter case corresponds to cases in which non-predation group benefits such as foraging gains or the experience of risk dilution favour the defector strategy. We use the term cooperation to encompass a range of scenarios with focus on the public type good cooperators produce of the probability of group evasion.

Specific choices of functions

The functional choices that we present (equations 2.1,2.2) to capture predator evasion are flexible enough to approximate linear, saturating [6, 149] and constant relationships with the number of group members (C strategists) which adopt a proactive anti-predator behaviour. By examining either of equations 2.1,2.2 with $p_1 = p_2 = 0$ the evasion function collapses to the simplest of possible functions; a constant. The probability of a failed attack is then independent of group size and the frequency of the so termed cooperative strategy. A constant group evasion is explained as follows. Consider an additional member joining the group. According to the group vigilance hypothesis [49, 188], each individual in the group should accordingly lower it's own vigilance. If vigilance is independent between groups members, and the combined reduction individual vigilance rates of existing group members, is equal to the additional vigilance from the new group member, then whole group vigilance does not change with size. Poysa [133] for example, found no significant relationship between group size and vigilance in the duck teal *Anas crecca*.

If the only argument of the evasion function is fraction of individuals playing the cooperate strategy, we apply either of equations 2.1 or 2.2 with $p_1 = 0$. In this case the probability that no group member is attacked given that all group members play the cooperate strategy, is determined by p_2 alone. The variable k_2 specifies how sensitive the probability of group evasion is to the expected proportion of C strategies in the group. So long as $p_2/(k_2 + 1) \leq 1$ one can use $k_2 >> 1$ and $p_2 >> 1$ to approximate a linear relationship [128, 129, 132] whereas $k_2 \in (0, 1)$ can reflect saturating returns of increased group safety with an increase in the frequency of the C strategy [6, 130, 149]. We do not explicitly draw a relationship between those group members who adopt a cooperative strategy and the confusion effect (but see [169, 179]) since not enough is known about how the confusion effect operates on the predators cognitive abilities. Instead, we assume that group size affects predator confusion [117, 170, 180] and that confusion effects dominate the n dependent components of equations 2.1 or 2.2. The Holling Type II structure of the function can accommodate linear and concave relationships between group size and confusion dependent evasion. One could also consider a threshold function which would be applicable when the probability of group evasion changes at a critical amount of cooperators, or a critical group size. The extreme being that only one cooperative individual is needed for group protection [7, 8].

Although we do not draw an explicit relationship, with caution, it is interesting to postulate how what we refer to as the cooperative strategy relates to a prey's other behavioural phenotypes. Coordinated behaviours in a group, as one example, can confuse predators and thereby increase the probability that no group member is killed. Examples are found in dragonfly larvae *Aeshna cyanea* [183] and stickleback [179] preying on *Daphnia magna*. Confusion effects are also known to vary on

spatial scales [154]. Denser groups, and parts of groups may be easier for a predator to detect, and receive an increased attack rate. During the targeting stage of an attack, predator attack error may increase with local group density [179]. Our model accounts for group size effects on risk dilution and predator evasion, but does not consider density effects. More generally, complex movement rules of prey are often theorised as adaptive responses to predation [155, 166]. Morrell [164] demonstrates that on a temporal scale, group members are safer by moving towards their nearest neighbour when subjected to short attacks, whereas during longer attacks, prey are safer moving to the densest part of the group. Incorporating spatial structure into the model presented in this research would most likely result in intractable mathematics: the function describing individual risk dilution would become highly nonlinear if group's contain denser subsets.The results of an unequal risk dilution of group members previously found [164, 256], although premised on movement rules, support the results presented in this chapter which models unequal dilution through ϵ: without explicit spatiality.

Final remarks

Previous models have considered that the effects of dilution and detection are not mutually exclusive in producing selective pressures for behaviours in grouping prey [31, 32, 146] however to our knowledge none have explored the possibility that dilution of risk is a function of behavioural strategy. Our model is unique in differentiating how adopting a proactive cooperative behaviour can directly affect a group members relative per-capita risk as compared to not doing so. The consequences of unequal dilution effects will have profound consequences emergent behaviours of groups of prey, and could be factored in to future models predicting such behaviours.

There are limitations and possible extensions of our model which we outline here. The payoffs of section 2.2 simply assume that average fitness in the group monotonically increases with group size because the probability of surviving a predator attack does so. In the Introduction section 1.4.2 we explain how group size can affect the predator attack rate. Our model does not account for this relationship. Predator attack rate may increase with group size [115] or with other factors that are enhanced by increasing group size, for example increased movement [138] or coordination [11]. Thus larger and more conspicuous groups may be targeted more relative to smaller groups. Alternatively, increasing in group size may lead to a reduced attack success rate for the predator [134, 147]. In this respect, we believe that our modelling could be enhanced by considering attack abatement [115] and encounter-dilution [4, 138] effects. By assuming that dilution effects are controlled by a group size dependent attack rate, individual survivorship, and thereby fitness, would not necessarily increase with group size. Attack abatement effects could be most simply incorporated into the model by considering a strategists probability of survival (big curly brackets in equations 2.4) as the product of the conditional probability of surviving given an attack, and the probability that the group is attacked as depending on a group size dependent attack rate $\alpha(n) \in [0, 1]$. This approach would be flexible enough to encompass situations whereby increases in group size strictly lead to increases in predator attack rate [143], when group size negatively affects attack rate [25], and other more complex, e.g., quadratic [25] relationships observed in species.

Another way in which group size can affect individual fitness is mediated through within group competitive interactions (see Introduction section 1.4.5).The group size effect that we find in the

case of $\delta, \epsilon > 0$ may be more pronounced when considering factors such as food density, which often dwindles with increasing group size [124]. In other words, if the difference between general grouping benefits in terms of foraging is a function of group size and the importance of foraging gains increase with group size increases. Group size dependent foraging could magnify the rate at which the payoffs to cooperators increase relative to the payoff to defectors, as group size increases. To capture a more complex group size-fitness relationship, we could investigate group size effects on resource competition [257] by setting $\gamma_i = \gamma_i(n)$. These functions could be (i) monotonic decreasing if increases in group size strictly led to increases in competition and decreases in foraging benefits, (ii) concave with group size if increasing group size initially procures greater foraging rewards but later reduces foraging rewards [258], or (iii) Holling type II if foraging gains asymptote with group size. It could be that foraging benefits monotonically increase with group size: in models of information transfer for example, it is shown that the shared knowledge of resources that benefit group members depends on the cooperative behaviour of individuals that actively search out food, which in turn depends on the size of the group [259, 260]. However in this case there must be some other cost associated with grouping that bounds group size or groups would grow indefinitely.

We show that it is possible for anti-predatory behaviours to change with variables such as group size, predation pressures and other grouping benefits. The model is able to demonstrate how ecological changes may influence the state and size of a group including when viable groups undergo sudden changes. We show that it is important to consider multiple anti-predation effects affecting grouping animals: each anti-predation effect will vary in its effectiveness against different predators. Importantly, disentangling such effects will help understand the appropriate behavioural responses to differing predators [18, 151] and the value of identifying attacking predators [24, 261]. Incorporating individuals and their heterogeneous life histories and preference into this picture remains a challenging problem for the field.

Chapter 3

A field data assessment of the model of chapter 2

Abstract

Mathematical models can be powerful tools in predicting and explaining anti-predator behaviours in animal groups. The purpose of this chapter is to show that the model presented in chapter 2 can provide novel insights into the anti-predator behaviours in wild systems. We present and inform the model of chapter 2, and make a series of explicit predictions based on the vigilance behaviours of the wading bird, redshank *Tringa totanus*. Subsequently, we theoretically predict how vigilance in redshank should change with factors relating to predation risk, as well as climactic conditions. An analysis of observational data elucidates which of the theoretical predictions are qualitatively reflected by the behaviours of redshank. We demonstrate that only when at least two of the three principle anti-predation effects are theoretically considered, can the theoretical model describe the behaviours of redshank. We clarify that a broad theoretical model can capture subtle and complex state dependent variables which influence behaviours in a particular grouping species. Our findings show the importance of theoretical considerations in explaining the anti-predator behaviours in animal groups.

3.1 Introduction

In this chapter, some of the general results of section 2.3.1 are qualitatively examined. The research presented here is motivated to show the importance of the theoretical framework of chapter 2: that the three principal anti-predation effects of group vigilance, dilution and confusion are interdependent forces influencing prey behaviour. As such, we seek to demonstrate that the general theoretical results anticipated in section 2.3.1 are relevant in predicting behaviours in observational systems, and that when any of the three principal anti-predation effects are not theoretically considered, the predicted behaviours do not capture the complex behavioural choices of grouping animals.

In this research the tasks described above are achieved by qualitatively assessing predictions made from the results of chapter 2, with observational data on the wading bird redshank, *Tringa*

totanus. Grouping redshank are well documented to benefit from the effects of vigilance, dilution and confusion whilst under predation [18, 25]. The foraging behaviours of individual birds within flocks are well documented in an environment where they feed within proximity of predatory sparrowhawk, *Accipiter nisus* [1, 18, 25, 123, 151]. Time spent feeding by redshank strikes a compromise between scanning for the predator and foraging [204]: vigilance therefore reflects a trade-off between starvation and predation risk.

In the following research, section 3.2.1 justifies our use of, and section 3.2.2 parametrises, the model of chapter 2 to predict behaviours in redshank flocks. The general results of sections 2.3.1 and 2.3.3, under the parametrisation, produce a series of species specific predictions given in section 3.3. The first prediction of section 3.3 intends to highlight the importance of modelling all three principal anti-predation effects. To show this importance, we provide theoretical predictions when independently removing each of the key anti-predation effects from the theoretical model. The goal here is to contrast the subsequent predictions of each theoretical model configuration and compare these to the data. Another motivation of this chapter is to show that the broad theoretical results presented in chapter 2 are able to describe relatively complex features in grouping animals. Predictions based on some of the key parameters analysed in chapter 2 are made explicit with regards to redshank behaviour. In section 3.4 the observational data on redshank is presented. We use this section to conduct new analysis on existing data sets for the specific purposes of model comparison. We quantify trends in the data, and assess whether such trends are consistent with the predictions of each theoretical model configuration, and each model prediction. This approach shows which of the anti-predation effects are theoretically required for the model to be able to qualitatively predict vigilance behaviours in redshank. We therefore emphasise the importance of considering the multiple effects of vigilance, dilution and confusion as potentially interdependent factors influencing the behaviours in groups of prey.

3.2 Methods

3.2.1 The mathematical model

We apply the model of chapter 2 with general assumptions therein to reflect the behavioural dynamics of vigilance in redshank. As vigilance and feeding are largely exclusive for redshank [18], we assume that a bird's phenotype is sufficiently plastic to produce two distinct and discrete behaviours. Birds either choose to play the "cooperate" (C) strategy of section 2.2 and prioritise predator defence by adopting a vigilant behaviour (henceforth simply "vigilant" birds). These birds contribute to collective detection benefits for the entire group. Otherwise, birds prioritise feeding (henceforth "non-vigilant" birds) and choose to play strategy D. During a predation event, a bird's choice of vigilance behaviour has a marginal effect on it's expected future reproductive success (EFRS) through mortality risk and time spent not foraging. Over the modelling timescale the combination of many predation events give rise to a bird's EFRS [212, 262].

We use equations (2.4) in section 2.2 to specify EFRS. Foraging benefits for vigilant (C) and non-vigilant (D) birds are respectively given by γ_C and γ_D. Predator evasion is given by $g(x, n)$. For more details of the payoff functions see section 2.2. We assume that the frequency of vigilant birds either changes through a pure imitation process outlined in section 2.2.1 which we refer to as $\mu = 0$,

or that small rates of behavioural exploration, $\mu > 0$, affect the imitation process. Exploration is interpreted as random sampling of either of the vigilant or non-vigilant behaviour (animals often sample alternative behaviours which do not maximise fitness [21]), or random behavioural mutations between the vigilant and non-vigilant behaviours. On the ecological timescale, which we suppose to be a winter, equations (2.7) (without exploration) or (2.5b) (with exploration) give equilibrium frequencies of vigilant birds defined by population states.

To justify our use of the model we make the following assumptions. Variation in flock size is not assumed to be a proactive anti-predator behaviour: we adopt the hypothesis that changes in flock size are responsive to longer term cues about risk [263]. We assume instead, that short term cues about a predator give rise to reactive responses of a bird's vigilance. Because we do not consider variable flock size as an anti-predator behavioural response, we do not consider differentiated attack rates for different sized flocks. Nor do we consider the optimal flock size [264]. We assume that flocks may consist of both adult and juvenile birds [1, 265], although juvenile birds often form separate flocks [266]. We do not consider differences in age or sex as factors that affect vigilance.

3.2.2 Parametrising the Model

The data sets used to qualitatively assess the model (introduced in section 3.4) are not used in informing the model parameters. We use previous research on the specific predator-prey system [18, 25, 206, 267] and a more general wider literature of research on wading birds (specified when used) to infer key model parameters. Foraging benefits are informed using general literature on the metabolic demands of wading bird's, and other field data, regarding the quality of food at the site location. The evasion function; $g(x, n)$ of equations (2.4) is fitted by empirically supported inferences from [18, 206, 267]. The parameter ϵ is fitted using simple calculations and guided observations [266].

Parametrising the foraging benefits functions: γ_C and γ_D

We consider that general grouping benefits (γ_i in equations 2.4) are solely and explicitly specified by foraging gains. Following McNamara & Houston [262] we assume that a redshank's EFRS is proportional to surviving a winter. Over the winter we assume there are approximately α predation events. Let Ψ be a function that maps energy to fitness over this period. The effects of one predator attack on EFRS can be expressed by $\Psi^{1/\alpha}$. Guided by Cresswell & Quinn [151] we set $\alpha = 100$ attacks.

We propose that $\Psi(E)$ is sigmoidal with an individual redshank's hourly energetic intake, E [250, 268]. As risk prone and risk averse foraging behaviours can characterise both increasing and decreasing returns to fitness [250, 262, 268], the gains in fitness through $\Psi(E)$ depend on the initial state of satiation E_0 of a bird [262, 269, 270]. Due to harsh weather conditions, starvation is a real possibility for individual redshank on the saltmarsh, and redshank only forage in this habitat when experiencing negative energy budgets [1, 204]. Therefore, we assume that foraging redshank are initially risk prone. Let E^* be the minimum hourly intake to avoid starvation. Before energetic needs are met (for $E < E^*$), a bird's survival probability is low, and each extra amount of energy is highly valuable to a bird (birds are risk prone). As E approaches E^* a bird's survivorship increases

at an increasing rate. When $E = E^*$, energetic needs are balanced, and for $E > E^*$ any additional amount of energy is likely to be decreasing in value (to an already satiated bird) [250] (birds are risk averse). To capture this mapping of energy to marginal fitness we use

$$\Psi(E^i, k) = \left(\frac{(E^i)^\beta}{k^\beta + (E^i)^\beta} \right)^{\frac{1}{\alpha}}. \tag{3.1}$$

Here, E^i distinguishes hourly energetic intake for vigilant ($i = C$) and non-vigilant ($i = D$) birds. The parameters β and k respectively describe the slope of the function, and energetic cost per hour in terms of a birds metabolic rate. A parameter choice of $\beta = 8$ reflects a steep slope about the point of inflection where energy is balanced, E^*. We choose this functional form to reflect that while a birds energetic intake is less than metabolic costs, fitness increases at an increasing rate with feeding, but once energetic needs are surpassed, the relationship should plateau. Eventually a bird cannot intake any more energy whether it be limited by time constraints [200] or digestive physiology [269]. Of course other functional forms are possible. We adopt equation (3.1) as it is both analytically workable and biologically plausible; similar energy to fitness mappings have been used to model American gray jays *Perisoreus canadensis* [271].

An individual bird's energetic intake, E, depends on both it's behaviour and it's spatial location on the saltmarsh. We let E be the product of foraging intensity ϕ and quality of prey q. Using data from [265] which provides a site specific spatial map for the average densities of *Orchestia*, *Spaeroma* and *Hydrobia*, we find the average quality of prey items, $q(d)$ (in kJ) as a decreasing function of distance from (predatory) cover d. The relationship between q and d is found by calculating, for each distance from cover d, the abundance of each prey item along with it's average size and length. Using methods from Speakman [272], Goss-Custard [197] and Zwartds & Bloomert [273] the weight of each prey item measured in "Ash free dry mass" is transformed into calorific content. The average quality of an individual prey item follows the form

$$q(d) = 0.105 - \frac{(0.032(-40.038 + d))}{\sqrt{1 + (-40.038 + d)^2}}. \tag{3.2}$$

$E(q(d), \phi^i)$ is applied to represent energy intake, where ϕ^i is the number of swallows for vigilant ($i = C$) and non-vigilant ($i = D$) birds per minute. On average, we assume that vigilant birds make 4 swallows per minute yielding $E^C = 4 \times 60 \times q(d)$, and non-vigilant birds make 6 swallows per minute meaning $E^D = 6 \times 60 \times q(d)$ [123]. The parameter k of equation (3.1) is inferred using that energy assimilation in redshank is around $7.5 \times BMR$ (Basal Metabolic Rate) [198].

How metabolic costs affect fitness

The point at which a bird's energy intake equals its metabolic exertion is represented by the point of inflection of equation (3.1). Solving

$$\frac{\partial^2 \Psi}{\partial E^2} = \frac{8E^6 k^8 (7k^8 - 9E^8)}{(E^8 + k^8)^3} = 0 \tag{3.3}$$

we find $k = 3^{\frac{1}{4}} \times 7^{-\frac{1}{8}} \times E$ when a birds energetic needs are balanced. Therefore metabolic costs of a bird in $\mathrm{kJ\,h^{-1}}$ are expressed as a multiple of E. Using methods from Speakman [272] we apply

$$mc = 3.6(2.29 - 0.7333\ln(T) + 0.0123(M)) \tag{3.4}$$

to determine the energetic costs in $\mathrm{kJ\,h^{-1}}$ for foraging bird. Here, T represents the temperature (degrees) and M represents the bird's mass (in grams). Estimates of around 100 are appropriate in redshank [272].

Energetic needs are balanced when $E = mc$, i.e., when metabolic exertion equals energetic intake. Therefore substituting equation (3.4) into $E = mc$, and then into k we calculate metabolic exertion for given temperatures $\bar{T}$ and mass $\bar{M}$:

$$k = 3^{\frac{1}{4}} \times 7^{-\frac{1}{8}} \times E \tag{3.5a}$$

$$= 3^{\frac{1}{4}} \times 7^{-\frac{1}{8}} \times mc = 3^{\frac{1}{4}} \times 7^{-\frac{1}{8}} \times 3.6(2.29 - 0.7333\ln(\bar{T}) + 0.0123(\bar{M})) \tag{3.5b}$$

On average birds feed for 7.5 hours each day on the saltmarsh, however energy is also exerted at rest and whilst not foraging. To account for this we multiply their hourly metabolic costs by 24/7.5.

Parametrising the evasion function $g(x, n)$

We assume that the probability that the flock evades a predator attack depends on a disjoint probability union of attack failure resulting from predator confusion, and the effects of vigilance. As such, evasion takes on the additive functional form given by equation (2.1). Confusion effects are modulated by flock size and vigilance effects are modulated by the proportion of vigilant birds. We apply $g(x, n) = p_1 n/(k_1 + n) + p_2 x/(k_2 + x)$ with $0 < p_2, k_2 < 1, 0 < p_1, k_1$ to describe the evasion function where confusion effects are modulated by the n-dependent parameter p_1 and collective vigilance effects are controlled by p_2.

The probability of evading a predator also depends on how far from predator concealing cover, d, the flock feeds [25]. We use $p_1 = A_1(1 - \lambda_1 e^{\alpha_1 d})$ and $p_2 = A_2(n)(1 + \lambda_2 e^{\alpha_2 d})$ where $e^X = \exp(X)$ is the exponential function, such that

$$g(x, n, d) = \frac{A_1(1 - \lambda_1 e^{\alpha_1 d})n}{k_1 + n} + \frac{A_2(n)(1 - \lambda_2 e^{\alpha_2 d})x}{k_2 + x} \tag{3.6}$$

In redshank, overall collective (flock) vigilance is known to increase at a decreasing rate with additional individual vigilance [18]. Predator attack success asymptotes for flocks of sizes of 40 plus birds [25]. If all individuals are non-vigilant; $x = 0$, the first term in equation (3.6) describes the probability that the flock evades attack as a function of the confusion effects of flock size, and distance from cover only. In this case, the probability that the entire flock escapes capture plateaus for flocks of sizes larger than 20 birds and for distances greater than 30m from cover [151]. Otherwise if $0 < x < 1$ then the probability of evasion increases with the proportion of vigilant birds x, flock size n and distance from cover d.

TABLE 3.1: Evasion parameters

Parameter values							
A_1	$A_2(n)$	λ_1	λ_2	k_1	k_2	α_1	α_2
1.1	$1-\frac{1.1n}{1.6n+2}$	0.30	0.22	10	0.24	-0.05	-0.05

The parameter $A_2(n)$ captures how the effects of confusion and collective vigilance interact to produce overall probability of evasion. The relative magnitudes of effect that each collective detection and confusion have on overall flock evasion, depend on flock size. In larger flocks, confusion of the predator plays a dominant role in procuring safety for flock members through failed predator attacks [18, 25]. As such, in larger flocks we expect the n-dependent effects of confusion to dominate equation (3.6). In smaller flocks, the predator is unlikely to become confused and through the effects of collective vigilance, early detection offers the best probability of flock evasion [25]. We expect the relative magnitude of x-dependent effects to increase as flock size decreases. We use $A_2(n)$ to reflect how the magnitude of effects of vigilance and confusion change with flock size, and $(1-\lambda_2 e^{\alpha_2 d})x/(k_2+x)$ captures how overall collective detection increases with the proportion of vigilant birds x [18].

In line with previous research [25, 206, 207] equation (3.6) plateaus at 1 for both increases in flock size and distance from cover. To ensure that the probability of evasion is bounded by the closed unit interval, we define equation (3.6) to be piecewise continuous, such that $g(x,n,d) := 1 \ \forall \bar{d},\bar{n},\bar{x} : g(\bar{x},\bar{n},\bar{d}) > 1$ with parameter values given in Table 3.1. A full derivation of the evasion function is given in Appendix B.1

Parametrising ϵ

We are guided that birds with high rates of vigilance have only small advantages over birds with low rates of vigilance through early escape response [266]. As such we expect ϵ to be negative but of small magnitude. To find an appropriate value, we use the following findings: non-vigilant birds respond to a predator attack on average 0.07 seconds slower than their vigilant counterparts [255]. On average, birds feed between 10-20 meters to cover. Sparrowhawk fly at an average speed of $25\mathrm{m\,s^{-1}}$. With these factors in mind, an approximate benefit of being vigilant over non-vigilant is a 1.75 meter head start in escaping through an earlier response.

Redshank frequently fed less than 10m (meters) from the sparrowhawk in the study and for a reasonable chance of success, the sparrowhawk needs to be within a 15m proximity of the flock [255]. We propose that a upon an attack, a distance of 12m between flock and the sparrowhawk is a conservative estimate. As such, vigilant birds are $1.75/12 = 0.15$ times less likely to be targeted. We apply $\epsilon = -0.15$.

3.3 Predictions

The parametrisation of section 3.2.2 captures a trade-off between feeding and vigilance: the more a bird feeds the greater its energetic reserves, but the more susceptible to predation the bird is. In relation to the mathematical model, this trade-off reflects a situation in which $\delta, \epsilon < 0$ (see case 4 in Fig 2.1, section 2.2). The parametrised model combines the three key anti-predation effects of

collective vigilance, dilution and confusion. Vigilance effects dominate the x-dependent component p_2 of the evasion function (3.6): the probability that any flock member is attacked depends on the population frequency of vigilant birds, which translates to the expected number of vigilant birds in a flock. As such, by setting $p_2 := 0$ we can effectively remove the effects of collective vigilance from the model. The effects of an unequal dilution of risk between flock members are modulated by the ϵ parameter: by setting $\epsilon := 0$ we remove the fitness advantage of adopting a vigilant behaviour relative to adopting a non-vigilant behaviour through preferential predator targeting. Confusion effects are controlled by the n-dependent parameter p_1 of equation (3.6): confusion effects increase with flock size. By setting $p_1 := 0$ we remove confusion effects from the model.

In this section, we produce theoretical predictions of different model configurations. Each model configuration is defined by the removal of any one of these three anti-predation effects. We give predictions for the "full model" i.e., when all three effects are included, and reduced models, for which an effect is independently removed (as explained above). In doing so we aim to discover which combination of anti-predation effects considered by the model are required to qualitatively predict the anti-predator behaviours in redshank (predictions are to be qualitatively compared to the observational data (see section 3.4).

The subsequent analysis utilises the following mathematical equivalence, which is applied to the model results of section 2.3.1 and 2.3.3: a polymorphic population with respective frequencies q and $1-q$ adopting (pure) vigilant and non-vigilant behaviours is equivalent to a monomorphic population in which all individuals adopt the vigilant behaviour with probability q and non-vigilant behaviour with probability $1-q$ (explained in the introduction section 1.1.5 or [35, 43, 60]). Hence, the population states predicted by the model, i.e., frequencies $x \in [0,1]$ of vigilant birds are equivalent to the probability that any randomly chosen bird is vigilant. We leverage this result to predict the expected number of vigilant birds, nx, within in a flock of size n.

3.3.1 Prediction 1: vigilance decreases with flock size

The full model predicts that as flock size increases, the frequency of vigilant birds in the population will decrease. This prediction is equivalent to a decrease in the probability that any focal bird within a flock is vigilant with increases in the flock size, as illustrated in Fig 3.1 and Fig 3.2. In each figure, panels (i)-(iv) respectively show model predictions for the full model (i), and models with removed unequal dilution of risk (ii), vigilance (iii) and confusion (iv).

When the effects of dilution are removed from the model, the model predicts little (when $\mu > 0$) to no (when $\mu = 0$) variation in the probability that a bird is vigilant with changes in flock size. When $\mu = 0$, for all flock sizes the model predicts birds to be non-vigilant with probability one, and when $\mu << 1$, the model predicts birds to have a very high probability of choosing the non-vigilant behaviour. We explain this as follows: by removing the effects of (unequal) dilution, we model the risk dilution of all birds as the pure inverse of flock size. By choosing to be vigilant, birds do not receive any advantageous safety benefits relative to choosing to be non-vigilant. In this case, the only differences in fitness between the vigilance behaviours occur through differing foraging benefits, which favour those birds choosing the non-vigilant behaviour.

When confusion effects are removed from the model (Fig 3.1 and Fig 3.2 panels (iv)), the model predicts that a bird will be vigilant with probability one (for $\mu = 0$), or nearly one (for

$\mu > 0$), regardless of flock size. By removing confusion effects from the model, we remove flock size effects in our modelling of the predator evasion function (equation 3.6). This has two effects. First, for all flock sizes, there is a decrease in the probability of predator evasion. Given that all birds experience an increase in the probability of attack, any focal bird experiences an increased magnitude of dilution effects on it's fitness. Since choosing the vigilant behaviour yields birds a relative dilution advantage, the model predicts a greater probability for birds to be vigilant. Second, without group size effects, the probability of evasion only changes with changes in the number of vigilant flock members. This results in a relative increase in the magnitude of the (positive) effect that vigilance has on predator evasion, and thereby, on fitness.

When vigilance effects are removed, the model predictions are subtly different when we account for behavioural exploration compared to pure imitation. When $\mu > 0$ the predicted probability that a bird is vigilant decreases with flock size, but at a more gradual rate than the full model (see Fig 3.2 panel (iii)). By removing the effects of vigilance, with all other parameters equal, the model predicts a lower probability that the flock evades attack. This has the effect to increase the relative magnitude of dilution effects on individual fitness. Since vigilant bird's have a relative advantage through the dilution effect, the increase in the magnitude of dilution effects acts to increase the fitness of vigilant birds relative to non-vigilant birds. Hence, the model predicts a shifting out of the probability that any bird is vigilant, for given flock size class (comparing panels (i) to (iii)). Contrastingly, when we assume zero behavioural exploration ($\mu = 0$), there is a region of bi-stability in the population state. This corresponds to birds being vigilant with probability zero or one, depending on their initial behaviour. The range of flock sizes where the model predicts that either of two configurations describe the probability that a bird is vigilant (see Fig 3.1 panel (iii)) reflects a clear trade-off for individual birds: flock size is not sufficiently large for safety through numbers to select for a bird to be non-vigilant with probability one, yet flock size is not sufficiently small such that a bird prioritises vigilance. For sufficiently small flocks this model configuration predict birds to be vigilant, and in sufficiently large flocks, non-vigilant.

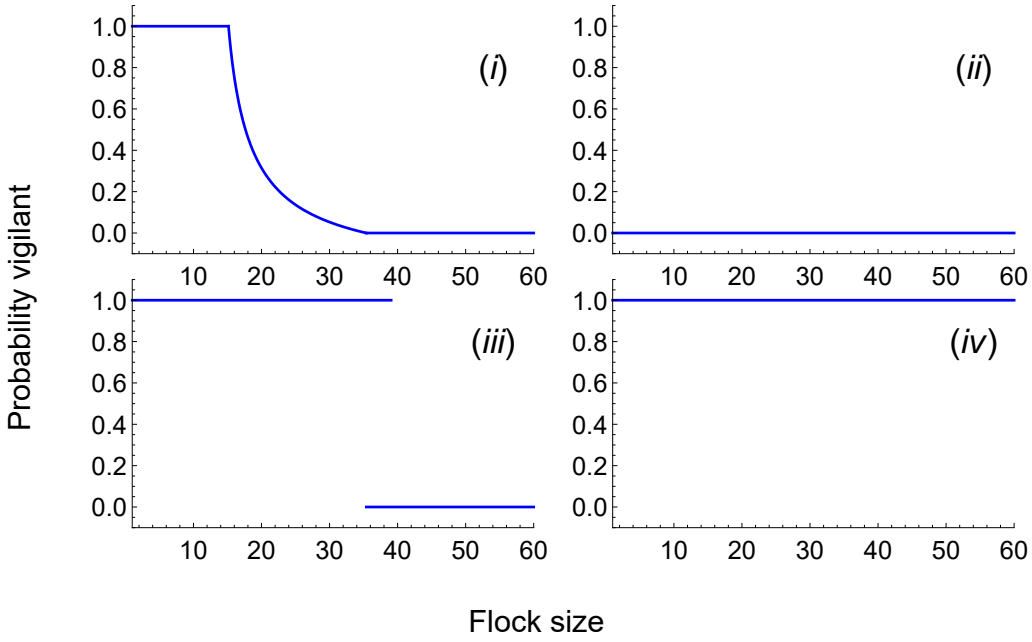


FIGURE 3.1: Model predictions for the probability a bird is vigilant given flock size, when $\mu = 0$. Each curve gives the probability that a focal bird is vigilant given flock size, and each panel (i)-(iv) shows model predictions when including different anti-predation effects. (i): the full model. (ii): the model predictions with dilution effects removed. (iii): predictions when vigilance effects are removed, and (iv): the model predictions when confusion effects are removed.

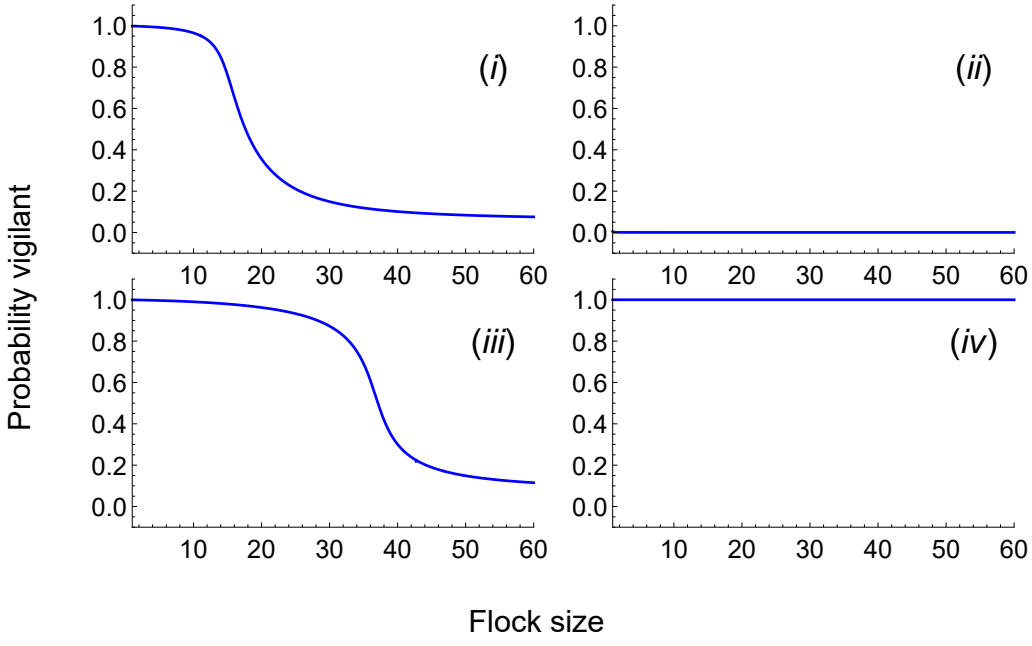


FIGURE 3.2: Model predictions for the probability a bird is vigilant given flock size, when $\mu = 0.001$. Each panel shows the model predictions when different anti-predation effects are included. (i): the full model. (ii): the model predictions with dilution effects removed. (iii): predictions when vigilance effects are removed, and (iv): the model predictions when confusion effects are removed.

To capture the model prediction of a mixture of vigilant and non-vigilant behaviours, we apply the mean and variation of inter-scan intervals within flocks of given sizes. To do this, we use the

marginal density function with variable distance to cover (see section 3.3.4). The intuition is that as the frequency of vigilant birds decreases with flock size, the average inter-scan interval should increase, and in cases where the model predicts a mixture of vigilant and non-vigilant birds, one would expect a higher variation in the inter-scan intervals within a flock.

Let ζ^v represent the inter-scan interval of a vigilant bird, and ζ^{nv} as the inter-scan interval for a non-vigilant bird. Inferring parameters of $\zeta^v = 1$ and $\zeta^{nv} = 10$ as suitable values from [123], we apply

$$\tilde{x}(d) := \hat{x}(d) \cdot \zeta^v + (1 - \hat{x}(d)) \cdot \zeta^{nv}. \tag{3.7}$$

Using equations 3.14, the average inter-scan interval within a flock is

$$\text{mean}[\zeta] = E[\tilde{x}(d)] = E[\hat{x}(d)] \cdot \zeta^v + (1 - E[\hat{x}(d)]) \cdot \zeta^{nv} \tag{3.8}$$

and using equations 3.15, the variation between inter-scan intervals within a flock is

$$\text{Var}[\zeta] = \text{Var}[\tilde{x}(d)] = (\zeta^v - \zeta^{nv})^2 \cdot \text{Var}[\hat{x}(d)] \tag{3.9}$$

where we use $R_- = 1$ and $R_+ = 60$ as a general average, since 85 percent of attacks in our data analysis (see next section) occurred within 60m from cover (dataset from [202]).

Fig 3.3 shows the theoretical model predictions for mean inter-scan interval in the flock, and the variation between individual birds' inter-scan intervals, given flock size. We average the inter-scan interval means and variances over flock size class (which are specified by the data the inter-scan intervals of individual birds). The full model predicts qualitatively similar results for the cases of $\mu = 0$ and $\mu > 0$: that mean inter-scan interval (blue bars) monotonically increases with flock size, and that variation between flock members' inter-scan intervals (black error bars) peaks for mid-sized flocks as shown in Fig 3.3a and Fig 3.3b. This prediction of the mathematical model can be explained as follows: the frequency of vigilant birds is predicted to decrease with flock size. For each increase in flock size, respectively fewer vigilant birds, and more non-vigilant birds are predicted in the flock. An increasing mean inter-scan interval is in this sense indicative of a decreasing proportion of vigilant birds with increasing flock size: flock members prioritise feeding and spend longer times between vigilant scans. A high variance of inter-scan intervals is indicative of multiple vigilant and non-vigilant birds within the flock: some birds prioritise vigilance and some prioritise feeding. The full model therefore predicts that small flocks encourage birds to be vigilant (low mean and low variance of inter-scan interval), mid-sized flocks promote a mixture of vigilant and non-vigilant birds (mid sized mean, large variance of inter-scan interval) and large flocks encourage birds to be non-vigilant (high mean, no variance of inter-scan interval).

The model's qualitative predictions in cases when the effects of vigilance, dilution and confusion are removed are shown in Fig 3.3c and Fig 3.3d. When dilution effects are removed, in the case of $\mu = 0$ the model predicts no change in the mean inter-scan interval with flock size (green bars in Fig 3.3c), and when $\mu > 0$, the model predicts that mean inter-scan interval is virtually invariant with flock size changes (green bars in Fig 3.3d). If vigilant birds do not receive a favourable dilution experience, then choosing to be non-vigilant always yields a greater fitness. The model predicts that all (when $\mu = 0$) or nearly all ($\mu > 0$) birds choose the non-vigilant behaviour. The effects

of unequal dilution are necessary for the model to predict a mixture of vigilant and non-vigilant behaviours, and thus the model predicts no variation in inter-scan intervals under pure behavioural imitation, and practically zero variation when $\mu > 0$ (errors over the green bars in Fig 3.3c,Fig 3.3d).

When the effects of confusion are removed we find that in the model with behavioural exploration ($\mu > 0$), there is a very gentle U-shaped relationship between mean inter-scan interval and flock size (red bars of Fig 3.3d). Whereas, without behavioural exploration, the model predicts that for all flocks of size greater than five, the mean inter-scan interval is invariant with flock size (red bars Fig 3.3c). Without group size effects, for each flock size, individual birds experience a greater probability of attack and an increase in mortality risk. The increased risk encourages birds to choose the vigilant behaviour in two ways. First, an increase in the relative importance of dilution effects favours those individuals adopting the vigilant behaviour, as explained previously through the unequal dilution of risk. At the same time, when group size effects are removed the probability of predator evasion decreases, and only depends on the number of vigilant birds. The relative influence that vigilance effects have on survival increase, and as such, birds have an increased propensity to be vigilant. In the respective cases of $\mu > 0$ and $\mu = 0$, birds are predicted to be vigilant with very high probabilities or probability one, and there is zero to very little variation in inter-scan interval.

If the effects of vigilance are removed (blue bars in Fig 3.3c and Fig 3.3d), the effect of random behavioural exploration produces qualitatively differing model predictions. When $\mu > 0$, the model predicts mean inter-scan interval to monotonically increase Fig 3.3c, but at a lower rate than the full model. We find that compared to the full model, increases in inter-scan interval are shifted out to larger flock sizes. The gradual increase in mean inter-scan interval is accompanied by a peak in variation of individual inter-scan intervals within the flock, for flocks of between $36-40$ birds. This prediction is because for this flock size class, the model expect the greatest mixture of vigilant and non-vigilant birds. Without exploration ($\mu = 0$), the model predicts a constant inter-scan interval representative of entire flocks adopting the vigilant behaviour for all flock size classes, except for flock size classes of $55-60$ at which the mean inter-scan interval is characterised by flocks being entirely composed of non-vigilant birds.

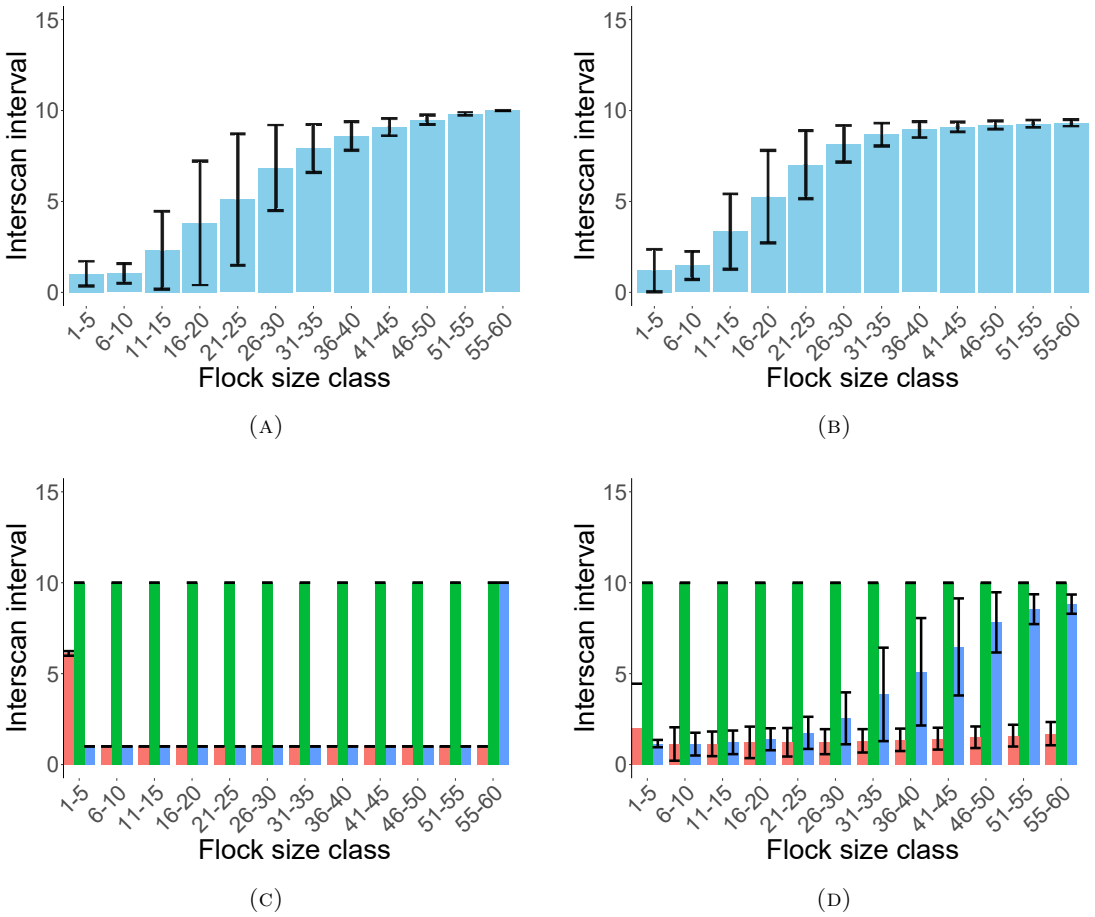


FIGURE 3.3: The model prediction for the mean and the variation of individual bird's inter-scan intervals. (A),(B): The full model. Blue bars show the mean predicted inter-scan interval and black error bars centred on the mean inter-scan interval show mean plus or minus the predicted standard deviation of inter-scan intervals within the flock. (A): Without exploration and (B): with exploration rate of $\mu = 0.001$. (C),(D): Model predictions when anti-predation effects are removed. Blue, green and red bars respectively show the mean inter-scan interval when the effects of vigilance, dilution and confusion are removed. Black error bars show standard deviation. (C): Without exploration and (D): with $\mu = 0.001$

3.3.2 Prediction 2: vigilance decreases with temperature

We expect that in environments which exacerbate state dependent metabolic demands experienced by the birds, the proportion of vigilant birds should decrease. In particular, the model is able to predict how for any given flock size, the proportion of vigilant birds should change according to differing temperatures. As the proportion of vigilant birds feeds back into the probability that the flock evades the predator (equation 3.6) the model predicts differing probabilities of capture as a function of the climate. For this prediction we apply the full model which includes all three key anti-predation effects.

The effect of temperature on a bird's state is modelled by applying the γ_i functions with different parameter values of k, derived using equation (3.5) from section 3.2.2. Changes in the parameter k reflects differing metabolic expenditures. We use k_{warm} to model metabolic demands during high temperatures, and k_{cold} reflect the effects of low temperatures. The parameters k_{warm} and k_{cold} are

found by taking average temperatures in the data analysis, for which attack recordings are grouped into temperature classes (see section 3.4.3). Model predictions are based on the average flock size in each of the flock size classes that the data is grouped into, and an average distance from cover in each of the flock size classes.

The model predicts that the proportion of vigilant birds in a flock should increase with increasing temperatures; Fig 3.4. At lower temperatures and increased metabolic demands, for fixed flock sizes we expect the proportion of non-vigilant individuals to increase; Fig 3.4a and Fig 3.4c respectively for the models without and with exploration. This result is robust for a range of flock sizes. For decreases in metabolic demands the model predicts increases in the proportion of vigilant birds for fixed flock size; Fig 3.4b and Fig 3.4d.

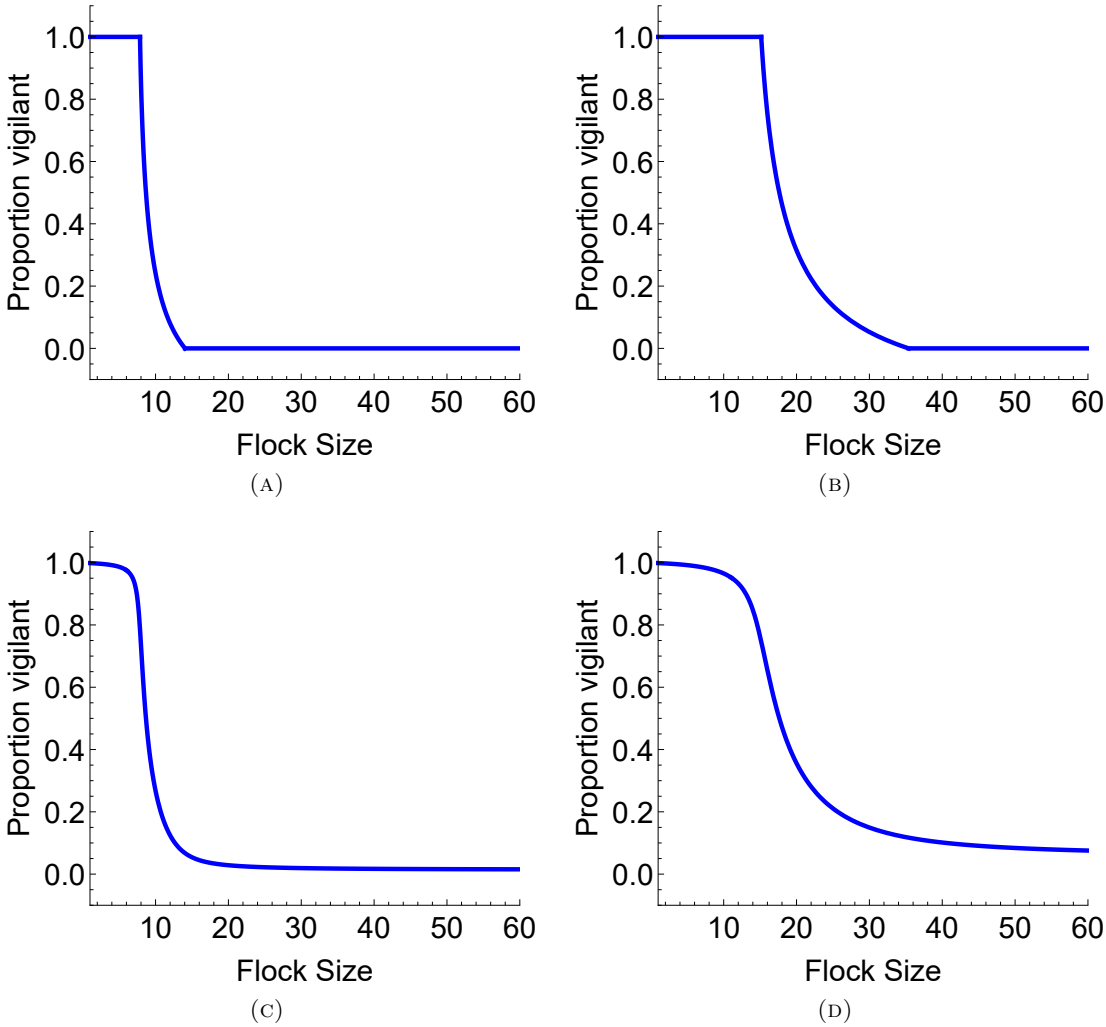


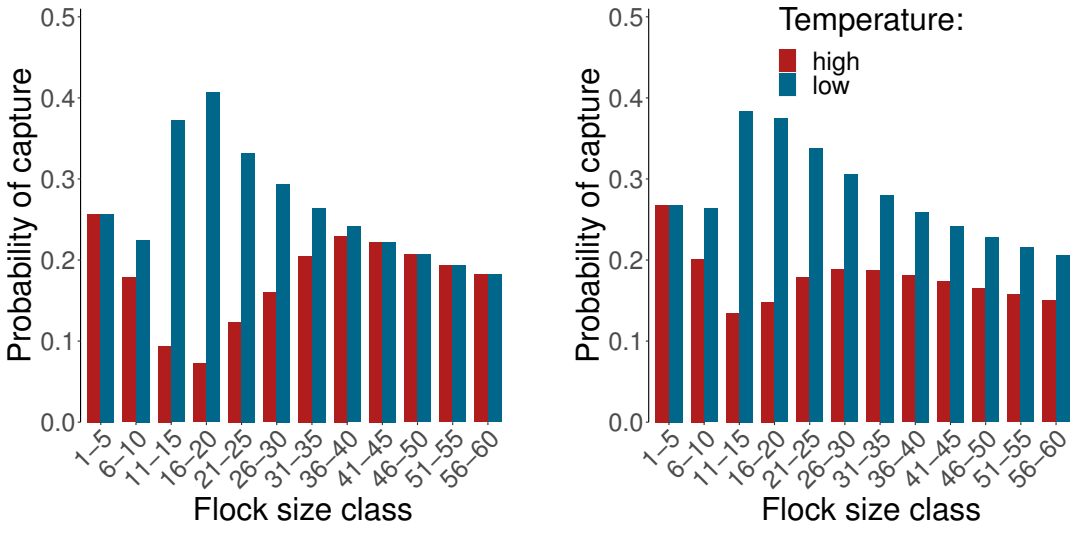
FIGURE 3.4: Model predictions for the proportion of vigilant birds as dependent on metabolic demands. (A),(B): without behavioural exploration and (C),(D): with behavioural exploration. (A),(C): The predicted proportion of vigilant birds given flock size at low average monthly temperatures of 4.2 degrees; birds face high metabolic costs, and (B),(D): the predicted proportion of vigilant birds during high average monthly temperatures of 6.2 degrees; birds face low metabolic costs. The shifting out from (A),(C) to (B),(D) illustrates a decrease in foraging behaviour with increasing temperatures.

By informing the evasion function (3.6) with the predicted frequency (proportion) of vigilant birds x, the model predicts a probability that a flock member is captured Fig 3.5. When $\mu = 0$, the model predicts that for a range of flocks sizes, the probability that at least one flock member is captured by a predator is higher given low temperatures than given high temperatures; Fig 3.5a. When behavioural exploration ($\mu > 0$) is included in the model, the qualitative prediction is similar, except in this case, flocks are never comprised of all vigilant or non-vigilant birds and as such capture probabilities do not reach the same extrema; Fig 3.5b.

In both cases of $\mu = 0$ and $\mu > 0$, at around flock sizes of 16-20 birds and given high temperatures, the model predicts that the probability of capture is at a minimum. Whereas given low temperatures, for the same flock sizes the model predicts a maximum probability of capture. We explain this result using the critical group sizes n_1 and n_2 (see section 2.3.1). We see that $n_1 \approx 16$ and $n_2 \approx 9$ in Fig 3.4a and Fig 3.4c) where flocks of size less than n_2 are entirely vigilant and flocks of size greater than n_1 size are entirely non-vigilant.

When the temperature is low there is a relatively steeper decrease in the proportion of vigilant birds between flock sizes n_2 and n_1 Fig 3.4a and Fig 3.4c. Since the model is parametrised such that vigilance effects dominate the predator evasion function for small flock sizes, there is a corresponding significant increase in the probability of capture within these flocks (blue bars increasing in Fig 3.5a and Fig 3.5b). For flocks of size greater than n_1 the model predicts no vigilant birds in the flock and the evasion function (equation 3.6) is independent of x. The only factor affecting the probability of capture in these flocks is flock size, which is parametrised to have a negative effect. As such, for flocks of size greater than n_1, the probability of capture is decreasing with increasing flock size. Overall, given low temperatures the resulting effect is for the probability of capture to increase and then decrease.

For high temperatures the model predicts a more gradual decrease in the proportion of vigilant birds with flock size (see Fig 3.4b and Fig 3.4d). In this case $n_1 \approx 37$ and $n_2 \approx 18$. While flock size is less than n_2, flocks are entirely vigilant, and the only factor affecting the probability of capture is flock size: as flock size increases the probability of capture decreases (initially red bars are decreasing in Fig 3.5a and Fig 3.5b). For flock sizes greater than n_2 and less than n_1, there is a relatively greater proportion of vigilant birds (compared to lower temperatures) within the flock (compare Fig 3.4b to Fig 3.4a). For this range of flock sizes, a relatively greater proportion of vigilant birds within the flock combined with the effects of flock size have the overall effect to minimise the probability of capture (see flock size class "16-20" in Fig 3.5a and "11-15" in Fig 3.5b). For increases in flock size between flocks of around 20 birds to n_1, the model predicts the number of vigilant flock members to be decreasing with flock size (as shown in both Fig 3.4b and Fig 3.4d), which has a positive effect on the probability of capture. For the same range of flock sizes, the models predicts the reduction in the probability of capture resulting from confusion effects – the flock size dependent effects – to become saturating. The overall effect over this range of flock sizes, is for the probability of capture to be gradually increasing with flock size. For flocks of sizes greater than n_1 there are no vigilant birds in the flock. The only factor affecting the probability of capture within these flocks is flock size, which has a negative effect, and the probability of capture is decreasing with flock size. Hence, for high temperatures the model predicts the probability of capture to have a cubic type structure: decreasing, then increasing, and finally decreasing.



(A) (B)

FIGURE 3.5: Model prediction for the probability that a bird is captured depending on temperature. Bars show the theoretical model's prediction for the probability that a flock member is captured. (A): Model without exploration. For a range of flock sizes (approximately between 5 and 41 birds), the model predicts a higher probability of capture for lower temperatures. (B): Including behavioural exploration. For all flock sizes, the model predicts that the probability of capture is higher for low temperatures.

The contrasting minima and maxima of the probability of capture are more pronounced for the model without behavioural exploration. In this model, for flock size roughly between 9 to 16 birds during cold months (Fig 3.4a), and 18 to 37 birds during warm months (Fig 3.4b), we find a predicted a mixture of vigilant and non-vigilant birds. The expected proportion of vigilant birds transitions from one to zero with increasing flock size, and as such the evasion function is fitted with respective maxima and minima of group vigilance effects. When there the frequency of vigilant birds is maximum, (equal to one), the probability of capture is minimised with respect to collective detection. This is predicted during the warm temperatures as shown by the red bars in Fig 3.5a. When the frequency of vigilant birds is minimum, i.e., zero, the probability of capture is maximised with respect to collective detection, as shown by the blue bars in Fig 3.5a reflecting model predictions for low temperatures.

In contrast, the model with behavioural exploration never predicts a flock of entirely vigilant or non-vigilant birds due to the birds' recurrent random sampling of behaviours (Fig 3.4c and Fig 3.4d). As such, as flock size increases the predicted frequency of vigilant birds transitions from a state in which nearly all are vigilant to one in which nearly all are non-vigilant. Importantly, because $\mu > 0$, flocks are never predicted entirely vigilant or non-vigilant. This prediction is reflected in the probability of capture never being minimum for the high temperatures (red bars Fig 3.5b), or maximum for the low temperatures (blue bars Fig 3.5b), with respect to collective detection, since the expected number of vigilant or non-vigilant birds within the flock is never equal to flock size.

3.3.3 Prediction 3: vigilance decreases with increases in distance to cover

The full model predicts that the probability that an individual bird is vigilant, and equivalently the overall number of vigilant birds in a flock, should decrease with increases in distance from cover for fixed flock size Fig 3.6. As both flocks sizes increases and distance from cover increases, we expect the probability that a focal bird is vigilant to decrease.

In cases without behavioural exploration ($\mu = 0$) the model predicts the following: for sufficiently large flocks, and even at distances very near cover, birds will be non-vigilant with probability one, as shown by the upper left white region (where $x = 0$) of panel (i) in Fig 3.6. The white region defined here shows distance to cover flock size combinations for which the model predicts that birds are non-vigilant with probability one; flocks are composed of entirely non-vigilant birds. In between the upper and lower white regions, the model predicts an expected flock composition containing both vigilant and non-vigilant birds. The relative proportions of vigilant and non-vigilant birds characterising the flock composition are respectively predicted to decrease and increase, as either distance from cover of flock size increases.

When an exploration rate is included the model ($\mu > 0$) predicts birds to non-vigilant with a very high probability but never probability one, as illustrated by the colour legend in panel (ii) of Fig 3.6. In this case, flocks are predicted to always feature a mixed composition of vigilant and non-vigilant birds, but the same qualitative finding that the vigilant proportion decreases with flock size and distance holds.

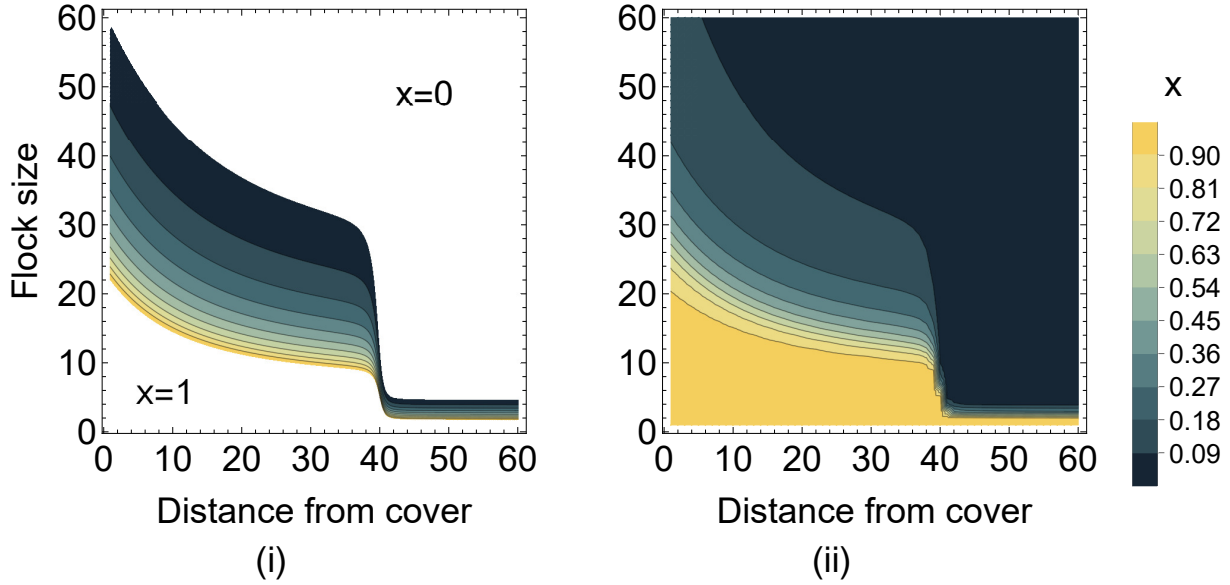


FIGURE 3.6: The model predictions for the probability, x, that an individual redshank is vigilant, given flock size and distance from cover. The same legend scale gives the probability that an individual bird is vigilant given (d, n) combinations in each figure. (i): Without exploration. The top right white region shows distance to cover and flock size combinations for which the model predicts that a bird is vigilant with probability zero, and the bottom left white region shows a probability of one. (ii): With exploration $\mu = 0.001$ the absorbing probabilities of $x = 0, 1$ are no longer valid.

A horizontal cross section in Fig 3.6 shows the full model prediction of how individual vigilance changes with distance from cover, for a flock of fixed size. Considering an average sized flock of

size 30, Fig 3.7 show the predictions of different model configurations, respectively for $\mu = 0$ and $\mu > 0$.

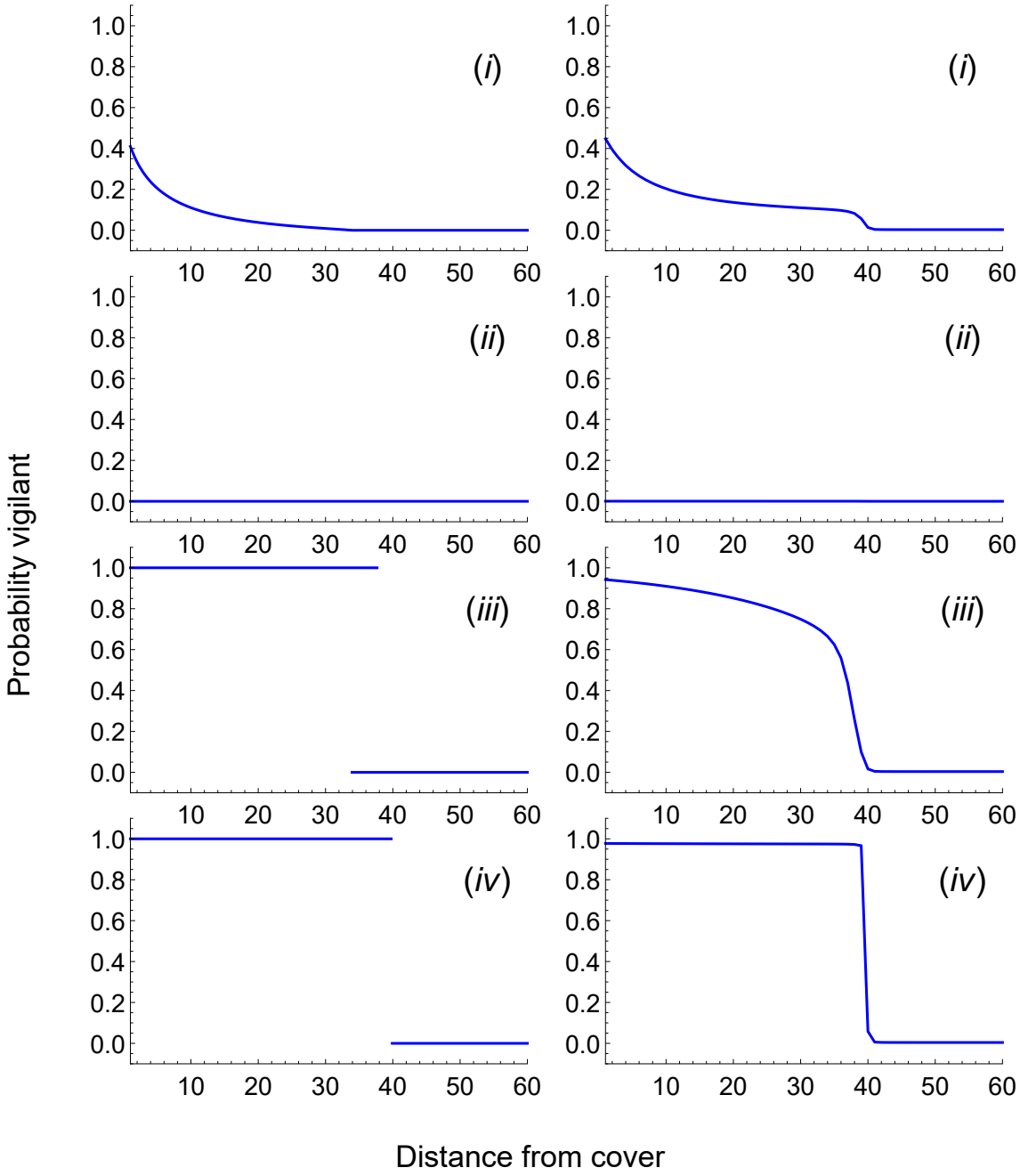


FIGURE 3.7: Theoretical predictions of the probability that an individual bird is vigilant changes with distance from cover. Each panel shows the theoretical predictions regarding how the probability that an individual bird is vigilant should change with distance from cover. Panel (i) shows the full model; (ii) shows the model without unequal dilution of risk; (iii) shows the model without vigilance, and (iv) shows the model without confusion effects. (Left) for $\mu = 0$; (right) for $\mu > 0$.

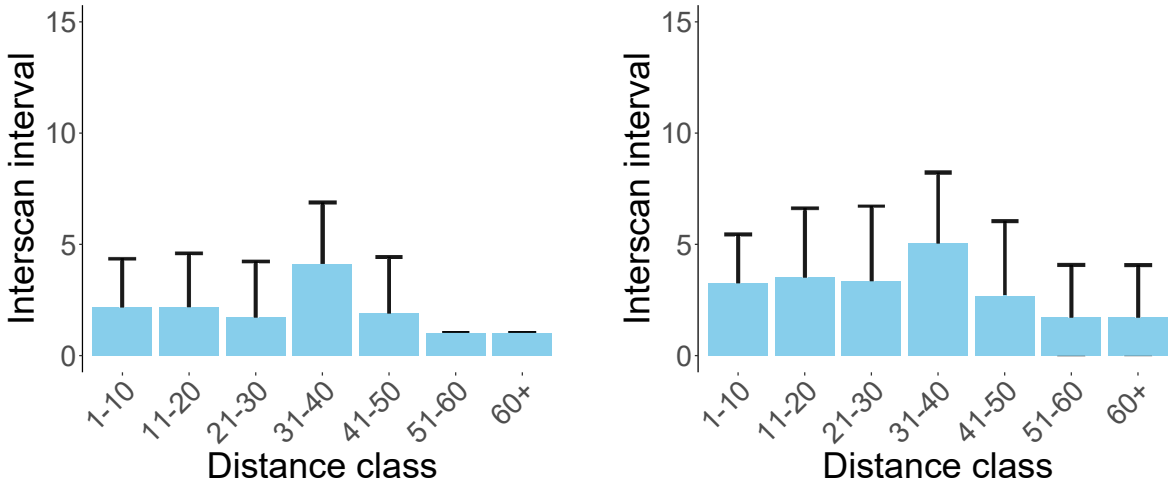
Fig 3.7 provides a good description of how we would expect vigilance to change with distance from cover, assuming that there are no concomitant changes in flock size. However, over the range of distances from cover, flock size may vary. To capture the possibility of differing flock sizes occurring at different distance to cover, we introduce the joint probability density function

to find the mean inter-scan intervals over specific distance classes. We divide the distance range $R = R_1, R_2, R_3, ..., R_6$ with R_i increments of 10 meters. We then use a series of probability densities, $\psi_i(\hat{x}(n,d))$, each integrated over the whole range of flock sizes $n \in Q = [1, 60]$, and over each $d \in R_i$:

$$\psi_i(\hat{x}(n,d)) := \frac{1}{\mathcal{N}_i}\hat{x}(n,d) \text{ where } \mathcal{N}_i = \int_{R_i}\int_{Q} \hat{x}(n,d)\mathrm{d}n\mathrm{d}d \tag{3.10}$$

which provides the mean and variation of inter-scan interval in each distance interval R_i, and over all flock sizes.

The predictions of the full model are shown in Fig 3.8 where blue bars show mean inter-scan interval and error bars give one way standard deviation (i.e., mean + sd). The model predicts that the average inter-scan interval peaks for flock sizes of around 31-40 birds. For flock sizes larger than this threshold there is a significant decrease in the model predicted mean inter-scan interval. Comparison with Fig 3.6 helps explain this finding. At around 40 meters from cover, the model predicts a significant decrease in the probability that a focal bird will be vigilant: individuals are predicted to be non-vigilant with probability near to one for all distances greater. For any given flock size, the model would predict a significant drop in the expected proportion of vigilant birds, as distance to cover increases past the threshold. However, over all possible flock sizes, the model predicts that the mean proportion of vigilant birds should gradually increase, corresponding to an decrease in the mean inter-scan interval.



(A) (B)

FIGURE 3.8: Mean and variation in the inter-scan intervals within flocks for increases in distance to cover. Blue bars show mean inter-scan interval, and black errors show mean inter-scan interval + standard deviation of inter-scan intervals. (A): The full model without behavioural exploration. (B): The full model with an exploration rate of $\mu = 0.001$.

3.3.4 General model analysis

Some of the predictions just made utilise an equivalence between the (population) state $\hat{x}(n,d)$ of the theoretical model and the probability that a bird is vigilant, given flock size n and distance to

cover d. We introduce the joint probability density function

$$\psi(\hat{x}(n,d)) = \frac{1}{\mathcal{N}}\hat{x}(n,d) \tag{3.11}$$

where

$$\mathcal{N} = \int_Q \int_R \hat{x}(n,d)\mathrm{d}d\mathrm{d}n \tag{3.12}$$

is the appropriate normalisation constant using the domain $\mathcal{D} = (n,d) \in Q \times R$. We find the marginal density functions $\psi(\hat{x}(n))$ and $\psi(\hat{x}(d))$ by respective integration over all $d \in R$ and $n \in Q$:

$$\psi(\hat{x}(n)) = \int_R \psi(\hat{x}(n;d))\mathrm{d}d \text{ and } \psi(\hat{x}(d)) = \int_Q \psi(\hat{x}(n;d))\mathrm{d}n. \tag{3.13}$$

Here, the distance boundaries R_- and R_+ are inferred from previous studies [207]. The flock size boundaries are given by Q_- and Q_+ can be inferred from [25].

Using such methods, the expected value for the mean proportion of vigilant birds over either d or n can be found by either of the marginal densities

$$E[\hat{x}(d)] = \int_{d\in R} \hat{x}(d)\psi(\hat{x}(d))\mathrm{d}d \text{ or} \tag{3.14a}$$

$$E[\hat{x}(n)] = \int_{n\in Q} \hat{x}(n)\psi(\hat{x}(n))\mathrm{d}n \tag{3.14b}$$

and the variance of the proportion of vigilant birds given by

$$\mathrm{Var}[\hat{x}(d)] = \int_{d\in R} \psi(\hat{x}(d))\big(\hat{x}(d) - E(\hat{x}(d))\big)^2\mathrm{d}d \text{ or}$$

$$\mathrm{Var}[\hat{x}(n)] = \int_{n\in Q} \psi(\hat{x}(n))\big(\hat{x}(n) - E(\hat{x}(n))\big)^2\mathrm{d}n \tag{3.15a}$$

3.4 Data

We analyse trends in observational data on redshank, and qualitatively compare to the predictions of the theoretical model. All data analysed for the purpose of comparison of model predictions is separate from other observational studies used to inform the model. The data was supplied by W. Cresswell and comprises information on redshank foraging over several winters. Redshank vulnerability to predation from sparrowhawks with respect to group size and position, as well as to predator concealing cover was assessed by observing attacks and kills. Time spent foraging or scanning for predators in redshank was also recorded. The study area consisted of saltmarsh habitat adjacent to mudflats and woodlands at Tynninghame Estuary, East Lothian, Scotland, U.K (see [267] for a more details).

Predation data were collected using the methods in [274]. Data were collected between September to early March 1989-1992 and 2001-2006. Sparrowhawk and peregrine falcon, *Falco peregrinus*, attacks were recorded during observation periods. There were sparrowhawk attacks recorded on 288 separate days over 8 winters with mean $\pm$ standard error $= 2.5 \pm 0.3$ attacks per day. Peregrine attacks were recorded on 102 separate days over 8 winters with mean $\pm$ standard error $= 1.7 \pm 0.3$

attacks per day. Over the 16 year period of the study, it is likely that numerous different predators were involved in attacks.

Attacks were defined as fast flights of the raptor towards a single redshank within a flock. Attacks were categorised as surprise versus pursuit. In a surprise attack, the raptor was obscured from the view of the redshank that it attacked, using visual cover to conceal some part of the attack. Usually, this occurred via the raptor using the adjacent woodland habitat. Pursuit (or "open" in [151]) attacks occurred when the raptor initiated attack in full view of the redshank it targeted. Usually such attacks happened when the redshank had already begun an escape response.

A kill was defined a when a raptor captured a redshank. For each kill, the flock size and distance from predator-concealing cover was estimated. A flock was defined as a congregation of 1-100 birds, for which the maximum distance from one bird to its nearest neighbour was less than 25 meters. In nearly all cases flocks were composed of the single species redshank, although on occasion flocks of size 30 or more birds contained a single bird of another species. The presence of other species was ignored in data collection on flock size, and the attacks were not recorded if the other species was targeted by the raptor. Distance to cover was estimated using regularly spaced markers around the edge of the saltmarsh. In total we used 642 attacks for which there were 107 captures. Capture rates are calculated over flock size classes as the total number of captures divided by the total number of attacks.

Foraging data were collected between December 2004 and February 2005, and from November 2005 to February 2006 at the same location. Observations of foraging redshank were conducted on the edge of the saltmarsh. Prey consisted of *Orchestia* and birds faced competition for food due to temporal depressions of prey availability. Video analysis was used to infer birds' distances from predators, position within the flock, number of prey consumed and proportion of time scanning for predators. See Sansom *et al.* [202] for full details on data collection. In total 159 observations on 21 different individual birds was conducted in the first winter period (2004-2005), and 160 observations on 25 individuals during the second winter (2005-2006).

Vigilance behaviour was defined using head-up versus head-down position. During the analysis, a bird was scored as vigilant if its head was above the horizontal body line and non-vigilant if its head was below the horizontal body line. Searching an probing for the *Orchestia*, a cryptic prey, was considered mutually exclusive with vigilance behaviour. We use the proportion of time spend a bird spent in a "head-up" position defined by the proportion of heads-up per second per individual bird, as one measure of vigilance. The inter-scan interval (inter-scan interval) of a bird, defined as the time spent between scanning bouts per individual bird, is used as another measure of vigilance.

3.4.1 Data analysis methods

Subsequent analysis applies the techniques of Bayesian regression. Details of the statistical analysis used in each section of the results are detailed in the following subsections. For all analysis, we report the posterior mean coefficients and the 95% credible interval (CI) around these coefficients. Credible intervals not overlapping 0 demonstrate 95% confidence that the mean of the coefficient is different from 0. All regression analysis is carried out using the brms package of R software [275]. The methods apply a Markov Chain Monte Carlo (MCMC) algorithm implemented in stan [276] via the brms package [277]. We checked that all models reached convergence of coefficients, absence

of divergent transitions, and good mixture of chains and effective sample sizes in the "bulk" and the tails of the distribution.

3.4.2 Predictions 1 & 3: data analysis

The theoretical model predictions 1 & 3 as presented in respective sections 3.3.1 and 3.3.3 are both assessed using the existing foraging dataset introduced in section 3.4. We conduct all relevant new data analysis relating to these predictions in this section.

Predictions 1 & 3: investigating how the proportion of time individuals spend vigilant changes with flock size and distance from cover

New analysis is conducted to assess the effects of flock size and distance from cover on the proportion of time individual birds spend head-up. Because we are dealing with a proportion which we cannot express as a sample of binary outcomes, we adopt a Beta regression [278]. That is, we assume that the proportion of time a bird spends head-up follows a beta distribution. For further details of the Beta regression see section B.2 of appendix B or [278].

Our theoretical model predicted that individual redshank vigilance depended on flock size and distance from cover (see section 3.3.1). As such, we model the proportion of time a bird spends vigilant (p.vig) to depend on flock size (Flock size) and distance from cover (Distance from cover) using a Bayesian Beta multilevel regression model, defined by

$$\begin{aligned}
\text{p.vig}_i &\sim \mathcal{B}(u_i, \phi_i) && \text{(3.16a)}\\
\text{logit}(u_i) &= \beta_{0,\text{bird ID}_i} + \beta_1 \text{Flock size}_i + \beta_2 \text{Distance from cover}_i && \text{(3.16b)}\\
\beta_{0,\text{bird ID}} &\sim \mathcal{N}(\beta_0, \sigma) && \text{(3.16c)}\\
\beta_0 &\sim \mathcal{N}(0, 0.25) && \text{(3.16d)}\\
\sigma &\sim \text{Exponential}(1) && \text{(3.16e)}\\
\beta_1 &\sim \mathcal{N}(-0.01, 0.01) && \text{(3.16f)}\\
\beta_2 &\sim \mathcal{N}(-0.01, 0.01) && \text{(3.16g)}
\end{aligned}$$

where $\beta_{0,\text{bird ID}}$ is a random intercept (group level effect) for each recorded bird, included to account for repeated observations of the same bird. Posterior checks indicate that the inclusion of a random slope for each bird ID did not improve the model fit, and as such we do not assume non-independence between the effects that flock size and distance have on each recorded bird. We do assume that each recorded bird may exhibit a unique intercept, reflecting variations between the proportion of time birds' spend vigilant. The random intercepts are drawn from a population distribution defined in line (3.16d). We set the prior for β_0 with the assumption that birds spend approximately half of the time head-down, and allow for variation; line (3.16d). The average population variation in the proportion of time birds spend vigilant, from one recorded bird to another is captured by σ in line (3.16e). Here, we use the exponential distribution as a weakly regularizing prior to estimate the variation across recorded birds (bird ID). The predicted effect that flock size has on the proportion of time a bird spends vigilant is defined by β_1, and the effect that distance from cover has on the proportion of time a bird spends vigilant is defined by β_2. We set prior distributions of β_1 and

β_2 in respective lines (3.16f) and (3.16g) to reflect our assumptions that the proportion of time a bird spends vigilant should generally decrease with flock size and distance from cover. Here, we also allow for possible positive relationships between the dependent variable and each regressor. The prior for ϕ is left uninformed. We ensured that all relationships predicted by the priors were reasonable on the (link function) untransformed scale, (i.e., the prior predicted proportion of time head-up remained bounded between zero and one) using the "rethinking" package in R [275, 279]. The model was fitted over a total of 292 observations, with 42 different recorded birds (bird IDs). We conducted posterior predictive checks using the bayesplot package in R [275, 280], and checked for convergence of the MCMC algorithm implemented in stan [276] for each coefficient by checking that the effective sample sizes, $\hat{R} = 1$, and by checking trace plots of the Markov chains.

Results of model (3.16) are summarised in Table 3.2 and illustrated in Fig 3.9. The posterior predicted distribution of β_1 has mean = -0.017 $\pm$ 0.003 (95% CI $[-0.02, -0.01]$), indicating that the proportion of time an individual bird spends vigilant (head-up) decreases with increasing flock size; Fig 3.9a. We find the sharpest declines in individual vigilance occur in smaller sized flocks. At average distance from cover (45.1 meters $\pm$ 22.1), with an increase in flock size from 1 to 10 we find a predicted decrease in the proportion of time a bird spends vigilant of 3.27% (95% CI, $[1.85, 4.84]$); when flock size increases from 11 to 20 we find a predicted decrease in the proportion of time a bird spends vigilant of 3.03% (95% CI, $[1.76, 4.39]$); and when flock size increases from 21 to 30 birds, we find a predicted 2.78 % (CI, $[1.67, 3.91]$) decrease. Whereas, for increases in flock size from 61 to 70 birds we find a posterior predicted 1.76% (95% CI, $[1.30, 2.11]$) decrease in the proportion of time a bird spends vigilant.

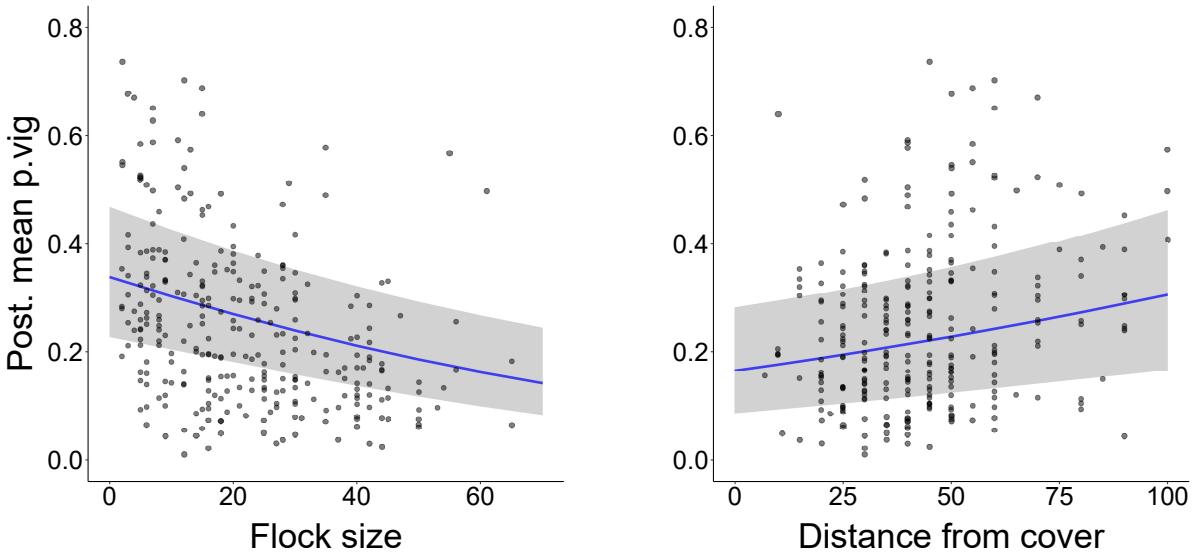
The posterior predicted distribution of β_2 has mean = 0.009 $\pm$ 0.002, (95% CI $[0.005, 0.011]$). This finding indicates that the proportion of time an individual spends vigilant increases with increasing distance from cover. The size of the effect that distance from cover has on the proportion of time a bird spends vigilant is relatively uniform over all distances from cover. We find that for an average sized flock (22 birds $\pm$ 13), an increase in distance from cover from 1 to 10 meters is predicted to result in an 1.26% (95% CI $[0.77, 1.61]$) increase in the proportion of time an individual spends vigilant. When distance from cover increases from 11 to 20 meters we find a predicted increase of 1.34% (95% CI $[0.8, 1.83]$), and when distance from cover increases from 21 to 30 meters we find a predicted increase of 1.42% (95% CI $[0.82, 1.98]$) in the proportion of time a bird spends head-up. At greater distances from cover the size of effect is similar: increasing distance from 51 to 60 meters results in a predicted 1.66% (95% CI $[0.89, 2.48]$) increase in the proportion of time an individual is vigilant, and increasing from 61 to 70 meters we predict a 1.74% (95% CI $[0.91, 2.61]$) increase.

We see from Table 3.2 that the error term of the intercept coefficient (second level residual standard deviation) is estimated to be 0.24 (95% CI $[0.08, 0.41]$) indicating deviation between the random intercepts; Fig 3.10.

TABLE 3.2: Assessing the beta regression of equation (3.16)

Regression results for equation (3.16)					
			95 % Credible interval (CI)		
Coefficient	Estimate	Estimate error	CI_{lower}	CI_{upper}	$\hat{R}$
Group level effects					
$\sigma(\beta_0)$	0.24	0.08	0.08	0.41	1
Population level effects with logit link function for u_i					
β_0	0.25	0.15	0.19	0.29	1
β_1	-0.02	0.00	-0.02	-0.01	1
β_2	0.01	0.00	0.00	0.01	1
Family parameters with identity link function for ϕ.					
ϕ	7.02	0.52	6.15	8.29	1

Overall, with regards to how the proportion of time a bird spends vigilant changes with flock size, model (3.16) provides support for the full theoretical model predictions (including vigilance, dilution and confusion effects) of section 3.3.1; see panels (i) in Fig 3.1 and Fig 3.2. We also find support for the reduced theoretical model in which the effects of vigilance are removed (see panels (ii) of Fig 3.1 and Fig 3.2). The data analysis does not support either of the theoretical models in which dilution or confusion effects are removed. When considering how the proportion of time a bird spends vigilant changes with distance from cover, the results of model (3.16) do not support our theoretical model predictions (see Fig 3.7).



(A) (B)

FIGURE 3.9: Posterior predicted mean (± 95% CI) proportion of time birds spend vigilant (Post. mean p.vig), depending on flock size and distance from cover. Plotted are predictions based on posterior predictive checks generating draws from the random effects distributions of model (3.16), using out of sample birds. (A): The posterior predicted proportion of time a bird spends vigilant (blue) ± 95% CI (shaded) with changes in flock size. (B): The posterior predicted proportion of time a bird spends vigilant (blue) ± 95% CI (shaded) with changes in distance from cover.

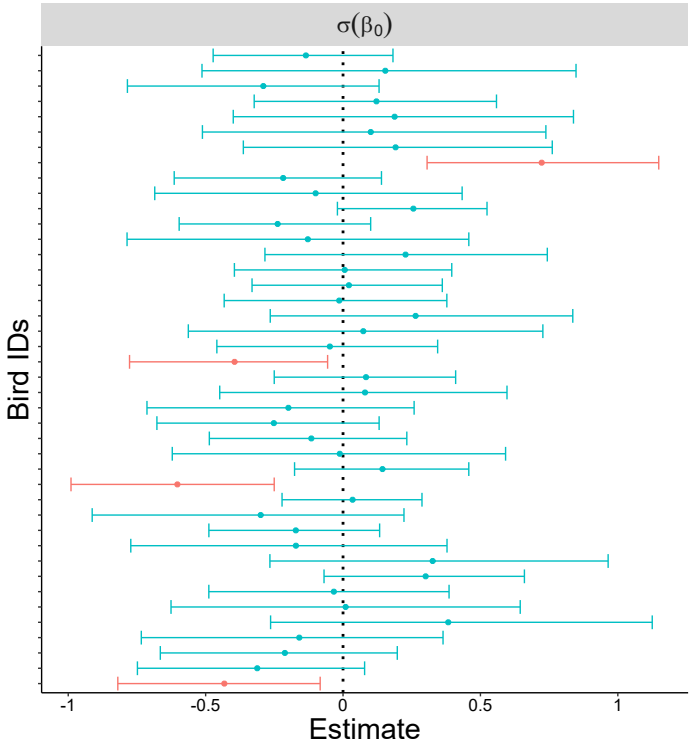


FIGURE 3.10: Random effects in model (3.16). Plotted are the posterior predicted deviations from the population intercept (points) ± 95% CI (error bars) of each recorded bird (bird ID). Red error bars show 95% CIs not overlapping zero. Blue bars overlap zero. Black dotted verticals show zero. With 95% credibility, four birds exhibit an intercept different from the population mean.

Prediction 1 & 3: investigating the variation between, and mean of, inter-scan intervals given flock size and distance from cover.

In section 3.3.1 the theoretical model predicts that the average inter-scan interval in a flock should increase with flock size, and that variation in the birds' inter-scan intervals should be lowest in small and large flocks, and highest in mid-sized flocks. For a simple qualitative comparison of this theoretical prediction to the data, we bin the data into "flock size classes" from 1 to 60 birds with increments of 5 birds. Summary statistics show that mean inter-scan interval increases with flock size class, and that the most pronounced variation between birds' inter-scan interval occurs in mid-sized flock size classes; Fig 3.11a and Fig 3.11b. These findings qualitatively support the predictions of the full model (see Fig 3.3a and Fig 3.3b), and the model configuration with the effects of vigilance removed (see the blue bars of Fig 3.3c and Fig 3.3d).

In section 3.3.3 the theoretical model predicts a relationship between the average inter-scan interval over a range of distances from cover, as well as the variation in inter-scan intervals in this range. For a simple qualitative comparison, we bin the data into "distance classes" from 1 to 60+ meters from cover, with increments 10 meters from cover. Summary statistics show that average inter-scan interval, as well as variation in inter-scan intervals, is relatively invariant with increasing distance class; Fig 3.11c and Fig 3.11d. These findings do not qualitatively support the theoretical model predictions (see Fig 3.8a and Fig 3.8b).

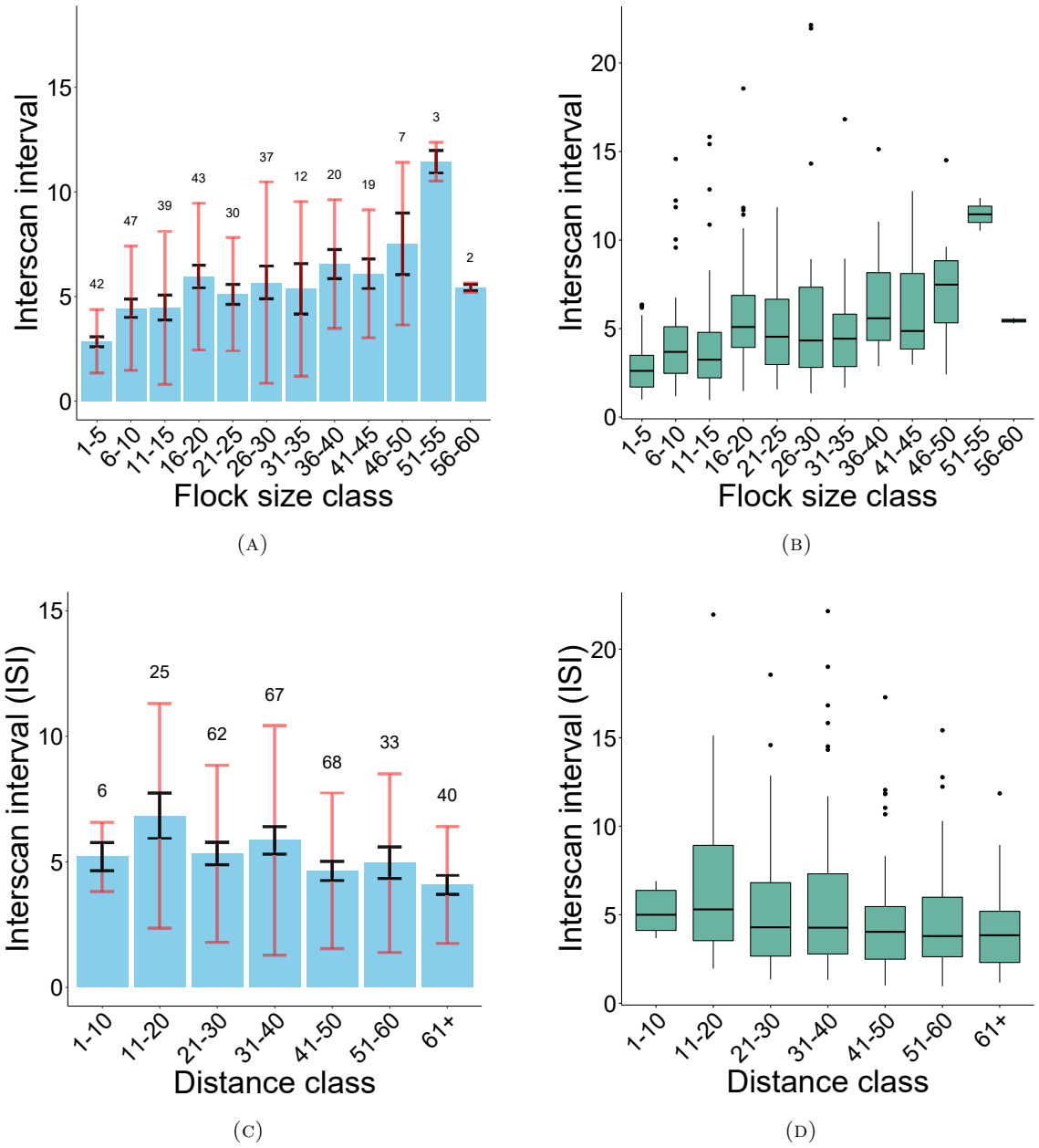


FIGURE 3.11: Summary statistics of inter-scan interval (ISI). (A): Blue bars show the average inter-scan interval for each flock size class. Red error bars show standard deviations and black error bars how the standard error. Sample sizes for each flock size class are indicated by the numbers above the bars. (B): Box-plot illustrating the quartiles of the inter-scan intervals with different flock size class. We see the largest inter-quartile range for flock size class 26-30. Outliers occur greater than the maximum for nearly all flock size classes. We also see that the median inter-scan interval tends to increase with flock size class. (C): Bar plot showing the mean (blue bars), and variation of (black standard error, and red standard deviation bars) inter-scan interval with distance class. Sample sizes over each bar. (D): Box-plot illustrating the quartiles of the inter-scan intervals with different distance from cover class. We see that the median inter-scan interval is similar across distance classes.

More robustly, we consider a Bayesian regression. We seek to assess whether the independent variables of flock size (Flock size) or distance from cover (Distance from cover) affect inter-scan

interval (ISI). The distribution of bird inter-scan intervals in the data is positively skewed (exhibits a large right tail). We find normally distributed residuals from a linear model with a log-transformed dependent variable; inter-scan interval, and above mentioned independent variables. As such, we regress the dependent variable of log(ISI). To assess the potential affects that flock size or distance from cover have on inter-scan interval, we model the mean of inter-scan interval u_i using a Gaussian model of the form

$$\begin{aligned}
\log(\text{ISI}_i) \sim & \mathcal{N}(u_i, \sigma_i) && \text{(3.17a)}\\
u_i = & \beta_{0,\text{bird ID}_i} + \beta_1 \text{Flock size}_i + \beta_2 \text{Distance from cover}_i + && \text{(3.17b)}\\
& \beta_3 \text{Flock size}_i \text{Distance from cover}_i && \text{(3.17c)}\\
\beta_{0,\text{bird ID}} \sim & \mathcal{N}(\beta_0, \rho) && \text{(3.17d)}\\
\beta_0 \sim & \text{Log-Normal}(0.15, 0.25) && \text{(3.17e)}\\
\rho \sim & \text{Exponential}(1) && \text{(3.17f)}\\
\beta_1 \sim & \mathcal{N}(0.01, 0.01) && \text{(3.17g)}\\
\beta_2 \sim & \mathcal{N}(0.01, 0.01) && \text{(3.17h)}
\end{aligned}$$

where we include $\beta_{0,\text{bird ID}}$ as a random intercept to account for pseudo-replication. The prior distribution of inter-scan interval intercept β_0 is defined by a log-normal distribution to ensure that the average length of time between scans remains strictly positive. We choose the log-normal distribution to have a mean, which upon exponentiating, constitutes what we consider to a be good estimate of the average length of time between scans, and set standard deviation to allow for considerable variation in inter-scan interval, while ensuring the variable is bounded by the minimum and maximum observed values. We assume a weakly regularizing prior to estimate the variation between the recorded birds' inter-scan intervals, given by ρ in line (3.17f). The respective effects of flock size and distance from cover on the mean inter-scan interval are described by the β_1 and β_2 terms. We assume that both flock size and distance to cover have weak positive effects on inter-scan interval, and as such define respective priors for β_1 in line (3.17g) and β_2 in line (3.17h). Through the coefficient β_3 we investigate whether either of the effects that flock size or distance from cover have on inter-scan interval, depend on the value of the other regressor. In other words, we check whether the effect that flock size has on inter-scan interval may vary with different distances from the predator, or whether the effect that distance from the predator has on inter-scan interval may vary depending on the flock size. The prior for β_3 is uninformed. All priors which we defined are checked using the rethinking package [279] in R [275].

The theoretical model predicts an upside down U-shaped relationship between the variation of inter-scan intervals within a flock, and flock size, as well as a nonlinear relationship between the variation of inter-scan intervals and distance from cover. As such, we are interested in a potentially nonlinear relationship between both predictors and the standard deviation of inter-scan interval. We regress the standard deviation according to

$$\begin{aligned}
\sigma_i \sim & \mathcal{N}(\alpha_i, \xi) && \text{(3.18a)}\\
\alpha_i = & a_0 + a_1 f_1(\text{Flock size}_i, \text{Distance from cover}_i) && \text{(3.18b)}
\end{aligned}$$

where f_1 is a bivariate tensor spline [281] which is included to model possible nonlinear effects that flock size and distance from cover have on standard deviation of inter-scan interval. The priors for the coefficients a_0, a_1 are uninformed. Interpreting the coefficient a_1 is difficult. We emphasise that we only consider the nonlinear relationship described above for a qualitative comparison to the theoretical predictions of section 3.3.1 (see Fig 3.3a and Fig 3.3b). The model (3.17)-(3.18) was fitted over 293 observations of recorded inter-scan interval, with 42 different bird IDs. We conducted posterior predictive checks using the bayesplot package in R [275, 280], and checked for convergence of the MCMC algorithm implemented in stan [276] by checking that $\hat{R} = 1$ and by visualising trace plots of the Markov chains.

The results of model (3.17)-(3.18) predict that inter-scan interval increases with increasing flock size; Fig 3.12a. The predicted posterior distribution of β_1 has mean = 0.012 ± error 0.006 (95% CI $[0.0005, 0.024]$). We find the predicted posterior distribution of β_2 has mean = -0.005 ± error 0.003 (95% CI $[-0.012, 0.0003]$). As β_2 has a 95% probability of falling within a range that overlaps zero, we cannot rule out the possibility that distance from cover has no effect on inter-scan interval; Fig 3.12b.

The potentially nonlinear effects that flock size and distance from cover have on the predicted standard deviation of inter-scan interval, defined by the coefficient of the fitted bivariate tensor spline cannot be easily interpreted; we therefore describe the output of the model graphically. We illustrate how the predicted posterior standard deviation changes with flock size; Fig 3.12c, and with distance from cover; Fig 3.12d. In Fig 3.12c we see a relatively flat relationship between the posterior predicted standard deviation of inter-scan interval and flock size. In Fig 3.12d we see a slight concave upward and downward shaped relationship between posterior predicted standard deviation of inter-scan interval, and distance from cover. These results do not qualitatively support the theoretical model predictions.

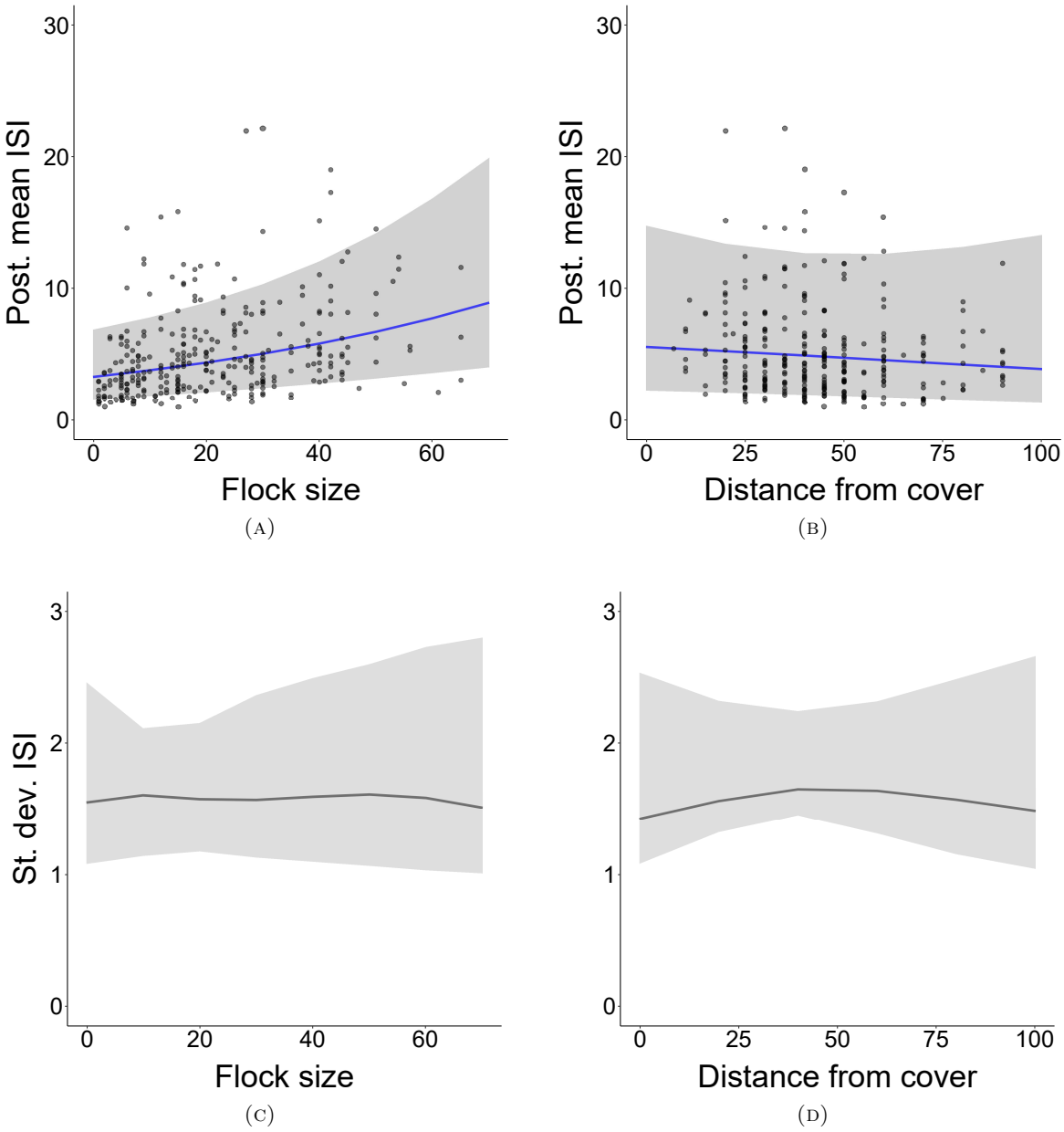
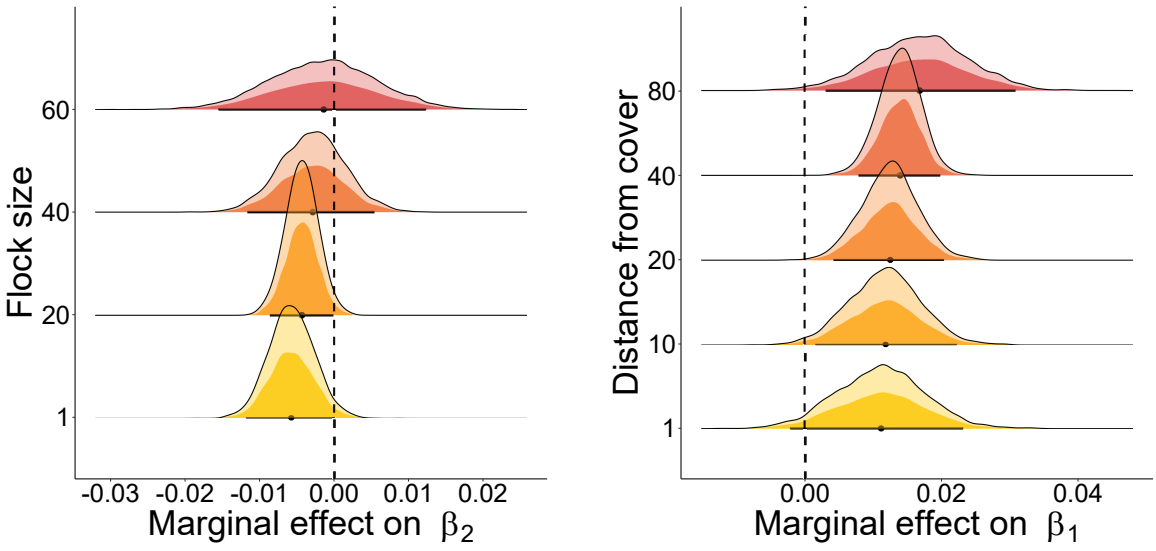


FIGURE 3.12: The posterior predicted (± 95% CI) mean (Post. mean) and standard deviation (Std. dev.) of inter-scan interval (ISI) predicted in model (3.17-3.18). To create the plots we conducted posterior predictive checks by generating data using draws of the random effects distributions with out of sample birds. (A): The posterior predicted inter-scan interval (blue line) ± 95% CI (shaded), with changes in flock size. (B): The posterior predicted inter-scan interval (blue line) ± 95% CI (shaded) with changes in distance from cover. (C): The posterior predicted standard deviation (solid black line) of inter-scan intervals with 95% CI (shaded), depending on flock size. (D): The posterior predicted standard deviation (solid black line) of inter-scan intervals with 95% CI (shaded), depending on distance from cover.

In Fig 3.13 we see the effect of β_3 through illustrating how distance from cover and flock size interact in their effects on inter-scan interval. In Fig 3.13b we see that for an average sized flock (22 birds), increases in distance from cover have positive effects on the mean of the posterior distribution of β_1. The marginal effect here indicates that flocks further from cover are predicted to exhibit a

greater increase in mean inter-scan interval with an increase in flock size, when compared to flocks nearer to cover. In Fig 3.13a we see that at the distance from cover (43 meters), an increase in flock size is predicted to have a positive effect on the mean of the posterior distribution describing β_2.



(A) (B)

FIGURE 3.13: Marginal effects of the interaction coefficient, β_3, of model (3.17). The plots are created by predictive posterior checks generating draws from the random effects distribution with out of sample birds. Both (A) and (B) illustrate the posterior predicted distributions describing marginal effects, with black dots showing the predicted mean marginal effect and horizontal bars showing 95% CI. The black dashed lines illustrate zero effects. (A): The marginal effect that flock size has on the mean of the posterior distribution of β_2. (B): The marginal effects of distance from cover on β_1. We see that the further from cover a flock is situated, the greater the increase in mean inter-scan interval with increasing flock size.

Overall, the results of Bayesian model (3.17)-(3.18) provide support for the theoretical prediction that inter-scan interval should increase with flock size. The results of the Bayesian analyses do not support the theoretical prediction that inter-scan interval should increase with an increase in distance from cover. The results presented here are not consistent with the theoretically predicted relationship between variation in inter-scan interval and flock size, nor do the results support the theoretical prediction regarding variation in inter-scan interval and distance from cover.

As shown in Fig 3.13a, the predicted effect that an increase in distance from cover has on inter-scan interval, varies depending on flock size. If flock size covaries with distance from cover, then distance from cover may not causally affect individual vigilance in redshank (as was found in model (3.16)). Instead, changes in flock size (which are known to influence vigilance [18]) may be the factor which affects individual vigilance in redshank. It is therefore of interest to check whether there is a relationship between flock size and distance from cover. To see whether distance from the predator has an effect on flock size we adopt a negative binomial regression (see section B.3 of appendix B for further details). We model flock size (Flock size) to depend on distance from cover

(Distance from cover) through the Bayesian Negative Binomial model defined by

$$\begin{aligned}
\text{Flock size}_i \sim & \text{NB}(u_i, \phi_i) && \text{(3.19a)}\\
\log(u_i) = & \beta_0 + \beta_1 \text{Distance from cover}_i && \text{(3.19b)}\\
\beta_0 \sim & \text{Log-Normal}(1.1, 0.1) && \text{(3.19c)}\\
\beta_1 \sim & \mathcal{N}(0, 0.001) && \text{(3.19d)}
\end{aligned}$$

with shape parameter ϕ given an uninformed (flat) prior, and defined by the identity link function. We choose the prior for β_0 because we assume an average sized flock should be of around 20 birds in size on the (link function) untransformed scale. Through our choice of a log-normal prior here, we allow variation in flock size, but ensure flock size remains non-negative. We choose the prior for β_1 as we have no reason to suspect a relationship between flock size and distance from cover. As such, we allow for either a positive or negative relationship, so long as the prior predicted changes in flock size (with distance from cover) do not result in flock size exceeding the maximum observed flock size in the data. The model was fitted over 292 observations of flock size and distance from cover combinations. We conducted posterior predictive checks using the bayesplot package in R [275, 280], and checked for convergence in the MCMC algorithm implemented in stan [276] by checking that $\hat{R} = 1$ and by visualising trace plots of the Markov chains.

Results of model (3.19) indicate that flocks tend to be of smaller size when foraging further from cover; Fig 3.14. We find that the posterior distribution of β_1 has mean= -0.006 $\pm$ 0.0018 (95% CI $[-0.01, -0.003]$), indicating that an increase in distance from cover has a negative effect on flock size. As distance from cover increases, we find the following predicted decreases in flock size: from 10 meters to 20 meters mean decrease = 1.54 $\pm$ error 1.86 (95% CI $[2.25, 9.54]$) birds; from 20 to 30 meters mean decrease = 1.43 $\pm$ error 0.49 (95% CI $[0.57, 2.48]$) birds, and for 30 to 40 meters mean decrease = 1.36 $\pm$ error 0.44 (95% CI $[0.56, 2.26]$) birds.

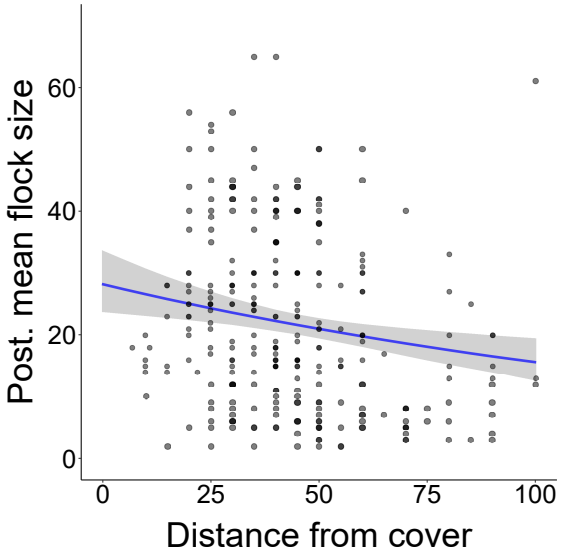


FIGURE 3.14: Predicted changes in flock size with distance from cover using the results of (3.19). Plotted is the posterior predicted mean (Post. mean) flock size (blue line) ± 95% CI (shaded) found by generating draws from the posterior distribution. We find a predicted decrease in flock size with an increase in distance from cover, as explained in text.

3.4.3 Prediction 2: data analysis

Prediction 2 of the theoretical model is assessed using the predation data set (described in section 3.4). Here, we conduct new analysis on this existing data set to assess whether temperature has an effect on the number of successful captures made, and the probability of a capture. We categorise observations of predator attacks into warm months (Oct, Nov, Dec 1989; Jan, Feb, Mar, Sep, Oct, Nov, Dec 1990; Oct, Nov, Dec 2001 and Jan, Feb, Oct, Dec 2002: N = 237 observations, mean temperature, $\langle T \rangle = 6.17 \pm 2.17$) and cold months (Jan, Feb, Dec 1991 and Jan, Feb, Mar 1992: N =134 observations, mean temperature, $\langle T \rangle = 4.32 \pm 0.78$).

TABLE 3.3: The proportion of captures across different temperatures

Proportion of captures over warm and cold months			
Flock size class	Sample size; N	Average temp; T	Proportion of captures; Prop_{cap}
Warm Months; overall weighted average $\langle T \rangle = 6.17$			
1-5	N =56	T=5.6	Prop_{cap}=0.161
6-10	N =26	T=6.5	Prop_{cap}=0.154
11-15	N=23	T=5.4	Prop_{cap}=0.217
16-20	N=22	T=6.1	Prop_{cap}=0.136
21-25	N=21	T=6.9	Prop_{cap}=0.000
26-30	N=18	T=6.8	Prop_{cap}=0.056
31-35	N=11	T=6.7	Prop_{cap}=0.000
36-40	N=13	T=6.7	Prop_{cap}=0.077
41-45	N=13	T=6.0	Prop_{cap}=0.077
46-50	N=12	T=6.8	Prop_{cap}=0.167
51-55	N=10	T=6.2	Prop_{cap}=0.100
56-60	N=12	T=6.5	Prop_{cap}=0.083
Cold Months; overall weighted average $\langle T \rangle = 4.32$			
1-5	N =35	T=4.2	Prop_{cap}=0.400
6-10	N =22	T=3.8	Prop_{cap}=0.318
11-15	N=11	T=4.3	Prop_{cap}=0.364
16-20	N=11	T=3.9	Prop_{cap}=0.273
21-25	N=13	T=4.5	Prop_{cap}=0.000
26-30	N=6	T=4.6	Prop_{cap}=0.333
31-35	N=14	T=4.2	Prop_{cap}=0.143
36-40	N=8	T=4.1	Prop_{cap}=0.250
41-45	N=5	T=4.0	Prop_{cap}=0.200
46-50	N=3	T=4.3	Prop_{cap}=0.333
51-55	N=5	T=4.0	Prop_{cap}=0.000
56-60	N=1	T=4.3	Prop_{cap}=0.000

A simple summary of the influence that temperature has on the proportion of successful attacks is given in Table 3.3. The table shows the proportion of successful captures made by the predator with flock size class, over the different temperature classes. Visualising these summary statistics indicates a broad and general increase in the proportion of captures during the colder climactic conditions; Fig 3.15. For comparison with model's theoretical predictions see Fig 3.5a and Fig 3.5b.

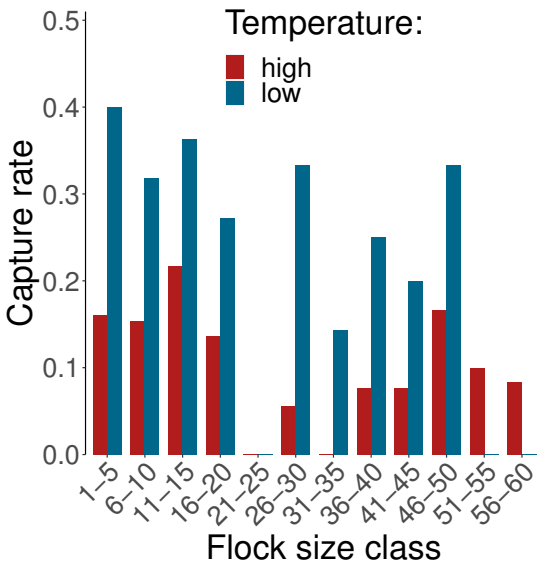


FIGURE 3.15: Observed proportions of birds captured given flock size, during warm and cold months. Data from Table 3.3. Warm months average temperature 6.2 degrees and cold months average temperature of 4.3 degrees.

Here we consider a Bayesian regression model with binary dependent variable; capture success "cap". Values of cap = 1 indicate a successful predator attack (the redshank is killed) and cap = 0 indicate a failed attack (the redshank is not killed). Temperature class is defined as a binary categorical variable; either "high" (reflecting warm months) or "low" (reflecting cold months). Assuming that temperatue class is the only factor which effects the probability of a capture, we apply a simple intercept only model where probability that a bird is captured, p_i, is given by

$$\text{cap}_i \sim \text{Binom}(n_i = 1, p_i) \tag{3.20a}$$

$$\text{logit}(p_i) = \alpha + \beta \text{Temp Class}_i \tag{3.20b}$$

$$\alpha \sim \mathcal{N}(-0.5, 0.25) \tag{3.20c}$$

$$\beta_2 \sim \mathcal{N}(-0.01, 0.01) \tag{3.20d}$$

with Temp Class = 0 used to indicate the high temperature class, and Temp Class = 1 used to indicate the low temperature class. Here, α corresponds to a average probability of capture given the high temperature class (i.e., Temp Class = 0), and $\alpha+\beta$ the average probability of capture given the low temperature class (i.e., Temp Class = 1). We set the prior for α to reflect our assumptions that the probability of a capture should be less than 0.5, and allow for variation. We leave prior for β uninformed so that temperature can have a positive or negative effect on the probability of capture. Posterior predictive checks were conducted using the bayesplot package in R [275, 280]. We checked for convergence in the MCMC algorithm implemented in stan by visualising trace plots of the Markov chains.

Results of the model (3.20) indicate that without consideration of any other factors which may influence the probability of a capture, there is a greater probability of capture during the colder climactic conditions; Fig 3.16. We find that the predicted mean probability of capture is 15.8% (95% CI $[7.8, 24.2]$) higher for the low temperature class, compared to the high temperature class.

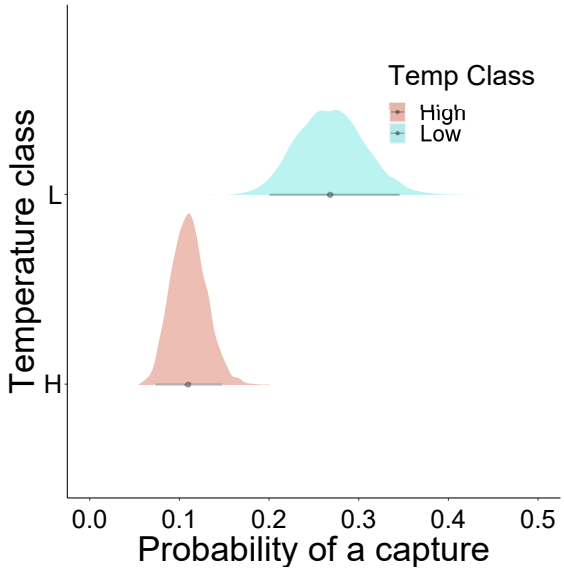


FIGURE 3.16: The posterior predicted probability of a capture depending only on temperature class. The plot is based on posterior predictive checks, conducted by generating data using draws from the posterior distribution. We see in blue the predicted posterior distribution of the probability of capture during the low temperature class, and in pink the predicted posterior distribution for the high temperature class. Points correspond to the posterior mean predicted probabilities of captures, and and error bars show 95% CI.

In greater detail, we theoretically expect that the probability of a capture should decrease with an increase in flock size, and decrease with an increase in temperature (see Fig 3.5a and Fig 3.5b). We additionally expect the probability of a successful capture to decrease with increasing distance from cover, as it is known that the majority of attacks come from an ambush predator which is most successful initiating an attack near a flock [25]. As such, we model the probability that a bird is captured p_i to depend on all three variables: flocks size, distance from cover, and temperature. We model the probability of a successful attack p_i using a Bayesian model of the form

$$\begin{aligned}
\text{cap}_i \sim &\text{Binom}(n_i = 1, p_i) && \text{(3.21a)}\\
\text{logit}(p_i) = &\alpha + \nu_i \text{Flock size}_i + \beta_2 \text{Temp Class}_i && \text{(3.21b)}\\
&+ \tau_i \text{Distance from cover}_i && \text{(3.21c)}\\
\nu_i = &\beta_1 + \beta_3 \text{Temp Class}_i && \text{(3.21d)}\\
\tau_i = &\beta_4 + \beta_5 \text{Temp Class}_i && \text{(3.21e)}\\
\alpha \sim &\mathcal{N}(-0.5, 0.25) && \text{(3.21f)}\\
\beta_1 \sim &\mathcal{N}(-0.01, 0.01) && \text{(3.21g)}\\
\beta_4 \sim &\mathcal{N}(-0.01, 0.01) && \text{(3.21h)}
\end{aligned}$$

where Temp Class = 0 indicates the high temperature class, and Temp Class = 1 indicates the low temperature class. Again, α corresponds to a average probability of capture given the high temperature class (i.e., Temp Class = 0), and $\alpha + \beta_2$ the average probability of capture given the low temperature class (i.e., Temp Class = 1). In this model flock size effects are also conditional

temperature class, defined by ν. The expression β_1Flock size$_i$ + β_3Flock size$_i$Temp Class$_i$ predicts the effect that flock size has on the probability of a capture through β_1 for the high temperature class (Temp Class = 0) and through $\beta_1 + \beta_3$ for the low temperature class. This allows us to compare the predicted posterior slope describing the relationship between capture probability and flock size, across temperature class. We assume that the probability of a capture decreases with flock size, and set the prior of β_1 to reflect this in line (3.21g). Model (3.21) indicates that the slope describing the relationship between the probability of capture and distance from cover from cover may vary across temperature class, as described by τ. The model predicts a mean of the posterior distribution describing the relationship between the probability of capture and distance from cover, β_4, when temperature class is high, i.e., the categorical variable of Temp Class = 0. For the low temperature class the effect that distance from cover has on the probability of a capture is given by $\beta_4 + \beta_5$. The expression τ allows us to check whether the effects of distance from cover confound the relationship between capture success, flock size and temperature class. If distance from cover does not confound the results we expect the slopes to be relatively uniform across temperature classes, i.e., the effect of the coefficient β_5 to be small with a 95% CI overlapping zero. We assume that the probability of capture should decrease with increasing distance from cover, and set the prior for β_4 accordingly in line (3.21h). We leave the prior for β_2 flat so that temperature can have a positive or negative effect on the probability of capture. The prior for β_3 is left flat so that temperature can have a positive or negative effect on the slope describing the change in the probability of capture with flock size, and we leave β_5 flat so that temperature can increase or decrease the rate at which the probability of capture changes with distance from cover. All of priors which we set were checked using the rethinking package [279] implemented in R [275]. We conducted posterior predictive checks using the bayesplot package in R [275, 280], and checked for convergence in the MCMC algorithm implemented in stan [276] by checking that $\hat{R} = 1$ and by visualising trace plots of the Markov chains.

The posterior predicted marginal effects that temperature class has on the probability of capture from model (3.21) are shown in Fig 3.17a. For an average sized flock of 25 birds, at average distance from cover of 31 meters, we find a predicted 14.58 % (95% CI [2.75, 27.22]) increase in the probability of a capture during colder climactic conditions. At 95% probability, we find that an increase in flock size is predicted to decrease the probability of a capture only in the low temperature class; β_1 (post. mean = -0.01 (95% CI $[-0.04, 0.01]$)) and $\beta_1 + \beta_3$ (post mean = -0.05 (95% CI $[-0.09, -0.02]$)); Fig 3.17b. We see the differences in the effects that an increase in flock size has on the probability of capture across temperature class, by comparing the posterior predicted distributions of β_1 and $\beta_1 + \beta_3$; Fig 3.18a. As flocks become sufficiently large the effects that an increase in flock size has on probability of capture are small across both temperature classes; Fig 3.17c and Fig 3.17b. In Fig 3.17c we see a bi-modal posterior distributions describing the marginal effects of flock size on the probability of capture for small to mid sized flocks. The two peaks reflect two slope coefficients of β_1 and $\beta_1 + \beta_3$. For larger flock sizes (see Flock size = 75 in Fig 3.17c) the distribution is uni-modal. This result supports the theoretical model prediction that capture probability should equalise over temperature classes as flocks become larger (see flock size greater than 40 birds in Fig 3.5a and Fig 3.5b). The posterior distributions of ν indicate that during colder months, there is on average a sharper decline in the probability of capture as flock size increases; Fig 3.18a.

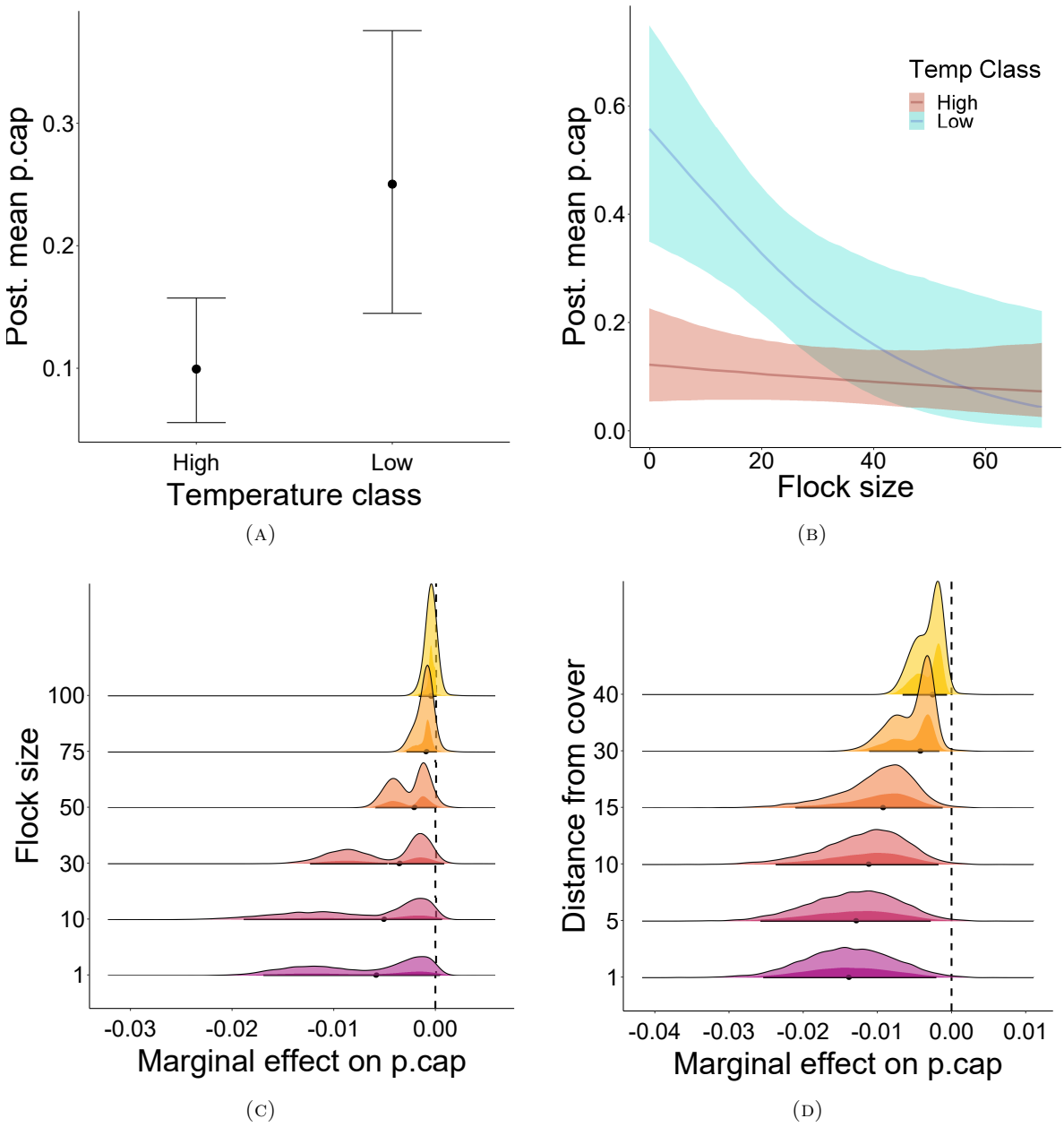
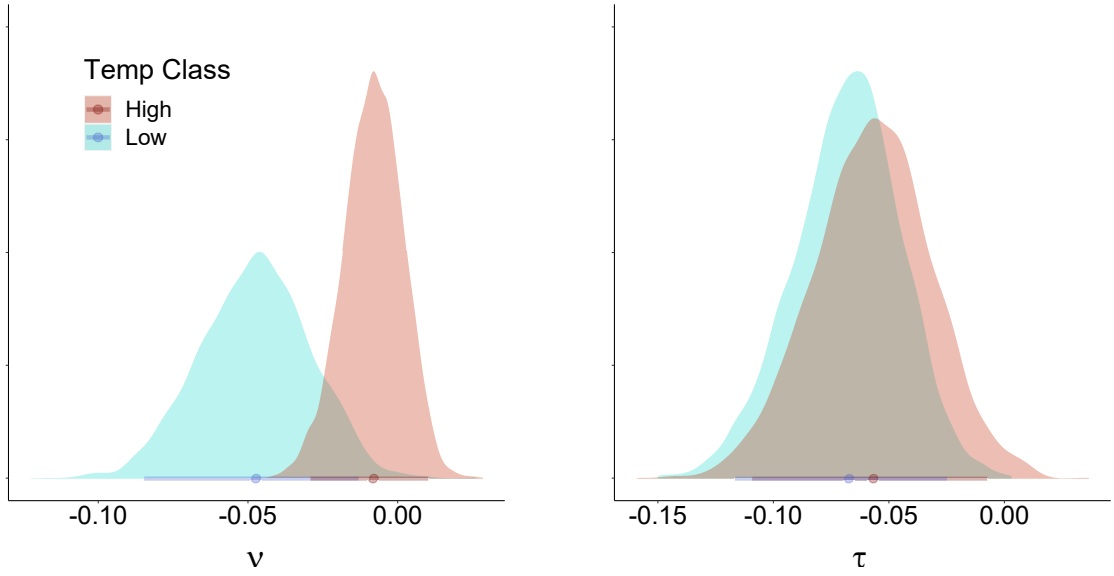


FIGURE 3.17: The marginal effects on the posterior predicted probability of a capture (p.cap), depending on (A) temperature class (B) flock size, and the marginal effects on the predicted probability of capture for (D) an average sized flock at varying distances from cover, and (C) an average distance from cover, with varying flock sizes. Plots (B),(C) and (D) are from posterior predictive checks, conducted by generating data using draws from the posterior distribution. Plot (A) is produced by calculating marginal effects. (A): Predicted posterior mean probability of capture (post. mean p.cap) for high and low temperature classes, based on an average sized flock and average distance from cover. Points correspond to the posterior mean predicted probabilities of captures, and and error bars show 95% CI. (B): The changes in the predicted mean probability of capture (lines) ± 95% CI (shaded) with increasing flock size, for each temperature class. (C) and (D) show distributions of the marginal effects (on the probability of a capture) as flock size and distance from cover increase. Black dots show posterior means, bars show 95% CI, and the black dashed lines describe zero effects. (C): The marginal effect on the probability of capture, given average distance from cover, for increases in flock size. (D): The marginal effect on the probability of capture, given average flock size, for increases in distance from cover.



(A) (B)

FIGURE 3.18: The posterior distributions of ν and τ of equation (3.21). The blue distributions show the predicted posteriors during the low temperature class. The pink shows the posterior distribution for the high temperature class. Respective posterior means are shown by the coloured dots, and 95% CI by horizontal bars. (A): The predicted posterior distributions of the coefficients $\beta_1+\beta_3$ (blue; low temperature class) and β_1 (pink; high temperature class), describing decreases in the probability of a capture with increasing flock size. We see that the distribution and mean of distribution are shifted left during the low temperature class. This shows a sharper decrease in the probability of capture with flock size, for the low temperature class. (B): Predicted effects on the probability of capture with changes in distance from cover are relatively uniform over temperature class. We see that the predicted distribution of $\beta_4+\beta_5$ (blue; low temperature class) is similar to β_4 (pink; high temperature class), as are the means. We expect the effects of distance from cover on the probability of capture to be similar over each temperature class.

During warm months birds fed on average at 16.1 meters from cover ($\bar{d} = 16.1, \sigma = 9.5, N = 237$) compared to 17.7 meters from cover during cold months ($\bar{d} = 17.7, \sigma = 8.3, N = 134$). Distance from cover therefore should not confound the results. Nonetheless, a potentially confounding relationship between flock size and distance from cover for either of the temperature classes, is captured by τ. The predicted effects that distance from cover has on the probability of a capture are defined through the coefficient β_4 during warmer months, and through $\beta_4 + \beta_5$ during colder months. We find that for both temperature classes, an increase in distance from cover is predicted to decrease the probability of a capture; the posterior distribution of β_4 has mean = -0.05 (95% CI $[-0.11, -0.01]$) and the posterior distribution of $\beta_4 + \beta_5$ has mean = -0.06 (95% CI $[-0.12, -0.02]$). We see that the effects that distance from cover has on the probability of capture appear uniform; Fig 3.19, as reflected by the similar posterior distributions describing the respective coefficients; Fig 3.18b. In Fig 3.17d we see that the posterior distributions describing the marginal effects that distance from cover has on the probability of capture are unimodal, except for around distances of 30m. This demonstrates that the effect that distance from cover has on capture probability is similar across temperature class. At 30m from cover, we see from Fig 3.19 that the slopes described by β_4 and $\beta_4 + \beta_5$ appear to differ by the most, explaining the bi-modal shape of the posterior

distribution at this distance as illustrated in Fig 3.17d. At 95% credibility, we cannot claim that the relationship between capture probability and distance from cover differs across temperature class. As the credible intervals of β_4 and $\beta_4 + \beta_5$ overlap we cannot rule out that the possibility that there is no difference in the effects that distance from cover has on the probability of capture, across temperature class. At 10 meters from cover it is 3.04% (95% CI $[-16.50, 11.39]$) probable that the effects that distance from cover has on capture probability differ across temperature class. At 20 meters from cover it is 4.18% (95% CI $[-26.42, 17.35]$) probable, and at 30 meters from cover it is 4.61% (95% CI $[-31.81, 18.75]$) probable. At each of 10,20 and 30 meters from cover, the posterior predicted probability that the slopes differ exhibits a wide 95% CI overlapping zero. Given that all successful captures occurred within 30 meters from cover, we therefore do not expect distance from cover to act as a confounding variable in the relationship between temperature class, flock size, and the probability that an individual bird is killed.

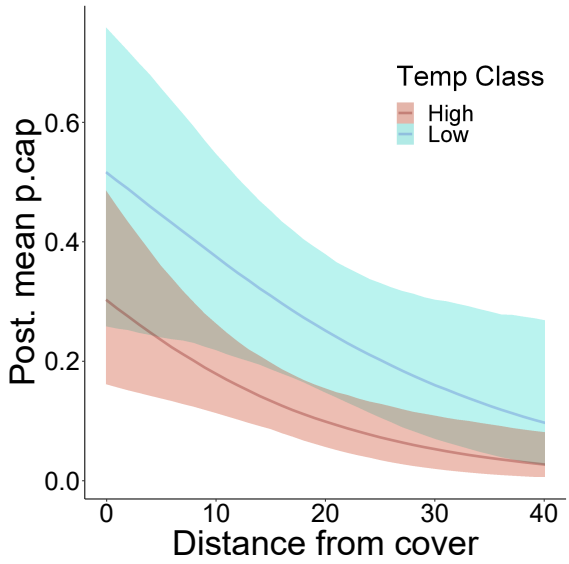


FIGURE 3.19: The predicted relationship between the probability of a capture and distance from cover, over each temperature class. Predictions are made from posterior predictive checks generating draws from the posterior distribution of model (3.21). The blue line describes the posterior predicted probability of capture (95% CI shaded) with changes in distance from cover for the low temperature class. The effect is defined by the mean of the distribution of $\beta_4 + \beta_5$. The pink line represents the predicted probability of a capture (95% CI shaded) given the high temperature class, with effect defined by β_4. The slopes are similar, suggesting that the decrease in the probability of capture with increasing distance from cover is not affected by temperature class.

Without identifying any other confounding factors, we conclude that the results of this section are qualitatively consistent with the predictions made by the theoretical model: that there is an increase in the probability of a capture during harsher climactic conditions.

3.5 Discussion

In this chapter we have compared the theoretical results of the model of chapter 2 to the vigilance behaviour of the wading bird redshank. We qualitatively assess the theoretical predictions of the model relating to how individual vigilance should change with group size, environmental conditions,

and perceived level of threat. By conducting new data analysis, we show that the theoretical model is able to qualitatively predict the anti-predator behaviours of redshank when all three principal anti-predation effects of collective vigilance, dilution and confusion, are modelled in concert, and when the effects of vigilance (collective detection) are removed. When accounting for all three principal effects, two of the three theoretical predictions are supported by data analysis: vigilance decreases with both flock size and with decreases in energetic state. The model prediction that vigilance should decrease with distance from cover is not empirically supported. Here, we provide an explanation for why this is the case, and argue that despite the lack of empirical support, the model prediction is still theoretically grounded.

3.5.1 Vigilance decreases with flock size

The theoretical model presented in section 3.2.1 and parametrised in section 3.2.2, predicts that individual vigilance in redshank should decrease with increasing flock size. Our data analysis indicates that individual vigilance in redshank decreases with an increase in flock size. The majority of field studies also find decreases in vigilance with increasing group size [2, 127, 146]. An important aspect of our theoretical prediction is based on the assumption of an unequal dilution of risk between vigilant and non vigilant birds. We explain this as follows: in small flocks, vigilant redshank react quicker and are less likely to be targeted, relative to non-vigilant individuals. Vigilant birds experience a significant fitness advantage through a preferable dilution of risk. In larger flocks, having many neighbours offers all individuals sufficient safety. Residing in a larger flock therefore decreases the relative magnitude of the effects that risk dilution exert on individual survivorship. With increasing flock size, the relative advantage of early detection (and vigilance) decreases.

The mathematical model that we present also predicts (by a mixed state of vigilant and non-vigilant birds) variation in the vigilance behaviours between individual redshank, for a range of flock sizes. In a summary analysis of the data we find most pronounced variation between the vigilance behaviours' of birds in mid-sized flocks; Fig 3.11a. Comparing the standard error terms of Fig 3.11a to the model predictions shown in Fig 3.3, we see a qualitatively similar pattern: small flocks contain little variation in vigilance levels, mid-sized flocks contain the highest variation in vigilance levels, and large flocks contain little variation in vigilance levels. We interpret this result as indicative of small flocks being predominantly composed of vigilant birds (low inter-scan interval, low variation of inter-scan intervals), mid-sized flocks containing a mixture of the vigilant and non-vigilant behaviours (mid inter-scan interval, high variation of inter-scan intervals), and large flocks being predominantly non-vigilant (high inter-scan interval, low variation of inter-scan intervals), as was predicted by the model. This result should be taken with caution due to smaller sample sizes in the smaller and larger flock size classes of the data. In applying a Bayesian analysis (see Fig 3.12c), we do not find a predicted relationship between standard deviation of inter-scan intervals and flock size which is qualitatively consistent with the theoretical model predictions. Overall, we interpret the results of the data analysis as very tentative support for the model prediction: mid-sized flocks contain the greatest mixture of two distinct behavioural phenotypes; one which prioritises vigilance, and the other which prioritises feeding.

The mixed vigilance flock predicted by the mathematical model is explained as follows: in mid-sized flocks, we model fitness benefits arising through both early detection (since flock size is

not too large) favouring the vigilant behaviour, and through foraging gains favouring non-vigilant behaviour. Mid-sized flocks are not large enough for risk dilution to offer sufficient safety for all birds: individuals which choose to be vigilant receive a significant relative fitness benefit through early detection, i.e., unequal dilution. At the same time, mid-sized flocks are not small enough for the fitness benefit of unequal dilution to be sufficient for all birds to choose the vigilant behaviour. Thus, some birds prioritise feeding, while others prioritise vigilance. Previous observational studies have noted that the relative magnitudes that the effects of collective detection and dilution effects have on predation risk, change with group size [31, 136, 146, 227]). Here we present an explicit theoretical framework which explains how, depending on group size, the magnitude of effect that each anti-predation mechanism has on predation risk varies. As such, we show how changes in the relative efficacies of the anti-predation effects, can predict changes in prey behaviour with changes in group size.

Through analysis of sections 3.3.1 and 3.4.2 the research presented here provides support for the hypothesis that collective vigilance, dilution and confusion effects are complementary anti-predation effects operating together to select for anti-predator behaviours in grouping individuals. We demonstrate that theoretically excluding either of the effects of unequal risk dilution or confusion results in theoretical predictions which the data are qualitatively less consistent with. In a similar theoretical grounding, Dehn [31] attempted to distinguish the importance of group vigilance and dilution effects. He found a good fit for a model in which non-vigilant individuals contribute to lowering individual risk through dilution effects, but not through vigilance. Dehn's conclusions support our findings that both of the effects of dilution and vigilance may be important in explaining vigilance behaviours in groups: dilution producing the dominant effect in larger group sizes, and vigilance effects dominating in smaller group sizes. When the effects of vigilance in the form of collective detection are not considered in our theoretical model, we still theoretically predict that an increase in flock size will a negative effect on the probability that an individual bird is vigilant; panels (iii) of Fig 3.1 and Fig 3.2. This prediction is qualitatively supported by the data; Fig 3.9a. Applying the theoretical model leads to a possible hypothesis that collective detection may not explain the observed decrease in individual vigilance in redshank with increasing flock size, as was found in the data. The negative group size effect associated with redshank vigilance could be explained by alternative mechanisms. Our model shows that with an increase in flock size, the effects of an inequitably shared safety in numbers (unequal dilution effect) and a lower probability of predator attack success (confusion effect) are sufficient to explain the reduction in individual vigilance. Moreover, we show that it is important to also theoretically consider confusion effects. When removing the effects of confusion, the theoretical model predicts an individual bird to have a very high propensity to be vigilant, and that vigilance is unaffected by changes in flock size (see panel (iv) of Fig 3.1 or Fig 3.2). This prediction is not supported by the data. Confusion effects are likely to reduce the need for individual vigilance in groups due to multiple group members reducing the predators ability to pick any one target [117, 170]. Our model predicts a higher probability for individual vigilance without such group size effects.

The qualitative similarities between the theoretical predictions of our model and the data should be taken with some caution. Several confounding factors are thought to moderate the effect of group size on vigilance. Perhaps the most studied confounding variable in the study of group vigilance is

food density. Foragers often aggregate in areas of high food density, and therefore food density could have a positive effect on group size. In particular, in redshank, feeding and vigilance are largely exclusive. In this case, a reduction in vigilance with group size could reflect the direct influence of food density on vigilance, rather than group size on vigilance. The parametrised theoretical model attempts to capture a direct relationship between food density and spatial location through using food quality as a function of distance from the predator (section 3.2.2). However, we do not consider the possibility that food density affects group size, i.e., areas of high food quality attract larger flocks.

3.5.2 Vigilance decreases with increased energetic needs

The model we present is able to capture complex state dependent effects on an individual's behaviour. By parametrising the benefits function (equation 3.1) with a fitness currency of energetic uptake rate [250, 262] the theoretical model is sensitive to the effects that harsher environmental conditions have on a bird's metabolic demands.

For a range of flock sizes the theoretical model predicts that an individual bird will be vigilant with a lower probability. Because the probability that a bird is vigilant x, translates to an expected number of vigilant flock members, nx, equation (3.6) gives a probability that no flock member is attacked. The model predicts that any flock member will be vigilant with a lower probability in low temperatures, compared to high temperatures. This results in a greater probability that any flock member is captured in harsher climatic conditions. Our data analysis largely supports this prediction. We find a higher predicted mean capture rate during the low temperature class of averaged monthly temperatures; Fig 3.17a. Under more demanding metabolic conditions, to avoid starvation redshank are forced to prioritise feeding [204]. Increased time head-down increases the ability of the sparrowhawk to initiate an attack undetected. The result is a greater number of captures within any given flock. Results of the data analysis also show a steeper decrease in the probability of capture with flock size for the low temperature class; Fig 3.17b. This finding indicates that vigilance effects are most important in small flocks facing harsh climactic conditions. In small flocks, confusion effects do not significantly affect the raptors ability to make a kill [25], and dilution effects do not affect the probability that a capture is made. It would seem that the sharp reduction in the probability of capture predicted in the data analysis, is explained by the effects of collective detection. Given that low temperatures force birds to spend lengthy time periods head-down, any additional pair of eyes joining the flock may be extremely valuable in increasing the flocks ability to detect a predator. In contrast, in climactic conditions which subject redshank to less severe metabolic demands, i.e., at higher temperatures, birds do not have to forage with the same intensity. While birds do not spend relatively as much time head-down, they are less reliant on the vigilance of others', and additional vigilance to the flock is less valuable. Our finding that the reduction in capture rate during milder climactic conditions is less steep than during harsher climates indicates that vigilance effects are state dependent.

Other studies have examined the effects that energetic state have on individual vigilance by considering time constraints on foraging [62, 200]. For example, McNamara & Houston [32] included metabolic constraints in their model of vigilance, and in doing so found a better qualitative fit to the data of Pulliam [49]. In many grouping species the risk of starvation cannot be discussed

separately from vigilance, and predation risk. Foraging with metabolic constraints will therefore have significant implications for vigilance, and more general anti-predator behaviours. The simple reality that each biological problem will have it's own complex relationship between energetic state and predation risk makes it difficult to produce a simple model to encapsulate these effects. In this research, we show that the broad theoretical model of chapter 2 is able to reduce to a specific species dependent problem and capture state dependent effects. Further research on the effects of state on behaviour will be valuable to the field.

3.5.3 How vigilance changes with distance to predatory concealing cover

In line with the expectation that individual vigilance should increase with predation risk, our theoretical model predicts that the probability an individual redshank is vigilant decreases with increasing distances from the predator. In particular we predict that for any fixed flock size, the vigilance rate of a bird increases with an increase in distance from cover. This model prediction is not qualitatively reflected in the observational data. Our analysis of the data shows a slight increase in vigilance with increasing distance from cover (see Fig 3.9b).

One possible explanation why the theoretical model fails to capture the relationship between distance to cover and individual vigilance draws from the confounding variable of flock size [18, 136, 191]. Our data analysis showed that at a signifiant level, flocks which feed further from cover tend to be smaller than flocks that feed nearer cover (see Fig 3.14). While flock size decreases with increasing distance from cover, individual vigilance could be a sensitive reactive response to both flock size and spatial location. In smaller flocks, a reduction in the dilution of risk increases the need for (and benefit of) individual vigilance [25, 116]. Even at greater distances from the predator, due to the smaller flock sizes, individual birds within the flock may still be at considerable risk. Therefore, as flocks forage further from cover, there are two compensatory effects at work which modulate an individual bird's predation risk: a reduction in safety through a reduction in flock size, and an increase in safety through an increase in distance to cover. If the overall effect is an increase in predation risk, then individual birds may compensate for the increased risk by maintaining or increasing their vigilance levels. Alternatively, it might only be flock size that influences the vigilance behaviour in redshank without any causal relationship between vigilance and distance from cover. Vigilance is most valuable when near the predator (<30 meters). Within this spatial range, flock size decreases with increases in distance from cover. Decreases in flock size are predicted, and shown in this system (see Fig 3.9a), to result in an increase in individual vigilance.

A different hypothesis is that competition between flock members may increase as flocks become larger, resulting in a greater need for individual birds to invest in time spent feeding. If there is a decrease in prey availability with an increase in flock size, then individual birds are subject to lower intake rates in larger flocks. A lower intake rate could have the effect to encourage an individual bird to increase it's foraging intensity and potentially decrease it's vigilance [146]. With respect to redshank, this is very possible. Given that vigilance and feeding are largely exclusive, one could expect that an increase in foraging intensity directly results in a lower rate of vigilance for an individual bird. However, it has been previously shown in the system that there is no significant relationship between flock size and intake rate in the birds [202].

Other studies have found that vigilance decreases with increasing distance from predatory cover. For example, Beauchamp [282], using pecking rate as a measure of vigilance, found that semi-palmated sandpipers decrease their pecking rate when further from cover. Beauchamp concluded that pecking success rate decreased nearer to cover, and that the increase in pecking rate (and decrease in vigilance) nearer cover could be due to less food availability near the predator. Food availability could be a possible factor which helps explain the models failure to capture the relationship between vigilance and distance from cover. A dependency between food density and distance was explicitly parametrised in section 3.2.2.

It could well be that the discrepancy between the theoretical model prediction and the data analysis is a simple case of the biological problem being more complex than the mathematical approximation. The fundamental reason for which we argue this is due to the presence of another raptor on the site location [266]. Redshank face a constant rate of attacks from peregrine falcons, regardless of proximity to cover [23]. In the theoretical model, despite equation 3.6 being well suited function to describe flock evasion under a sparrowhawk attack (see [206] or [207] for justification of this), the function does not account for peregrine attacks. The peregrine threat changes a redshanks perception of risk over all distances from cover. Peregrines attack from all distances. In this light, our model parametrisation is almost certain to overestimate the significance of the reduction in predation risk experienced by redshank, as flocks forage further from cover.

A subtle change in the parametrisation of theoretical model could account for peregrine attacks. Given that sparrowhawks only attack flocks within the 30m from cover boundary, the evasion function (equation 3.6) could be re-parametrised as piecewise continuous with distance from cover. In such a regime, the effects of vigilance and confusion would both control the probability of flock evasion along with distance from cover while $d < 30$. Whereas for $d \geq 30$ the probability of evasion would be independent of distance from cover (peregrines attack from all spatial locations) and relatively unaffected by vigilance effects (since vigilance is not an appropriate defence against the peregrine); the only dominant effect controlling evasion would be group size (confusion effects). On this line of reasoning, the model would capture how the relative effects of group size and vigilance are sensitive to spatial location when influencing vigilance behaviours in redshank. As such, for $d \geq 30$ the model would predict that a reduction in group size only has the effect to increase the importance of unequal dilution effects, which favour vigilant behaviours. If flock size decreases with increasing distance to cover, this would explain why individual vigilance increases.

The explanation we provide here emphasises the theoretical motivation of our modelling: anti-predation effects in the form of vigilance, confusion and dilution are intertwined and must be considered to operate interdependently. If the only predator is a peregrine then (collective) vigilance in redshank is a less effective defence strategy. The importance of the effects of group size (confusion effects) on predator evasion increase relative to vigilance. While flock size is the dominant factor determining the probability that any individual is targeted, changes in flock size result in substantial changes in individual per-capita risk. This results in an increase in the importance of risk dilution. While sparrowhawks are the prominent predator, vigilance is always an effective anti-predator strategy because collective vigilance in redshank can generally provide protection for the entire flock. The effects of unequal dilution are only significant in small flocks which are less likely to detect the predator.

3.5.4 Concluding remarks

Overall, when used to make specific predictions in a real biological system, the model of chapter 2 has received mixed empirical support. We have demonstrated that our general theoretical model can be applied to a specific system and correctly predict key relationships between prey behaviour and predation risk. Our qualitative analysis to the data demonstrate the importance of considering all three anti-predation effects when predicting the behaviours in prey groups. Previous experimental studies have shown that at least two of the principal anti-predation effects may interact when influencing the behavioural choices of group members, although these studies tend to be on specific species [31, 146]. A general and widely applicable theory is lacking. To our knowledge, this is the first time that a broad theoretical framework, able to predict the anti-predator behavioural responses in groups as the result of three interacting prominent anti-predation effects, has been compared to observational data. By capturing the trade-off between preferable risk dilution and unfavourable foraging returns that arise from a group member's behaviour [31, 146, 263], our model provides a general framework to explain the behaviours for a wide range of grouping species. We conclude that the model of chapter 2 provides an important step in understanding the anti-predator behavioural decisions for wide ranging grouping animals.

Chapter 4

The adaptive dynamics of anti-predator strategies

Abstract

Many grouping animals demonstrate highly flexible behaviours when responding to predators. In this chapter we analyse the continuous extension of the model presented in chapter 2. We focus on a foraging-predation trade-off given by case 4 of Table 2.1, here defined by continuous smooth functions relating strategic investment to individual payoff. Subsequently, conditions are clarified under which the mixed strategy and the continuous strategy interpretation of flexible behaviour produce the same behavioural outcomes, and when not. Through our analysis of a continuous strategy, we show that there are three qualitatively distinct possibilities which characterise the behaviours emerging from groups. We clarify when all group members choose an extremal behaviour, an intermediate behaviour, and when behaviours within the group become differentiated. Our key findings are threefold. First, we show that a richer array of behaviours emerge in groups when an individual's experience of predation risk and foraging gains continuously varies with its strategy. We show that when individuals experience binary levels of predation risk and foraging gains with continuously varying probabilities, fewer behavioural outcomes are possible. Second, we show how simple properties of payoff functions that define the above mentioned trade-off govern the continuous strategy dynamics. Third, we show how there is a subtle but important difference in behavioural diversity within groups that result from mixed and continuous strategies. Overall, we show how a simple mathematical model can reproduce highly flexible and complex anti-predator behaviours in groups of prey.

4.1 Introduction

In chapter 2 a model is presented that predicts the strategic behaviours of individuals within groups of prey, subject to predation. The model of chapter 2 assumes that group members have a fixed and discrete choice of strategies. Anti-predator behaviours in foraging groups are however most likely to be highly flexible [4, 27, 219], and liable to continually change with predation pressures

and foraging demands [18, 32, 46] (and see the results presented in section 3.4.1 of chapter 3). Flexible behaviours may be better captured by a continuous strategy. In this chapter, we build on the theoretical approach of chapter 2 in the mutual modelling of multiple anti-predation effects and grouping benefits, but allow individuals to flexibly choose and adjust their pro-active anti-predator behaviours.

Adaptive dynamics, introduced in section 1.3, provides a powerful set of techniques to model changes in continuous strategies. Despite traditionally being applied to study the slow evolution of quantitatively fixed traits, adaptive dynamics can model biological problems on faster timescales (see the Introduction section 1.3 or [60, 103, 112, 283]). Here, we apply adaptive dynamics to model the anti-predator behaviours that emerge in groups of prey. In lieu of the binary choice of strategy presented in chapter 2; "to contribute towards anti-predator activities" or not, here we consider a continuous strategy of "contribution investment" into anti-predator activities. For example, a continuous strategy could be an individual's investment effort into predator detection [32, 49, 188], the amount of energy it spends on raising alarm calls [59] or its contribution towards coordinated motion [179].

In contrast to the continuous strategy modelled under adaptive dynamics, behavioural flexibility can also be considered in the sense of a discrete strategy space [41]. Using the discrete strategy space, one can approximate flexible behaviours by assuming that individuals "probabilistically" adopt a particular fixed behaviour, i.e., the mixed strategy [35]. The model of chapter 2 may be interpreted in the sense of mixed strategies: with some probability, individuals invest all (contribute and play C) or nothing (do not contribute and play D) into anti-predator activities. Here, we refer to this mixed strategy as the "contribution probability". The contribution probability and the contribution investment thus provide two different mechanisms for modelling flexible anti-predator behaviours. Both interpretations of strategy are used to reflect the behaviours in groups of prey (for example see [59] for the "probability" to raise alarm calls, or [32] for "investment" into vigilance). Nonetheless, how the predicted behaviours emerging from groups under each mechanism possibly differ is unclear.

The key research tasks of this chapter are (i) to find "how much" effort group members should choose to invest into anti-predator behaviours, and (ii) to show that by allowing groups members a flexible effort level, we see qualitatively different behaviours compared to when individuals can only invest all or no effort with flexible probabilities. We aim to compare the contribution probability (presented in chapter 2) and the contribution investment (analysed in subsequent sections) mechanisms. To achieve these tasks we focus on case 4 of the trade-off Table 2.1, chapter 2, whereby the more an individual invests into an anti-predator strategy the more that individual experiences both reductions in its per-capita predation risk and in its general grouping benefits. Such foraging-predation trade-off's are common in groups of prey [18, 31, 284]. We analyse a range of possible different qualitative structures that reflect this trade-off. The respective emerging behavioural dynamics are analysed and compared to the results of chapter 2. We show which trade-off structures result in a qualitative equivalence between the mixed strategy of chapter 2, and the continuous strategy of this chapter. Importantly, we aim to show which structures do not produce qualitatively similar results, and the consequences for group members. We explore a range of possible

qualitatively different behavioural outcomes and hope to provide a purely mathematical explanation of the anti-predator behavioural phenotypes in groups, as well as how, and why, groups can come to exhibit distinct behavioural phenotypes.

4.2 Methods

We consider a population of size N. Within the population, individuals interact in randomly formed groups of size n. Each individual adopts a one-dimensional strategy z from a continuous space $S = [0, 1]$. As explained above, under the contribution investment, z is a measure of an individual's investment into anti-predator activities. Minimal (or "defector") investment corresponds to a strategy value of 0, and maximum (or "cooperative") investment corresponds to a strategy value of 1. Alternatively, when interpreted as a contribution probability, z indicates the probability that the individual invests 1, and otherwise 0. For comparison of the results of this chapter to the results of chapter 2 we note that the strategies C and D therein respectively correspond to investing 1 and 0 here. Each strategy, $z_i \in S$ has a population frequency given by $x_i(t) \in \mathcal{P} = [0, 1]$ at time t. Within the population we assume that there exist $M \geq 1$ distinct strategies at any one time. The state of the population is characterised by the strategy vector $\mathbf{z} = (z_1, z_2, ..., z_M) \in S^M$ and the frequency vector $\mathbf{x} = (x_1, x_2, ..., x_M) \in \mathcal{P}^M$, where $\sum_i^M x_i = 1$.

The strategies that we consider are purely behavioural. As such, we provide the following rational for our use of adaptive dynamics. Initially, we assume that all individuals in the population choose some strategy in S. Occasionally, a novel strategy appears in the population. We assume that novel strategies come into existence through individuals "exploring" their strategic choices. Exploration can be interpreted as an individual's propensity to randomly alter their strategy [40], sample alternative actions [21], or learn new behaviours not already exhibited in the population. We do not need to specify. What is of importance in the model is the emergence of a novel strategy. After an exploration event, a negligibly small section of the population adopt the novel strategy. If the novel strategy has an advantage over the existing resident strategy, then it spreads through the population (for example through imitation [60, 61]); if not the resident strategy remains dominant. The model therefore involves three distinct processes, each operating on a distinct timescale. First, on the very short "interaction" timescale individuals choose their preferred level of strategic investment which defines payoff. We assume that the population reaches a dynamic equilibrium in which all members choose some strategic level of investment. Second, on the intermediate "ecological" timescale we consider how upon an exploration event, the relative frequencies of a rare novel (or invader) strategy and the resident strategy (or strategies) in the population change via competitive interactions. The respective payoffs of each strategy determine whether either strategy tends to fixation. Finally, on the slowest "exploration" timescale, exploration events are assumed to occur infrequently. This ensures that contests between the previous resident and invader strategy are resolved before the next exploration event, i.e., before the next novel strategy appears.

4.2.1 Ecological dynamics

In this section we outline how competitive interactions between strategies in the population result in a dynamic equilibrium. We apply a continuous time ordinary differential equation of the form

$$\dot{x_i} = x_i\big(G(z_i;\mathbf{z};\mathbf{x}) - \bar{G}\big) \quad i = 1,,,\rho \tag{4.1}$$

to describe the success of any strategy z_i, i.e., the growth rate of x_i, in a population of ρ competing strategies. Here, $G(z_i;\mathbf{z};\mathbf{x})$ is the G-function [73, 74, 79] (and see the Introduction section 1.1.6) associated with the population. The dynamics of equation (4.1) reach equilibrium on the intermediate ecological timescale, which we assume almost always occurs before any new novel strategy can enter the population. All strategy values, z_i, in the population are therefore viewed as constants. We denote the mean payoff in the population as $\bar{G} = \sum_{i=1}^{\rho} x_i G(z_i;\mathbf{z};\mathbf{x})$.

For simplicity we assume that there are initially two strategies in the population; $\mathbf{z} = (z_1, z_2)$ and $\mathbf{x} = (x_1, 1-x_1)$ with z_2 is arbitrarily close to z_1. We henceforth refer to z_1 as the resident and z_2 the invader. Equation (4.1) admits two trivial hyperbolic equilibria: $\hat{x}_1 = 0$ and $\hat{x}_1 = 1$. There may also be interior solutions $\hat{x}_1^{\text{int}} \in (0,1)$. When $\hat{x}_1 = 1$ we have that

$$\begin{aligned} &G(z_1;\mathbf{z};\hat{\mathbf{x}}) - \bar{G} = \\ &G(z_1;z_1,z_2;1,0) - \bar{G} = 0 \end{aligned} \tag{4.2a}$$

and on the occasion that we consider an initially dimorphic resident population characterised by $\hat{x}_1^{\text{int}} \in (0,1)$, we have

$$\begin{aligned} &G(z_1;\mathbf{z};\hat{\mathbf{x}}) - \bar{G} = \\ &G(z_1;z_1,z_2;\hat{x}_1^{\text{int}},1-\hat{x}_1^{\text{int}}) - \bar{G} = 0. \end{aligned} \tag{4.3a}$$

Since the strategies do not change on the ecological timescale, equations (4.2) and (4.3) are equivalent to the replicator dynamics of chapter 2.

Payoff

Payoffs are based on equations (2.4) of section 2.2, chapter 2, but with a continuously varying level of strategic investment replacing the binary strategy choice. We apply the following G-function to describe the payoff for a focal group member with investment level y:

$$G(y;\mathbf{z};\mathbf{x}) = \underbrace{\gamma(y)}_{\text{Foraging benefits}} \left\{ \underbrace{g(y;\mathbf{z};\mathbf{x})}_{\text{Evasion}} + (1-g(y;\mathbf{z};\mathbf{x})) \underbrace{\left[\frac{(nx_y - 1)(1+\epsilon(y)) + \sum_{j=1}^{\rho} nx_j(1+\epsilon(z_j))}{\sum_{j=1}^{\rho} nx_j(1+\epsilon(z_j))} \right]}_{\text{Risk dilution}} \right\} \tag{4.4a}$$

where x_y is the frequency of strategy y in the population characterised by $\mathbf{z}$ and $\mathbf{x}$.

In line with equations (2.4), section 2.2 of chapter 2, the function $g(y;\mathbf{z};\mathbf{x})$ reflects the probability that the whole group evades the predator. We denote the term of equation (4.4a) under-braced

"Risk dilution" by $\mathcal{D}(y;\mathbf{z};\mathbf{x})$. The idea here essentially reflects the assumption of an unequal dilution of risk as presented in chapter 2. However, in this chapter targeting risk no longer depends simply on the relative frequencies of "cooperators" and "defectors" (see chapter 2). Instead we use $\epsilon(z_i)$ to ascribe a unique relative risk of dilution to each continuous level of investment (strategy) z_i. The function $\mathcal{D}(y;\mathbf{z};\mathbf{x})$ captures how a focal group member's relative dilution of risk depends on its strategy as well as all of the other strategies present in the group, i.e., the expected group composition. Mirroring chapter 2, the $\gamma(z_i)$ terms describe the foraging benefits a focal individual experiences as a continuous function of its strategic investment z_i.

Henceforth, we explicitly consider the payoff structure presented in chapter 2, Case 4 of Fig 2.1. Such a payoff structure is characterised by $\epsilon'(y) \leq 0$ and $\gamma'(y) \leq 0$. For comparison between the contribution probability and contribution investment we set $\epsilon(1) = \epsilon$ and $\epsilon(0) = 0$ as well as $\gamma(1) = \gamma_C$ and $\gamma(0) = \gamma_D$.

4.2.2 Adaptive dynamics

In this section we are interested in changes of the distribution of the investment strategies in the population on the exploration timescale. The investment strategy distribution over the population changes by imitation of the successful strategies, where every now and then a novel strategy enters the population and competes. We apply the adaptive dynamics [77, 103] to formalise these sequential resident-invader competitions. We begin with a monomorphic resident population with strategy z_1. Upon an exploration event, consider that a very small fraction ξ in the population adopt a novel strategy y. The invasion fitness of the novel strategy (see the Introduction section 1.3.3) in the resident population defined by $\hat{x}_1^{\text{res}} = 1$ is derived from equation (4.2), or equally the standard replicator equation:

$$\dot{\xi} = \xi(1-\xi)\big(G(y;z_1;1) - G(z_1;z_1;1)\big). \tag{4.5}$$

The invasion fitness can be expressed as

$$\lim_{\xi\to 0}\frac{\dot{\xi}}{\xi} = G(y;z_1;1) - G(z_1;z_1;1). \tag{4.6}$$

When $\hat{x}_1^{\text{res}} = 1$ we define

$$\lambda : \mathbb{R}_{[0,1]} \times \mathbb{R}_{[0,1]} \to \mathbb{R}_+ \tag{4.7a}$$

$$\lambda(z_1, y) := \begin{cases} G(y;z_1;1) - G(z_1;z_1;1), \text{ for } 0 \leq y, z_1 \leq 1 \\ 0, \text{ for } 1 < y, z_1 \text{ or } 0 > y, z_1 \end{cases} \tag{4.7b}$$

as the "monomorphic" invasion fitness [60, 61, 81]. Equation (4.7) is appropriate in the case of one resident strategy being challenged by one novel strategy, i.e., model (4.2) when $y = z_2$. In this case the invasion fitness is defined by the single eigenvalue of the monomorphic population equilibrium. Therefore the sign of $\lambda(z_1, y)|_{y=z_2}$ determines invasibility: if $\lambda > 0$ the population is liable to be invaded. Clearly, $\lambda(z_1, z_1) = 0$. In cases where the novel strategy enters a polymorphic population

with respective strategies $\mathbf{z} = (z_1, ..., z_{x_\rho})$ and frequencies $\mathbf{x} = (x_1, ..., x_\rho)$, we define the invasion fitness by $\lambda(\mathbf{z}, y) = G(y; \mathbf{z}; \mathbf{x}) - \bar{G}$.

While the resident population is monomorphic, the adaptive dynamics of z_1 are fully determined by the selection gradient $D(z_1) = \partial_y \lambda(z_1, y)|_{y=z_1}$. We identify singular strategies $\hat{z}_x$ as potential outcomes for z_1. In this research we classify singular strategies as evolutionary stable strategies (ESS or ESS-stable) in the sense of Maynard Smith [35], convergence stable (CS) [101], and continuously stable strategies (CSS) [285] (see conditions 1.57b and 1.62; Introduction section 1.3.5). Invasion analyses of the monomorphic adaptive dynamics are illustrated by pairwise invasibility plots (PIPs) and we provide simulations of invasion events by utilising the G-function. This approach allows us to simultaneously analyse both strategy and population dynamics [74] through the following system of equations

$$\dot{z}_i = \frac{\partial}{\partial y} \lambda(\mathbf{z}, y)|_{y=z_i} \tag{4.8a}$$

$$\dot{x}_i = x_i(G(z_i; \mathbf{z}; \mathbf{x}) - \bar{G}) \ , \ \ i = 1, ..., \rho \tag{4.8b}$$

Equations (4.8) produce the same limiting behaviour as the adaptive dynamics in monomorphic populations, and also provide a differential equation which captures branching phenomena [80, 81]. Implicitly we assume a separation of timescales, but we do not explicitly show this in the subsequent invasion simulations.

4.3 Analysis in monomorphic populations

Two complementary approaches can be used to study the interpretation of a flexible anti-predator strategy. The replicator dynamics of chapter 2 model the contribution probability. The approach of this chapter (as outlined in section 4.2.2) focuses on the contribution investment. Under the contribution probability an individual's strategy $\mathbf{z} = (z, 1 - z)$ is a probability (distribution) to invest 1 (or 0). Under the contribution investment, z is a continuous level of investment. To see how the results of each approach might differ it is useful to compare the respective gradients of selection and fitness functions. Henceforth we consider a monomorphic population with strategy defined by z.

4.3.1 A comparison of selection gradients

Under the contribution probability the invasion fitness of a novel strategy $\mathbf{y}$ is

$$\theta(z, y) := y f_C(z) + (1 - y) f_D(z) - \big(z f_C(z) + (1 - z) f_D(z)\big). \tag{4.9}$$

where $f_C(z) \equiv G(1; 1, 0; z, 1 - z)$ and $f_D(z) \equiv G(0; 1, 0; z, 1 - z)$ are respectively the cooperative and defector strategies of chapter 2. One finds that equation (4.9) leads to the selection gradient:

$$\frac{\partial}{\partial y} \theta(z, y) = f_C(z) - f_D(z) \tag{4.10a}$$

$$= \delta g(z) + \delta(1 - g(z))\{\mathcal{D}_C - \mathcal{D}_D\} \tag{4.10b}$$

where $\mathcal{D}_C, \mathcal{D}$ are defined by equations (2.3), section 2.2 of chapter 2. It is well known from the Bishop Cannings theorem [42] (see also [286, 287]) that any ESS $\hat{z}$ resulting from equation (4.10) is neutral to invasion. Convergence stability corresponds directly to the asymptotic stability of $\hat{z}$. Given case 4 of Table 2.1 we find the following conditions:

$$\frac{\partial^2}{\partial y^2}\theta(z,y)\bigg|_{y=z=\hat{z}} = 0 \text{ (ESS neutral)} \tag{4.11a}$$

$$\left[\frac{\partial^2}{\partial z^2}\theta - \frac{\partial^2}{\partial y^2}\theta\right]\bigg|_{y=z=\hat{z}} = \epsilon(1-g(\hat{z})) + (1+\epsilon\hat{z})g'(\hat{z}) > 0 \text{ (CS)}. \tag{4.11b}$$

Condition (4.11a) ensures that under the probabilistic interpretation of strategy, branching cannot occur. If $g(z)$ is a constant then condition (4.11b) shows that any interior solution must repel all nearby strategies (recall $\epsilon < 0$ in the case 4 of Table 2.1). Indeed, we show this to be the case when $p_1 = 0$ in section 2.3.1. Strategies diverge to the boundaries of 0 and 1. While $g(z)$ is monotonic increasing, it is possible for a singular strategy to be attracting (convergence stable). Conditions (4.11a)-(4.11b) characterise a singular strategy $\hat{z}$, and thereby the behaviours of group members, in one of two ways; if $\hat{z}$ is not convergence stable then all group members will invest 1 with a probability equal to either zero or one, alternatively if $\hat{z}$ is convergence stable then all group members will invest 1 with a probability defined by $\hat{z}$.

Contrastingly, under the contribution investment the invasion fitness of novel strategy y in a monomorphic resident population adopting z is defined by equation (4.7). In this case, the selection gradient is given by equation (4.8a). In a monomorphic population we use $\partial_y\lambda(z,y) = \partial_y G(y;z;1)$ which produces non-trivial ESS and CS conditions.

In general, the adaptive dynamics of two distinct scalar strategies y and z with respective invasion fitnesses $S(z,y)$ and $S(y,z)$ can be characterised by four possible different outcomes: (i) $S(y,z) < 0$ and $S(z,y) < 0$, (ii) $S(y,z) > 0$ and $S(z,y) < 0$, (iii) $S(y,z) < 0$ and $S(z,y) > 0$ and (iv) $S(y,z) > 0$ and $S(z,y) > 0$. Cases (i)-(iv) respectively correspond to bi-stability, z dominating, y dominating, and co-existence. Condition (4.11a) ensures that only cases (i) and (ii) can be observed under the contribution probability. Whereas, all four possible scenarios are possible under the contribution investment.

4.3.2 A comparison of fitness

Differences in the selection gradients (4.10) and (4.8a) are attributable to how we view a strategy to influence fitness. The invasion fitness of a probabilistic strategy, equation (4.9), treats a group member's experience of risk dilution as a sum of two constants. From equations (2.3) in chapter 2, we see that with a probability y the group member experiences a risk dilution of $\mathcal{D}_C$ and with probability $1-y$ it experiences $\mathcal{D}_D$ (where the extensive variable j is replaced by the intensive variable z). As such, as an individual alters its strategy y (here defined as the probability of investing 1), the changes in the dilution of risk it experiences are constant:

$$\mathcal{D}_C - \mathcal{D}_D = -\frac{\epsilon}{n(1+\epsilon z)} \tag{4.12}$$

regardless of the individuals initial probability to invest 1.

The invasion fitness of the contribution investment strategy, equation (4.7) treats an individual's experience of risk dilution as a continuous function of its strategic investment. As a focal group member alters its investment level, the change in dilution risk it experiences is characterised by

$$\frac{\partial}{\partial y}\mathcal{D}(y,z,1) = -\frac{\epsilon'(y)}{n(1+\epsilon(z))}. \tag{4.13}$$

Naturally, it is of interest to compare the changes in risk dilution an individual experiences with changes in strategy, for each interpretation of strategy. Only when $\epsilon(y)$ is linear in y does equality between equations (4.12) and (4.13) hold. Similarly, $\gamma'(y)$ and $\delta = \gamma_C - \gamma_D$ are only equal when $\gamma(y)$ is linear in y. Even when the functions of $\gamma(z)$ and $\epsilon(z)$ do reflect the constants used in f_C and f_D, the resulting selection gradients differ (see equation C.2 of appendix C.1). When nonlinear terms are used in equation (4.4a) these differences become more pronounced.

4.4 Results

As explained in section 4.3.2, the choice of functional forms of $\gamma(z)$ and $\epsilon(z)$ distinguish the continuous investment strategy dynamics from the (discrete) mixed strategy dynamics of the contribution probability. In this section we analyse how the choice of such functions affect the dynamics of the continuous investment strategy, and provide explicit examples of the adaptive dynamics of the contribution investment. For each choice of functions, we compare the results to the contribution probability interpretation of strategy as presented in chapter 2. The results presented here exhibit all four qualitatively distinct outcomes of the adaptive dynamics. Each different outcome encapsulates a different structure of the trade-off between foraging and risk dilution.

We start with the simplest of assumptions in which $\epsilon(z)$ and $\gamma(z)$ are linear, and the probability of group evasion is a constant. This case, analysed in section 4.4.1 will be used as a reference. We then respectively relax each assumption in turn. The functional choices in sections 4.4.1-4.4.3 all consider a constant probability of group evasion. We know that the replicator dynamics of the probability investment yield no stable interior solution (see equation 2.13 of section section 2.3.1, chapter 2). In such cases, the functional choices of $\epsilon(z)$ and $\gamma(z)$ specify the adaptive dynamics of the contribution investment. Analysis of models 4.4.1-4.4.3 show that combinations $\gamma(z)$ and $\epsilon(z)$ with different properties are able to produce 3 qualitatively distinct outcomes. We see that singular strategies may be neither ESS nor CS, ESS but not CS, and both ESS and CS.

In sections 4.4.4-4.4.5 we consider cases where the probability of group evasion depends on the strategic choices of group members. In this case, the probability of group evasion reflects a production function of a public good. The probability that the predator's attack fails depends on each group member's investment into anti-predator defences, but benefits the whole group equally. We know from section 2.3.1, chapter 2 that a monotonic increasing, concave choice of $g(x)$ can result in a pure strategy polymorphism (i.e., a frequency x_4 invest 1; and $1 - x_4$ invest 0). This finding is equivalent to a mixed strategy monomorphism, and can be interpreted as a convergence stable intermediate probability of adopting the maximum investment strategy. All individuals display the

same probability (less than 1; greater than 0) to invest 1. The results presented in sections 4.4.4-4.4.5 show how the adaptive dynamics of a continuous investment strategy relate to the findings presented in chapter 2.

4.4.1 Case 1: All functional choices are linear

In this case, individual benefits from grouping decrease linearly and individual benefits in risk dilution increase linearly with investment. We note here that in the case of linear $\gamma(z)$ and $\epsilon(z)$ at the boundaries of strategy space, $\theta(z,y)|_{y=0} = \lambda(z,y)|_{y=0}$ and $\theta(z,y)|_{y=1} = \lambda(z,y)|_{y=1}$. Moreover, because $\theta(z,y)$ is linear in y and $\lambda(z,y)$ is strictly concave for $y \in (0,1)$ we see that the invasion fitness of a novel contribution investment strategy is always greater than the invasion fitness of a novel contribution probability strategy.

Application: linear choices of $\gamma(z)$ and $\epsilon(z)$ and constant $g(y;\mathbf{z};\mathbf{x}) = p_0$

Consider that an exploration event results in a novel strategy y appearing in the population playing z_1. A condition is imposed on the novel strategy's region of feasibility through the requirement that its dilution experience (probability that the individual is targeted) remains strictly non-negative. The requirement is met by the imposition of a minimum group size; $\bar{n} = (1-\epsilon y)/(1-\epsilon z_1)$ which groups must be greater or equal to. In practise, $\bar{n} \approx 1$ since we only consider novel y strategies deviating slightly from z_1, and $\lim_{y\to z_1} \bar{n} = 1$. Using equation (C.5) (see appendix C.2) the invasion fitness of y is positive if

$$n(1+\epsilon z_1) - (1-p_0)\left\{1 + \frac{\epsilon\gamma_D}{\delta} + \epsilon(y+z_1)\right\} < 0. \tag{4.14}$$

Expression (4.14) defines a line such that

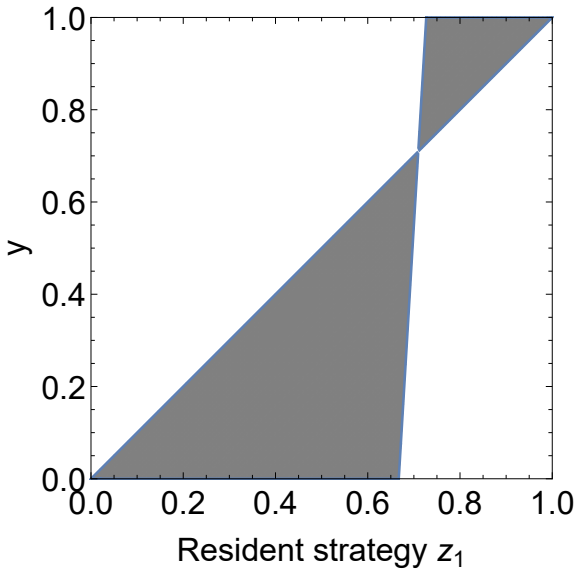
$$\mathrm{L} : y = \frac{n-(1-p_0)}{(1-p_0)} z_1 - \frac{(1-p_0)(1+\frac{\epsilon\gamma_D}{\delta}) - n}{\epsilon(1-p_0)}$$

Singular strategies of the adaptive dynamics are found by solving the intersection between L and the principle diagonal $z_1 = y$. We find one singular strategy given by

$$\hat{z}_x = \frac{\epsilon\gamma_D(1-p_0) + \delta(1-n-p_0)}{\epsilon\delta(n-2+2p_0)}. \tag{4.15}$$

Regimes for which $\hat{z}_x \in (0,1)$ are outlined in appendix C.2.1. The conditions of evolutionary (C.6) and convergence (C.7) stability simplify due to the zero second order derivatives of $\gamma(z)$ and $\epsilon(z)$. From condition (C.6) we see that $\hat{z}_x \in (0,1)$ is an ESS while $-\epsilon\delta/(\gamma_D + \delta\hat{z}_x) < 0$. Using (C.7) we see that $\hat{z}_x$ is CS when $\delta/\gamma_D > \epsilon$. From appendix C.2.1 one sees that the CS condition ensures $\hat{z}_x \notin (0,1)$ for $n \geq 2$. All physical singular strategies are evolutionary stable repellers.

Invasion simulations governed by the differential equations (4.8), shown in Fig 4.2 illustrate the dynamics of two arbitrary strategies, z_1 and z_2, initially co-existing at an approximately monomorphic population. Note that our choice of z_1, z_2 is purely to illustrate the use of the G-function in describing the strategy dynamics: the invasion simulation would produce the same result if we only



(A)

FIGURE 4.1: a): The singular strategy $\hat{z}_x$ is ESS but not CS. Grey regions indicate strategy combinations for which the invasion fitness $\lambda(z_1, y) > 0$, i.e., the y strategy can invade. White regions show strategy combinations for which the invader cannot invade (all subsequent PIPs follow this colour scheme). The intersection of the principal diagonal and the zero contour of $\lambda(z_1, y) > 0$ gives $\hat{z}_x$. White regions directly above and below $\hat{z}_x$ indicate that the strategy is an ESS. White regions above $\hat{z}_x$ on the right, and below $\hat{z}_x$ on the left indicate the solution is not CS. Parameter values: $\gamma_C = 0.82, \gamma_D = 0.97, \epsilon = -0.95, p_0 = 0.1, n = 16$

consider one strategy z_1 in the resident population. While both strategies z_1 and z_2 lie below the repeller $\hat{z}_x$ (blue dashed) then both converge to 0 in concert (A) of Fig 4.2. If both strategies lie above $\hat{z}_x$, then both converge to 1 in concert (not plotted). The singular strategy of the contribution probability (i.e., x_3 of section 2.3.1) is shown by the grey dashed line. Note that x_3 is a probability. We see that x_3 is not convergence stable, and is evolutionary neutral. Both singular strategies $\hat{z}_x$ and x_3 act to separate the basins of attraction of the respective strategies of 0 and 1. The adaptive landscape changes with changes in the investment strategy; (B) of Fig 4.2. In each panel of (B) the invasion fitness of a novel strategy y is plotted against the resident population at an ecological equilibrium. The strategies move in the direction of positive fitness, i.e., "uphill", quite similar to the landscape envisaged by Wright [13], albeit here on a behavioural timescale. While $z_1, z_2 < \hat{z}_x$ co-existing in the (approximately) monomorphic population, both strategies move in the direction proportional to increases in fitness, which results in them converging in tandem to 0 (see panel "t=15").

Viewing strategy as the probability to invest 1, if all individuals initially exhibit a probability less than x_3, then the outcome is for all to invest 1 with a probability 0. This outcome is equivalent to the outcome of the contribution investment where the initial strategy in the resident population is less than $\hat{z}_x$. It is possible for the outcomes of the contribution probability and the contribution investment to differ. We see differences in the strategy dynamics when initial strategies in the resident population are contained within the region of phase space defined by the differences in $\hat{z}_x$ and x_3; (A) of Fig 4.2 and Fig 4.3. For example consider a monomorphic resident population with strategy z_1, contained in this region (between the blue and grey dashed lines). While z_1 is viewed

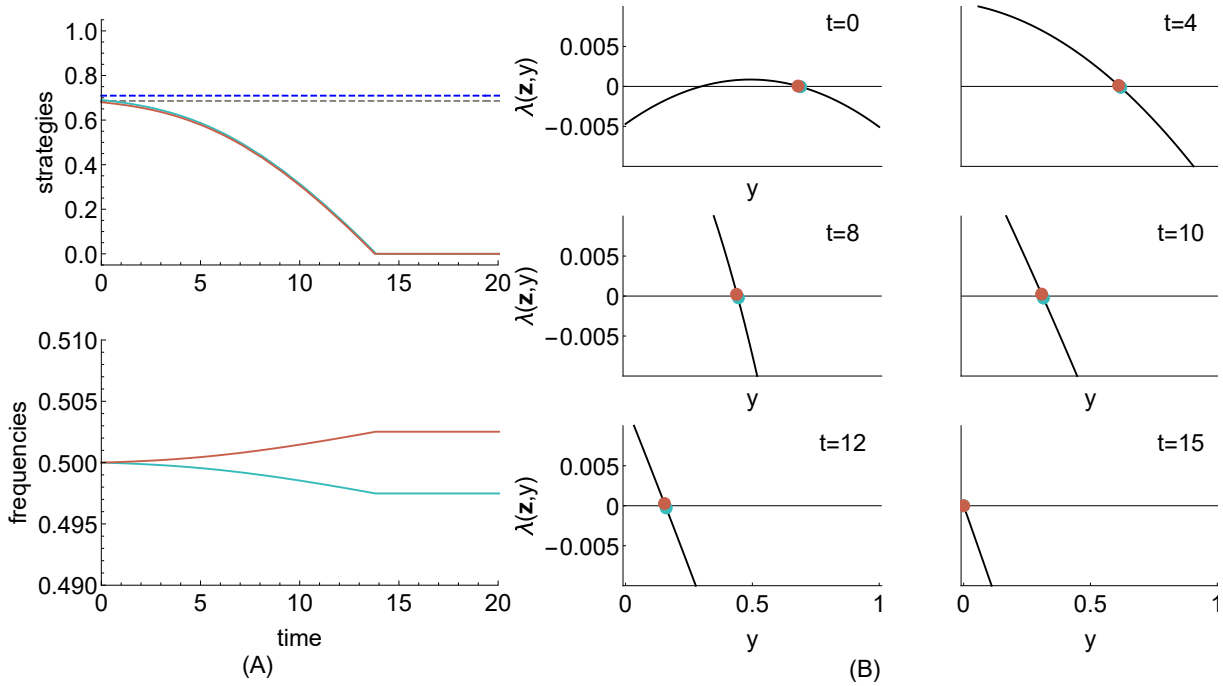


FIGURE 4.2: An invasion simulation. The interior singular strategy of the investment, $\hat{z}_x$, is shown by the blue dashed line. The grey dashed line indicates the interior singular strategy of the contribution probability interpretation. (A): Starting with two strategies z_1 (blue) and z_2 (maroon) with $z_1, z_2 < \hat{z}_x$ (we use $z_2 = \hat{z}_x - 0.02, z_1 = \hat{z}_x - 0.01$), both strategies converge to the zero investment strategy in tandem. (B): The adaptive landscape. Blue dots correspond to strategy z_1 and the maroon to z_2. Each panel shows the adaptive landscape at time t. Parameter values same as Fig 4.1. In all subsequent invasion simulation plots, stabilities of singular strategies are illustrated as follows: solid (dashed) for CS (for not CS), an ESS is shown blue, not an ESS red, and ESS neutral grey.

as a mixed (probability) strategy then the outcome is for all group members to invest 1 with a probability 1. Whereas if z_1 is a continuous investment strategy then all group members invest 0. The difference in $\hat{z}_x$ and x_3 governs the differences in these behavioural outcomes, and can be explained by comparing equations (C.1) and (C.2) of appendix C.1.

The derivatives of $\epsilon(z)$ and $\gamma(z)$ condition the outcome of the investment strategy. If

$$\epsilon'(z) = \epsilon < -\frac{\delta(1-n-p_0)}{\gamma_D(1-p_0)} \tag{4.16}$$

then $\hat{z}_x < 0$ in which case all strategies converge to 1, whereas if

$$\gamma'(z) = \delta < \frac{\epsilon\gamma_D(1-p_0)}{n-1+p_0+\epsilon(n-2+2p_0)} \tag{4.17}$$

then $\hat{z}_x > 1$ and all strategies converge to 0. The singular strategy $\hat{z}_x$ passes through the boundaries of phase space via transcritical bifurcations. While physical, the partial derivatives of $\hat{z}_x$ indicate how the investment levels of group members change with model parameters. We find $\partial_\epsilon \hat{z}_x > 0$ meaning that the basin of attraction of 1 increases with $|\epsilon|$; Fig 4.3c. We see this because for any given investment level, an increase in the slope $\epsilon = \epsilon'(y)$ increases the probability that any focal individual is not targeted by the predator. Investing 1 becomes relatively more beneficial.

Similarly, we find that the basin of attraction of 0 increases with increasing $|\delta|$; Fig 4.3b. As $|\delta|$ is increased, foraging benefits decrease at a greater rate, with increases in strategy investment. Investing 0 becomes relatively more beneficial. We find that increases in group size increase the basin of attraction of the 0 investment level ($\partial_n \hat{z}_x > 0$); Fig 4.3a. Similar negative group size effects are found for example in models of vigilance [32]. Lastly, we find that as p_0 increases there is an increase in the basin of attraction of 0; Fig 4.3d. An increase in p_0 results in all group members experiencing a reduction in predation risk. In this case, group members are able to decrease strategic investment while maintaining the same relative risk profile.

These results found in this example reflect a case in which it is better for all group members to invest all or nothing into anti-predator activities. These findings mirror the results of chapter 2 where individuals are viewed to play either strategy C (invest 1) or D (invest 0) with probability 1, depending on initial strategy. When there is no singular solution in $(0, 1)$, the contribution probability and the contribution investment mechanisms give rise to precisely the same dynamical outcomes. In each interpretation of strategy we find the same extremal strategy monomorphic population. As such, all group members invest 0 or 1 regardless as to whether it be by them probabilistically choosing an extremal investment level, or flexible altering their investment level. While there does exist an interior singular strategy it's quite possible for the outcomes of the respective dynamics to differ; Fig 4.3. Deviations between $\hat{z}_x$ and x_3 quantify the differences. We expect such differences to only occur in a small region of parameter space, due to the similarities between the singular strategies (compare $\hat{z}_x$ defined in equation (4.15) to the fixed point x_3 defined in equation (2.13), section 2.3.1 of chapter 2).

4.4.2 Case 2: Changes in foraging benefits are nonlinear

Here, we assume that the benefits of not being an explicit predator target increase linearly with increasing strategic investment, and that foraging benefits are some nonlinear function of investment. For example, foraging rewards may be convex decreasing with investment if intake rate is associated with a Holling type II functional response with foraging intensity [191]. By increasing effort into activities such as watching for predators (decreasing foraging rate), one would expect saturating decreases in foraging benefits with respect to increases in anti-predator strategy investment.

In this case we see that $\lambda(z, y)$ is affected by the properties of $\gamma(z)$. Using conditions (C.6) and (C.7) we see that a singular strategy $\hat{z}_x$ is an ESS if

$$\gamma''(z) < \frac{2\epsilon\gamma'(z)(1-p_0)}{(n-1+p_0)(1+\epsilon z)} \tag{4.18}$$

and is CS if

$$\gamma''(z) < \frac{(1-p_0)\epsilon\big(\gamma(z)\epsilon - (1-\epsilon z)\gamma'(z)\big)}{(1-\epsilon z)^2(p_0+n-1)} \tag{4.19}$$

when $z = \hat{z}_x$. The RHS of condition (4.18) is always positive and so it is possible for both convex and concave (and sigmoid) functional choices of $\gamma(z)$ to give rise to evolutionary stable strategies. The RHS of condition (4.19) is positive when

$$\frac{\epsilon}{1+\epsilon z} < \frac{\gamma'(z)}{\gamma(z)} \tag{4.20}$$

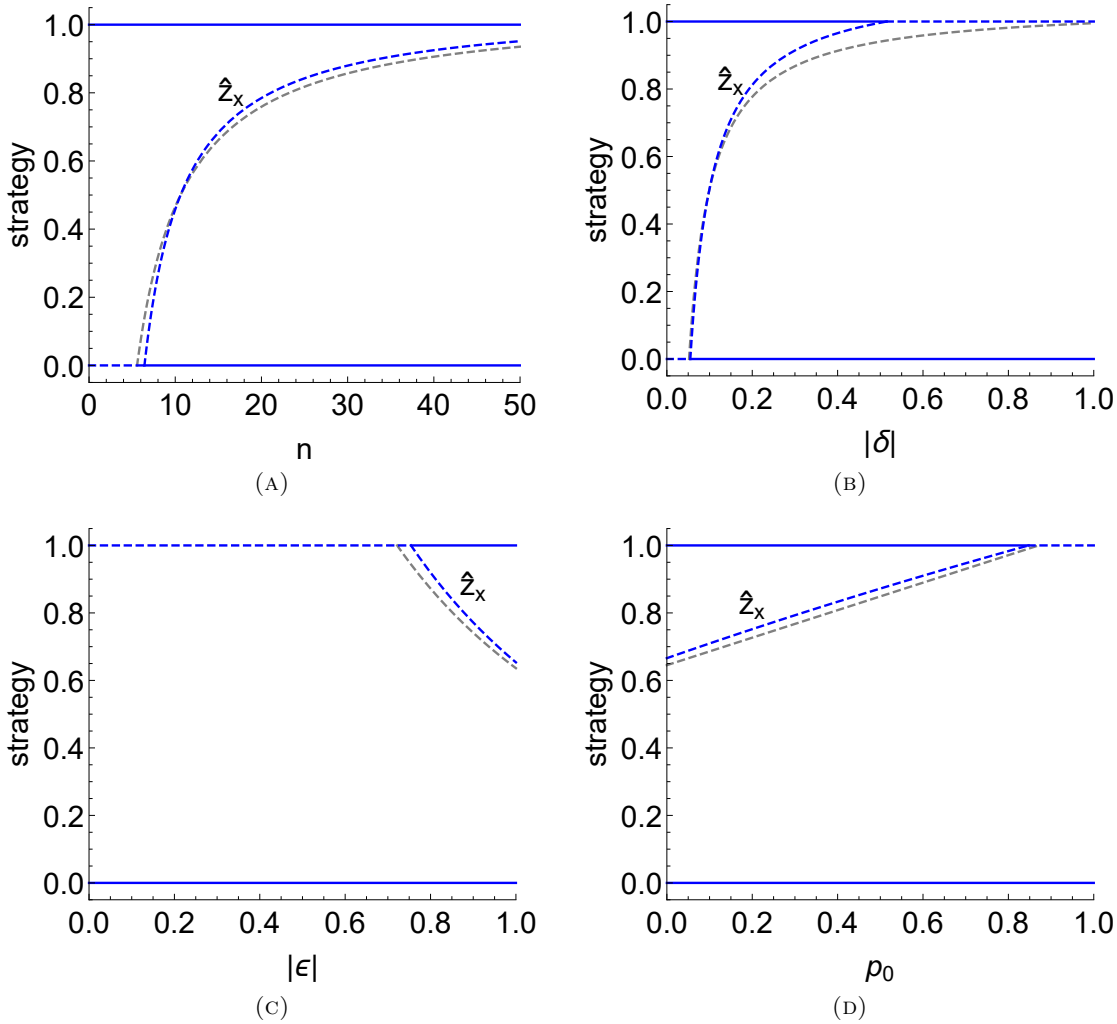


FIGURE 4.3: In plots solid (dashed) lines indicate convergence stable (unstable), and blue (red) lines indicate evolutionary stable (unstable) singular strategies. Grey lines indicate evolutionary neutral strategies. All subsequent phase portraits follow this scheme. In each plot we compare the ESS-stable $\hat{z}_x$ (blue) and ESS-neutral x_3 (grey). (A): The solution $\hat{z}_x$ increases with group size which increases basin of attraction of the zero investment strategy. (B): As $|\delta|$ increases the basin of attraction of zero investment strategy increases. (C): The basin of attraction for the maximum investment strategy increases with increasing $|\epsilon|$. (D): The solution $\hat{z}_x$ decreases with p_0 resulting in a increase in the basin of attraction of the minimum investment strategy. General parameter values same as in Fig 4.1, with respective free variables removed in each plot.

and negative otherwise. We see that if $\gamma(y)$ is convex on $y \in (0,1)$ at the singular strategy value, and condition (4.20) does not hold, then the singular strategy cannot be attracting. Conversely, while $\gamma(y)$ is concave at the singular strategy value then the singular strategy is CS when (4.20) holds. Conditions (4.18)-(4.19) ensure that branching cannot occur (the RHS of 4.19 is strictly less than the RHS of 4.18).

Application: $\gamma(z) = \gamma_D - \delta(k+1)z/(k+z)$, linear choice of $\epsilon(z)$ and constant $g(y;\mathbf{z};\mathbf{x}) = p_0$

While the resident population is monomorphic, a novel strategy y is able to invade if

$$n(1+\epsilon z_1) - (1-p_0)\left\{1+\epsilon\left(\frac{\gamma_D(k+z_1)(k+y)+\delta(k+1)\big(k(z_1+y)+z_1 y\big)}{\delta(k+1)k}\right)\right\} \lesseqgtr 0 \text{ for } y \lesseqgtr z_1\ . \tag{4.21}$$

Setting $y = z_1$, equation (4.21) defines a conic,

$$\begin{aligned} \mathrm{C}_1 :&\big(\epsilon(\gamma_D+\delta(k+1))(p_0-1)\big)z_1^2 - \big(\epsilon k(2(\gamma_D+\delta(k+1))(1-p_0)-\delta(k+1)n)\big)z_1 \\ &- \big((1-p_0)k(\delta(k+1)-\epsilon\gamma_D k)+\delta(k+1)kn\big) \end{aligned} \tag{4.22a}$$

There can be at most one root of (4.21) in the interval $(0,1)$. To see this we apply Budan's theorem. Let $\nu_i(z_1)$ be the number of sign changes in the sequence of coefficients of $\mathrm{C}_1(z_1+i)$. The quadratic coefficients of $\mathrm{C}_1(z_1+0)$ and $\mathrm{C}_1(z_1+1)$ are precisely the same, and linear coefficients are both strictly positive. The sequence of coefficients of $\mathrm{C}_1(z_1+1)$ and $\mathrm{C}_1(z_1)$ share at least one out of two potential sign changes. Let $R_{[0,1)}$ be the number of roots of $\mathrm{C}_1(z_1)$ in $[0,1)$. The quantity $\nu_0(z_1) - \nu_1(z_1) - R_{[0,1)}$ must be a non-negative even integer. From this we see that there is at most one root of C_1 in the unit interval. We define this singular strategy as $\hat{z}_x^+$. Full details of the singular strategies can be found in appendix C.2.

From condition (4.18) the singular strategy is an ESS if

$$-\frac{\epsilon(k-kp_0)+n+p_0-1}{\epsilon n} < \hat{z}_x^+ < 1. \tag{4.23}$$

Due to our choice of $\gamma(z)$ we know that $\hat{z}_x^+$ cannot be CS (see condition 4.20). There are therefore only two possible stability outcomes for the convergence unstable singular strategy. Either $\hat{z}_x^+$ is an ESS; Fig 4.4a. Alternatively, $\hat{z}_x^+$ is not an ESS; Fig 4.4b. While $\hat{z}_x^+$ is not an ESS it can be invaded on either side by nearby strategies y. In either of the just mentioned cases of evolutionary stability, $\hat{z}_x^+$ is not convergence stable, and all local strategies diverge to either of 0 or 1, depending on whether the strategy values are greater or less than $\hat{z}_x^+$ i.e., which basin of attraction the strategies lie in.

For the convergence unstable and evolutionary stable singular strategy depicted in Fig 4.4a, an invasion simulation of a novel strategy entering the population is shown by Fig 4.5. Initially, two virtually indistinguishable arbitrary strategies $z_1 \approx z_2$ co-exist within the (approximately monomorphic) population. As $\hat{z}^+$ is not CS it acts as a repeller. If $z_1, z_2 > \hat{z}_x^+$ then both strategies converge to 1 (not plotted) whereas if $z_1, z_2 < \hat{z}_x^+$ then both strategies converge to 0; (A) of Fig 4.5. Because the two strategies never deviate far from each other, the population remains approximately

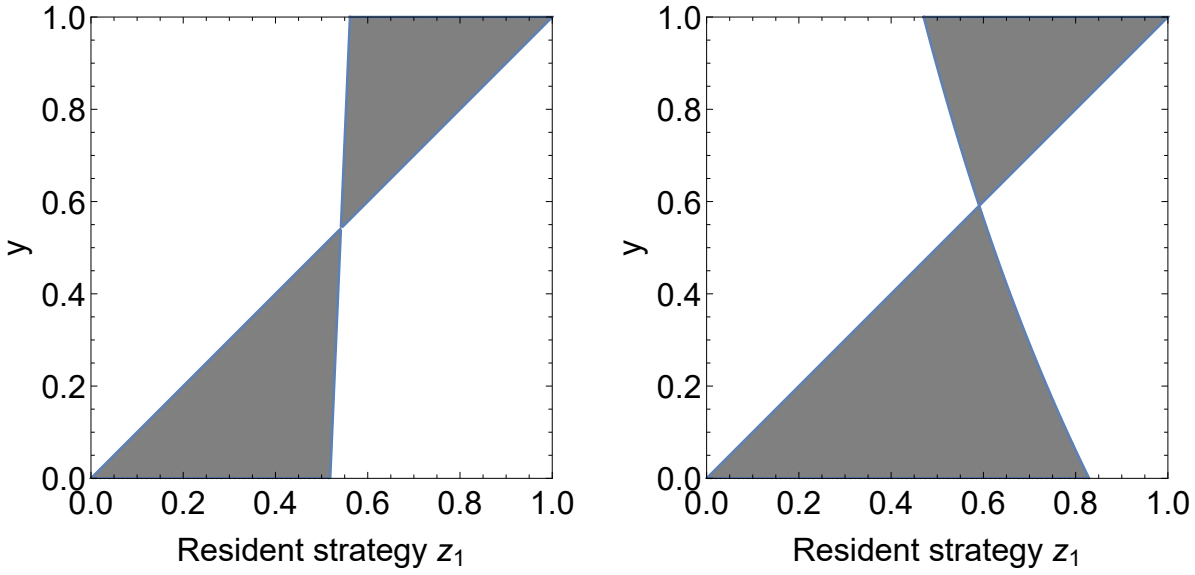


FIGURE 4.4: PIPs. Singular strategies shown in intersections of the principle diagonal and zero contours. (A): Example where the singular strategy is an ESS but not CS. Directly above and below $\hat{z}^+$, white regions indicate that $y > \hat{z}_x^+$ and $y < \hat{z}_x^+$ cannot invade $\hat{z}_x^+$, i.e, the solution is ESS . (B): The strategy $\hat{z}_x^+$ is neither ESS nor CS. Grey regions above and below the singular strategy indicate that it is vulnerable to invasion from above and below. Parameter values: $\gamma_C = 3.2\gamma_D = 4, k = 8, \epsilon = -0.99, p_0 = 0.1, n = 10$ for (A) and $\gamma_C = 4.3, \gamma_D = 4.6, k = 0.5, \epsilon = -0.85, p_0 = 0.45, n = 20$ in (B).

monomorphic. Both z_1 and z_2 maintain a positive frequency. We see that each strategy climbs the adaptive landscape moving in the direction of the positive invasion fitness, $\lambda(\mathbf{z}, y)$; (B) of Fig 4.5.

In the alternate case where $\hat{z}_x^+$ is neither CS nor an ESS (depicted in Fig 4.4b), two strategies $z_1 \approx z_2$ co-existing in an approximately monomorphic population simultaneously converge to either 1 if $z_1, z_2 > \hat{z}_x^+$, or to 0 if $z_1, z_2 < \hat{z}_x^+$ (neither plotted). If the resident population is dimorphic prior to invasion, with $\hat{x}_1 \in (0, 1)$ frequency of z_1 and $1 - \hat{x}_1$ frequency of z_2 such that $z_2 < \hat{z}_x^+ < z_1$, then the strategies z_1 and z_2 respectively diverge to 1 and 0; (A) in Fig 4.6. We see that z_1 becomes dominant and the frequency of strategy z_2 goes to zero. For this outcome, the resident strategies must be sufficiently distinct so that they cannot co-exist at the hyperbolically stable ecological equilibrium of $\hat{x}_1 = 1$ (if this latter holds then both residents converge to the same strategy). The changing adaptive landscape shows a fitness valley develops between the strategies; (B) of Fig 4.6. As such, the strategies undergo bi-directional change with $\dot{z}_2 < 0$ and $\dot{z}_1 > 0$. At each exploration event, the strategies move in directions proportional to increases in fitness, however, we see that because the adaptive landscape is constantly changing, even though z_2 moves in the direction of increasing fitness (i.e., uphill), it ends at a fitness minima; panel (t=5). Hence, the frequency of z_2 goes to zero.

The results found in this example clearly indicate a selective pressure for group members to prioritise either of maximum strategic investment, or no investment. These findings are in agreement with the results of chapter 2 wherein individuals choose the maximum investment strategy (i.e., play C) with a probability zero or one. The payoff structure of the game, defined by the explicit

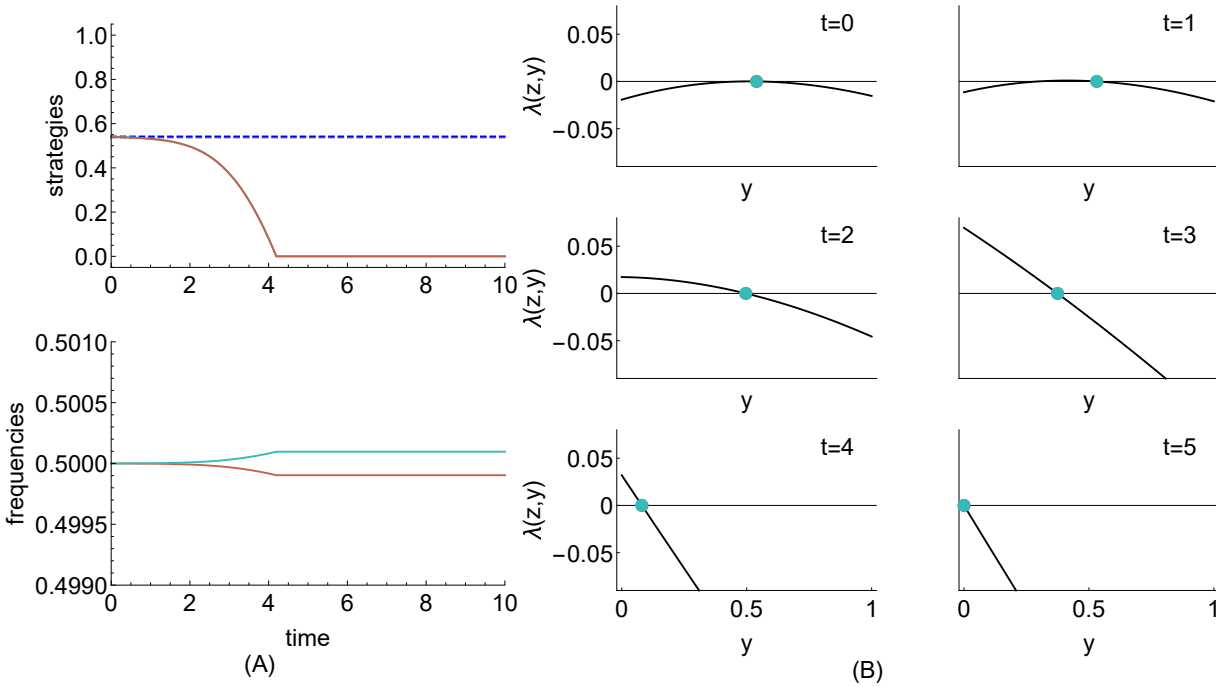


FIGURE 4.5: Invasion simulations. Here $\hat{z}_x^+$ (blue dashed line) is an ESS but not CS and $x_3^{(p_1=0)}$ (grey dashed) is not CS. (A): Starting with a resident population composed of two co-existing very similar strategies $z_1 = \hat{z}_x^+ - 0.01$ and $z_2 = \hat{z}_x^+ - 0.02$, both strategies converge to 0. (B): The adaptive landscape. Each strategy changes in a direction defined by the positive invasion fitness, λ. Parameter values: $\gamma_C = 3.2\gamma_D = 4, k = 8, \epsilon = -0.99, p_0 = 0.1, n = 10$.

choice of $\gamma''(z) > 0$, reflects a situation in which the greatest payoff is achieved by investing either all or nothing into anti-predator activities. If an individual invests zero, it maximises its foraging rewards; if it invests one then it maximises its relative probability of not being targeted. Group members are unable to balance the trade-off between foraging and predation by a compromising behaviour in between investing all or nothing into anti-predator activities.

4.4.3 Case 3: Changes in relative individual risk dilution are nonlinear

Here, we assume that the foraging benefits which group members receive decrease linearly with strategy investment, while the $\epsilon(z)$ functions (describing the probability that a focal with strategy z is a predator target) are nonlinear.

The stability of a candidate singular strategy depends on the form of $\epsilon(z)$. Using conditions (C.6) and (C.7) we see that a singular strategy $\hat{z}_x$ is evolutionary stable if

$$\epsilon''(z) > -\frac{2\delta\epsilon'(z)}{\gamma_D + \delta z} \tag{4.24}$$

and is convergence stable if

$$\epsilon''(z) > \epsilon'(z)\left(\frac{\delta}{\delta z + \gamma_D} + \frac{\epsilon'(z)}{1 + \epsilon(z)}\right) \tag{4.25}$$

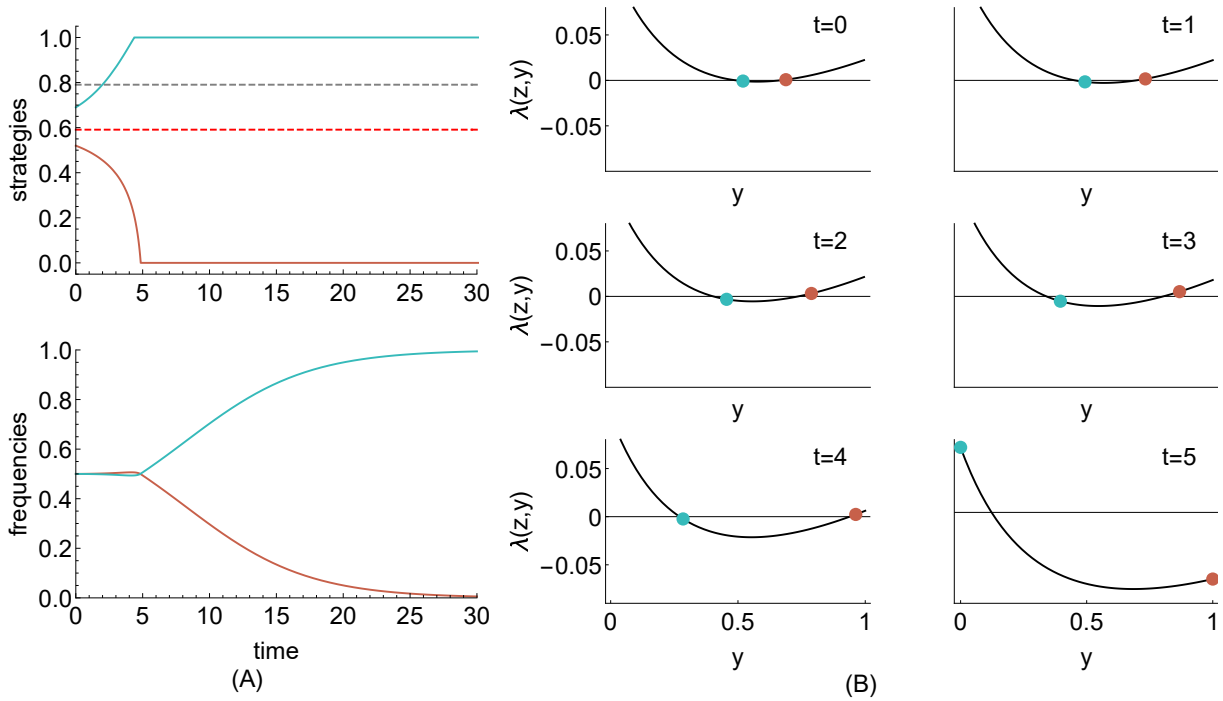


FIGURE 4.6: Invasion simulations. Here $\hat{z}_x^+$ (red dashed line), is neither an ESS nor CS and $x_3^{(p_1=0)}$ (grey dashed) is not CS. (A): Starting with an initially dimorphic population with $z_1 > \hat{z}_x^+ > z_2$, the strategies diverge to 1 and 0. (B): The adaptive landscape. We see that the strategies of z_1 and z_2 move in direction of increasing fitness. This leads z_1 to 1, and z_2 to 0. Eventually, $(t = 5)$, strategy z_2 lies at a local fitness minima, and its frequency tends to zero (as shown in (A)). Parameter values: $\gamma_C = 4.3, \gamma_D = 4.6, k = 0.5, \epsilon = -0.85, p_0 = 0.45, n = 20$.

when $z = \hat{z}$. From condition (4.24) one sees that if $\epsilon(z)$ is convex on $[0,1]$ at $\hat{z}$, then the singular strategy will be an ESS. By (4.25) we see that while $\epsilon(z)$ is concave at $\hat{z}$, the singular strategy cannot be CS. We see that branching cannot occur because convergence stability requires that $\epsilon(z)$ be convex, which also guarantees un-invasibility (ESS).

Application: $\epsilon(z) = \epsilon(k+1)z/(k+z)$, linear choice of $\gamma(z)$ and constant $g(y;\mathbf{z};\mathbf{x}) = p_0$

While the resident population is at an ecologically stable equilibrium $x_1 = 1$, a novel strategy y can invade if

$$1 - \frac{1-p_0}{n} - \frac{(1-p_0)\big(\epsilon k(1+k)(\gamma_D + \delta y)\big)}{\delta n(k+y)(k+z_1+\epsilon(1+k)z_1)} \lesseqgtr 0 \text{ for } y \lesseqgtr z_1 \ . \tag{4.26}$$

Setting the novel and resident strategies equal, equation (4.26) defines a conic C_2:

$$\begin{aligned}\mathrm{C}_2 :&\big(\delta(1+\epsilon+\epsilon k)n(n-1+p_0)\big)z_1^2 + \big(\delta k n\big(2(n-1+p_0)+\epsilon(1+k)(n-2+2p_0)\big)\big)z_1 + \\ &\big(kn\big(\gamma_D\epsilon(1+k)(p_0-1)+\delta k(n-1+p_0)\big)\big)\end{aligned} \tag{4.27a}$$

In the appendix we show that there can exist at most one singular strategy in the open unit interval at any one time. In certain parameter regimes we find a singular strategy that is CS and an ESS, $\hat{z}_x^+$ (physicality depends on two values of group size given in appendix C.2.3 due to space constraints). By inspecting condition (4.24) we know that $\hat{z}_x^+$ is an ESS, due to the property of $\epsilon''(z) > 0$. From condition (4.25) we see that $\hat{z}_x^+$ is CS if

$$0 < \hat{z}_x^+ < \min\left\{ -\frac{\gamma_D}{\delta} - \sqrt{\frac{(\delta k+\gamma_D)(\delta k - \epsilon\gamma_D + \gamma_D)}{\delta^2(1-\epsilon)}}, 1 \right\}. \tag{4.28}$$

depending on the values of ϵ and k. Direct calculations show that $\hat{z}_x^+$ is CS whenever it exists in $(0,1)$. In other parameter regimes we find an evolutionary stable, but not convergence stable singular strategy $\hat{z}_x^-$.

Fig 4.7 illustrates how an invasion event might occur through a PIP. The grey regions above left and below right $\hat{z}_x^+$ indicate that it is CS; the white regions directly above and below indicate that it is an ESS. A monomorphic resident population with strategy $z_1 \neq \hat{z}_x^+$ can be invaded by y while $y > z_1$ and $z_1 < \hat{z}_x^+$ or $y < z_1$ for $z_1 > \hat{z}_x^+$; Fig 4.7a. The Mutual invasibility plot (MIP) is a superposition of the PIP and its reflection in the principal diagonal; Fig 4.7b. To the right of $\hat{z}_x^+$ there is a "+" below the diagonal and a "-" above: the local fitness gradient from above points towards $\hat{z}_x^+$. Left of the singular strategy there is a "+" above the diagonal and a "-" below: the local fitness gradient from below also points towards $\hat{z}_x^+$. Two transcritical bifurcation curves which are drawn as dashed curves. The maroon dashed curve shows strategy values for a zero valued invasion fitness of a invader strategy with zero frequency, and the blue dashed curve shows the same for a zero sized frequency resident strategy. The two curves intersect at $\hat{z}_x^+$. Plotted also are the nullclines (zero isoclines) of the selection gradient evaluated for the resident strategy (blue solid) and the invader strategy (maroon solid). The arrows indicate the direction of small strategy changes. Notice that at $\hat{z}_x^+$, either side of the principle diagonal are regions of unstable co-existence. This prevents branching occurring as neither strategy can maintain a stable frequency when not

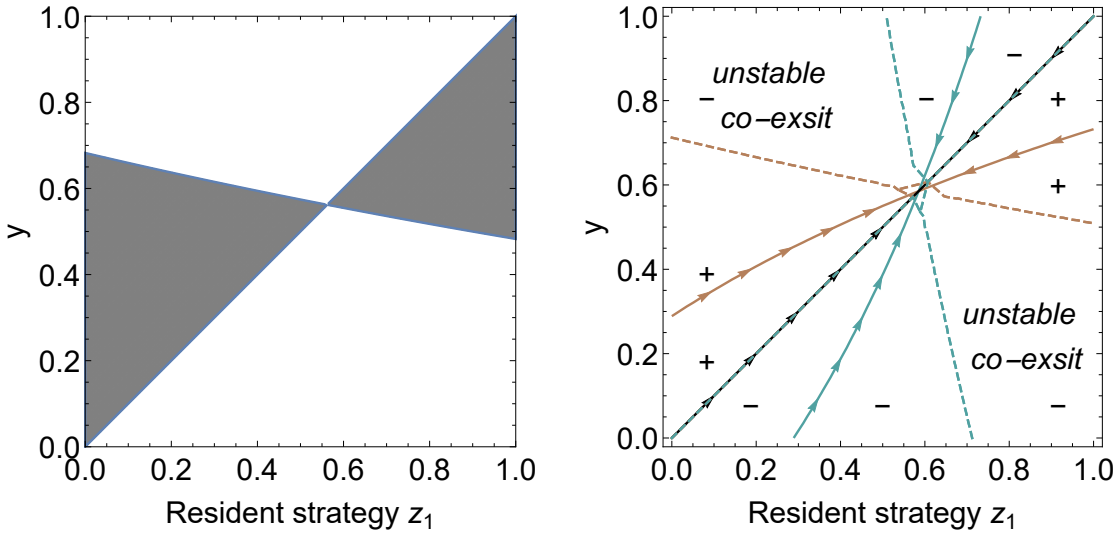


FIGURE 4.7: (A): PIP. The singular strategy $\hat{z}_x^+$ is CS and an ESS. (B): MIP. There is no mutual invasibility. The "+" show regions where the invader strategy can invade and replace the resident, and the "-" describes regions of strategy space for which the resident cannot be invaded. The dashed curves show the transcritical bifurcations, and solid lines the nullclines of the selection gradient (maroon for resident and blue for invader). Arrows show the direction of strategy changes. Parameter values: $\gamma_C = 6.51, \gamma_D = 6.82, \epsilon = -0.54, k = 0.62, p_0 = 0.1, n = 12$

equal to the singular strategy. Once the singular strategy is reached, no further changes in strategy occur.

When present in $(0, 1)$, the singular strategy defines the investment level of all group members. Starting with two similar strategies z_1, z_2, we see that both converge in tandem to $\hat{z}_x^+$. As the strategies do not deviate far from each other, the population remains approximately monomorphic and the strategies can co-exist; (A) of Fig 4.8. The dynamic outcome for the investment strategy is considerably different from the contribution probability strategy. The singular strategy of the contribution probability x_3 (grey dashed) acts to separate the basins of attraction of the CS probabilities of 0 and 1. The outcome is that all group members either invest all (if their initial probability of investing all is greater than x_3) or none (if their initial probability is less than x_3). For no regions of phase space do the outcomes of the contribution probability and the contribution investment mechanisms coincide.

The location of the singular strategy in phase space defines the unique investment level of all group members. We see that if $\hat{z}_x^+ > 1$ and condition (4.28) is satisfied (i.e., the first term of the minimum is greater than $\hat{z}_x^+$), then all group members adopt 1. If $\hat{z}_x^+ < 0$ then all group members adopt 0. Varying model parameters alters the location of $\hat{z}_x$ in phase space. We find that an increase in group size results in a decrease in the investment level of all group members ($\partial_n \hat{z}_x^+ < 0$); Fig 4.9a. For examples of this group size effect see [136, 188]. We find that $\partial_\epsilon \hat{z}_x^+ < 0$; Fig 4.9b. Increasing $|\epsilon|$ results in a greater magnitude of decrease in the relative probability that an individual is targeted as it invests in strategy. We see $\partial_\delta \hat{z}_x^+ > 0$; Fig 4.9c. Increasing $|\delta|$ increases the cost of investment into the strategy. Finally, we see that increasing k increases $\hat{z}_x^+$; Fig 4.9d. The smaller the value of k the sharper the increase in the relative safety of risk dilution, with

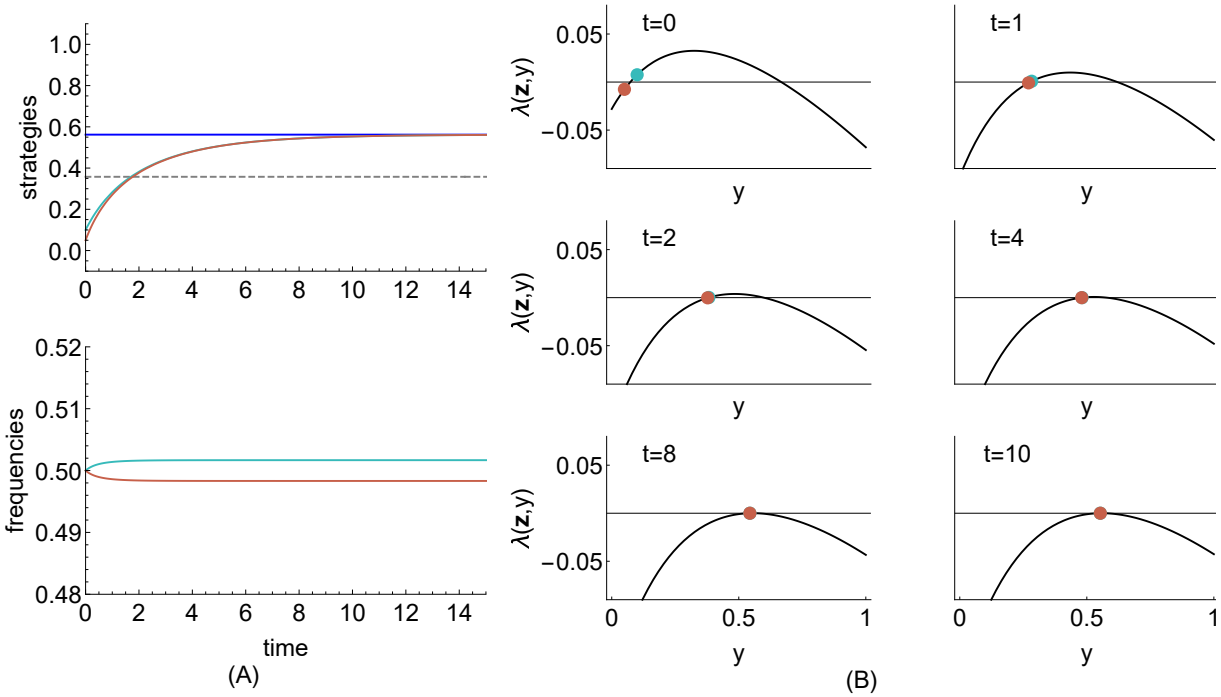


FIGURE 4.8: Invasion simulation. (A): Strategy and frequency dynamics. Starting with $z_1 = 0.1$ and $z_2 = 0.01$ both strategies converge in tandem to the CS solution $\hat{z}_x^+ \approx 0.59$ (blue solid line). The unstable interior fixed point of the contribution probability x_3 is plotted as a grey dashed line. Both strategies change in concert, neither deviating far from the other. As such z_1 and z_2 are able to co-exist at monomorphic ecological equilibrium, each with positive frequency. (B): The adaptive landscape.

Parameter values: $\gamma_C = 6.51, \gamma_D = 6.82, \epsilon = -0.54, k = 0.62, p_0 = 0.1, n = 12$.

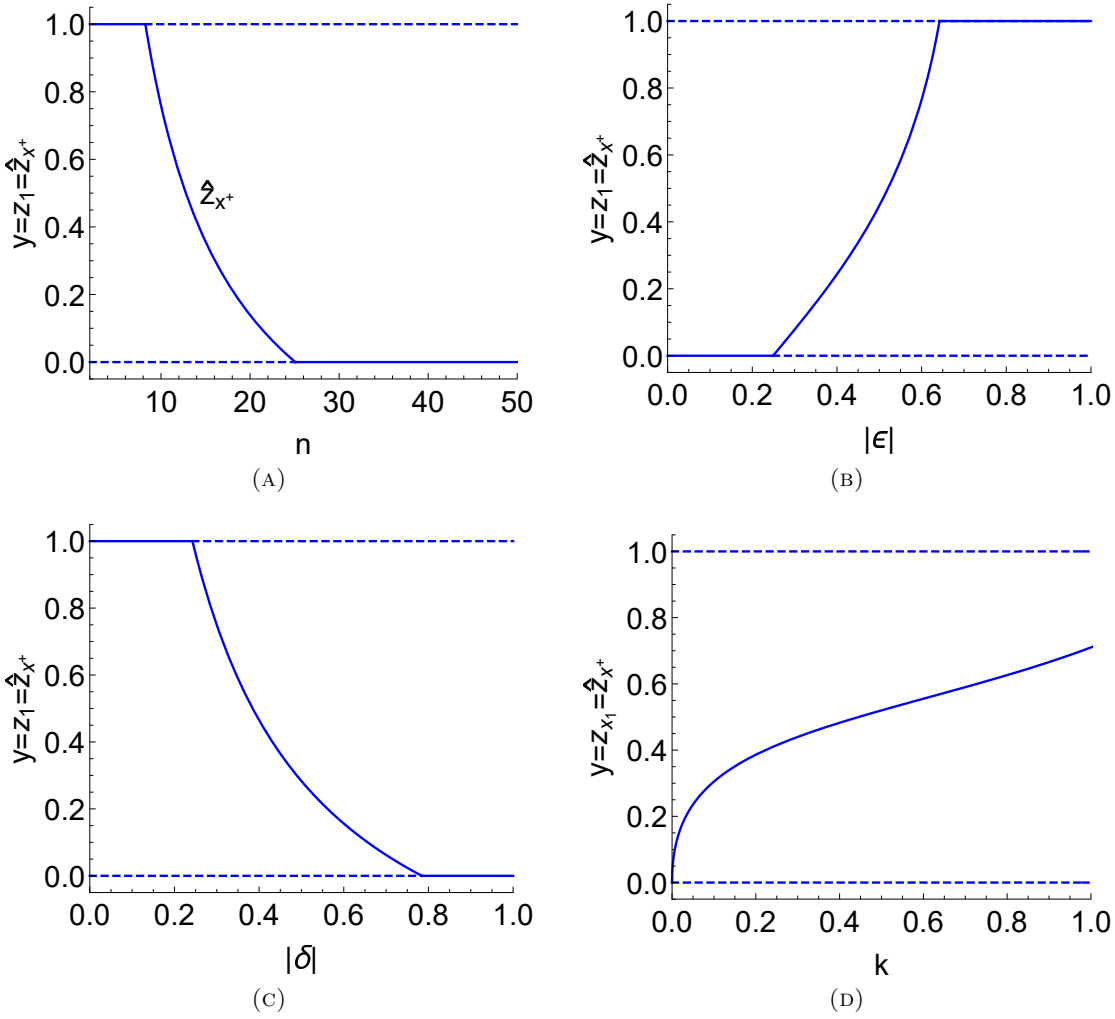


FIGURE 4.9: Bifurcation diagrams. As in other figures, solid lines show convergence stable strategies, dashed show non convergence stable outcomes, blue and red respectively show ESS-stable and ESS-unstable strategies. (A): The interior singular strategy $\hat{z}_x^+$ decreases with group size n, (B): increases with the increases in the magnitude of the unequal dilution of risk, $|\epsilon|$, and (C): decreases with increases in the magnitude of the decrease in foraging benefits, $|\delta$. (D): The singular strategy $\hat{z}_x^+$ increases with k, as explained in text.

increasing strategic investment. If $k << 1$ then an individual experiences almost all the fitness benefits from the unequal dilution of risk, while investing little into anti-predator activities. For increasing values of k, greater investments into strategy are required to maintain risk profile. While k is very large, the increases in safety through dilution are approximately linear, and functionally, our choice of $\epsilon(z)$ approximates the choice used in the example of section 4.4.1. In such cases, we find a value of k at which the singular strategy $\hat{z}_x^-$ bifurcates into phase space and closely resembles the singular strategy of example in 4.4.1.

Unlike the examples of sections 4.4.1-4.4.2, the results presented here show that a unique intermediate level of strategic investment is adopted by all group members. A convex decreasing choice of $\epsilon(z)$ allows individuals to reap a mixture of the fitness benefits gained through foraging and through dilution of risk. We note that linear or concave decreasing choices of $\epsilon(z)$ would not have

allowed for this (see condition 4.25). Rather than opting to maximise the fitness benefits of "either" foraging rewards or dilution of risk (as was the case of examples 4.4.1-4.4.2), group members are able to balance this trade-off using a flexible strategy. This result is at contrast to the results found when individuals are allowed to adopt the maximum investment with a flexible probability (analysed in 2). In the case of chapter 2, depending on their initial behaviours, individuals always adopt the zero or maximum investment strategy. Here, we have compared an interpretation whereby with a probability equal to its strategy an individual experiences one of two constant probabilities of being explicitly targeted by the predator, to an interpretation whereby an individual's experience of risk dilution continually changes with its strategy. We have shown that a continuous mapping between individual strategy and the dilution effect dramatically changes the expected composition of behaviours within a group.

4.4.4 Case 4: A nonlinear evasion function

In this subsection we consider that the strategic behaviours within the group contribute to a probability of group evasion. We introduce an averaging of strategy within the group; $r_g(y, \mathbf{z})$. In a population of ρ strategies at an ecological equilibrium, the expected strategic contribution in a group of size n is $\sum_{j=1}^{\rho} x_j z_j$. If we consider a potential novel strategy which contributes $y \in S$, then the average contribution in the group is defined by

$$r_g(y, \mathbf{z}) = \frac{1}{n}\left[ay + (n-a)\sum_{j=1}^{\rho} x_j z_j\right]. \tag{4.29}$$

Here, the factor a effectively scales novel strategy contribution to the average behaviour within the group. If $a = 0$ then the novel strategy does not influence the average group behaviour, whereas if $a = 1$ the contribution of the novel strategy is equal to the contribution of all other group members.

To describe the probability of group evasion given a monomorphic population with strategy z (i.e., $\rho = 1$ in equation 4.29), we use $g(y, z) = h\big(r_g(y, z)\big)$ where $h(x)$ is monotonic increasing, bounded by 0 and 1. In the limiting case of $a = 0$ we write $g(z)$. In this case the conditions for evolutionary stability are the same as in cases 1-4 above (since $g(z)$, independent of y and is treated as a constant in the selection gradient). Otherwise, while $a \neq 0$, $\epsilon(z)$ and $\gamma(z)$ are linear, we see that the condition for $\hat{z}$ to be an ESS becomes

$$\Big(2\delta\epsilon(1-g(y,z)) + 2(\delta + \gamma_D\epsilon + 2\delta\epsilon y)g_y(y,z) + (\gamma_D + \delta y)(1+\epsilon y)g_{yy}(y,z)\Big)\Big|_{y=z=\hat{z}} < 0 \tag{4.30}$$

(where we use notation $g_y(y, z)$ for the partial derivative of g w.r.t y). The first two terms of (4.30) are negative. From this we see that if $h''(x) \leq 0$ when evaluated at $\hat{z}$, then $\hat{z}$ is an ESS. In this case, we see no qualitatively new dynamical behaviour; the outcome of the adaptive dynamics is covered in one of the models presented in sections 4.4.1-4.4.3.

Application: Linear choices of $\gamma(z)$ and $\epsilon(z)$ and $g(y; \mathbf{z}; \mathbf{x}) = p_0 + p_1 r_g(y, \mathbf{z})/(k_1 + r_g(y, \mathbf{z}))$

In this example, we assume that the probability of group evasion is saturating, with a domain given by an averaged group behaviour. The expression for invasion fitness is omitted for space. Singular

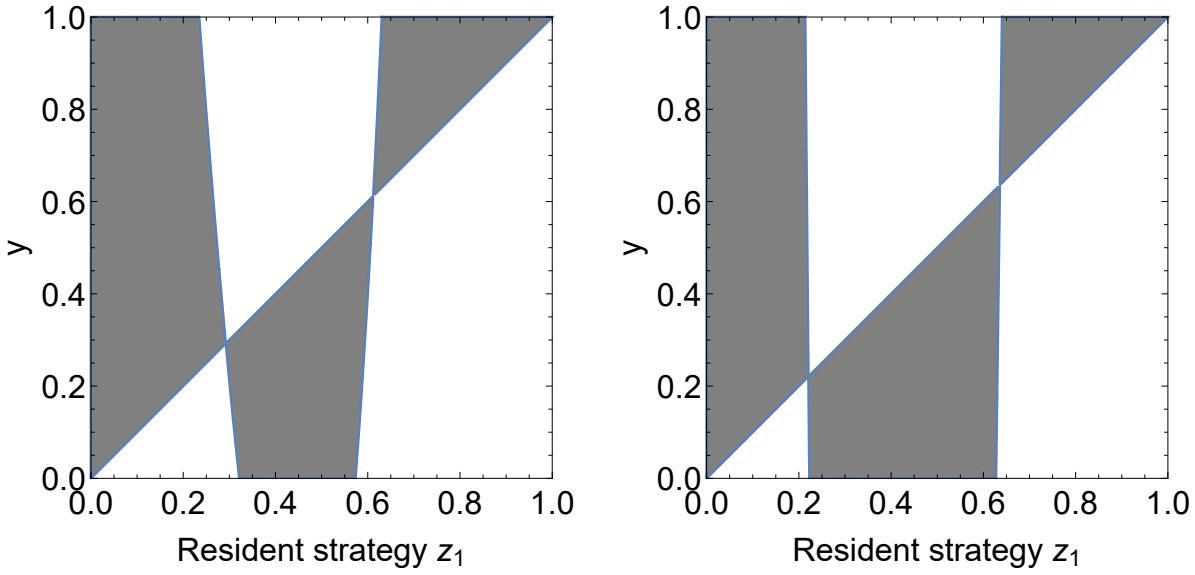


FIGURE 4.10: PIPs for the model with (A): $a = 1$ and (B): $a = 0$. In both PIPs we see that $\hat{z}_x^-$ is a CSS, whereas $\hat{z}_x^+$ is an ESS but is not CS. For comparison of results, parameter values are the same as in Fig 2.2a of chapter 2: $\gamma_D = 1.92, \gamma_C = 1.9, n = 25, \epsilon = -0.75, p_0 = 0.3, p_1 = 0.6, k_1 = 0.1$

strategies are the roots of the cubic:

$$\begin{aligned}\mathcal{C}_3 :&\delta\epsilon n(n + 2(p_0 + p_1 - 1))z_1^3 + \\ &\delta n(n - 1 + p_0 + 2\epsilon k_1(n - 2 + 2p_0)) + \delta(n + \epsilon(k_1 + 2k_1 n))p_1 + \gamma_D \epsilon n(p_0 + p1 - 1)z_1^2 + \\ &k_1(\gamma_D \epsilon(p_1 + n(2p_0 + p_1 - 2)) + \delta(p_1 + n(2(n - 1 + p_0) + \epsilon k_1(n + 2p_0 - 2) + p_1)))z_1 + \\ &k_1(\delta k_1 n(n - 1 + p0) + \gamma_D(p_1 - \epsilon k_1 n(1 - p_0)))\end{aligned}$$

We find at most two roots, $\hat{z}_x^+$ and $\hat{z}_x^-$, can co-exist within $(0, 1)$. From condition (4.30) we see that the singular strategies are always evolutionary stable. Convergence stability is determined by direct calculation.

Fig 4.10 provides PIPs for the respective cases of $a = 1$ and $a = 0$. In each case, two singular strategies, $\hat{z}_x^-$ and $\hat{z}_x^+$. The singular strategy $\hat{z}_x^-$ is continuously stable (CSS), while $\hat{z}_x^+$ is ESS but not CS. The singular strategies of the contribution investment found here closely resemble the singular strategies of the contribution probability; the interior fixed points x_3 and x_4 found in section 2.3.1 of chapter 2 (see Fig 2.2a). Note that x_3 and x_4 are neutrally evolutionary stable, but that x_4 is CS while x_3 is not. Viewed as an individual's strategy, x_4 is the probability that the individual invests 1, while x_3 determines the region of attraction of this probability, depending on the individual's initial behaviour. The adaptive dynamics of the continuos investment strategy result in all group members adopting an intermediate level of strategic investment defined by $\hat{z}_x^-$. Whereas, under the contribution probability, all individuals exhibit a probability, $x_4 \approx \hat{z}_x^-$, to choose the maximum level of investment.

While $a = 0$ the novel strategy y does not affect the probability that the group evades attack. In this case, with a monomorphic population the evasion function used here is identical to the one used

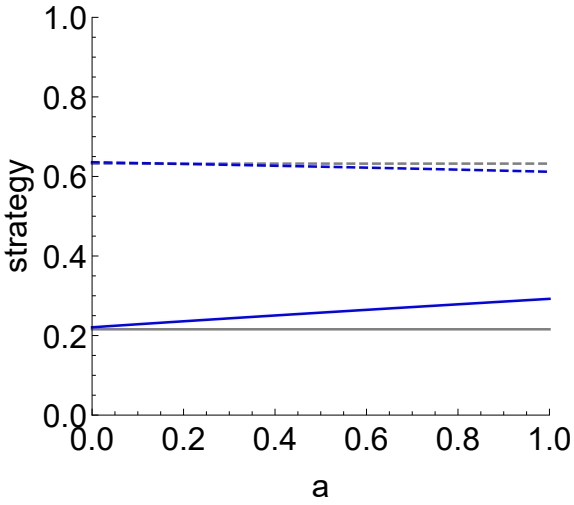


FIGURE 4.11: Comparison of the singular strategies of the contribution investment; $\hat{z}_x^+$ (blue dashed) and $\hat{z}_x^-$ (solid blue), and the singular strategies of the contribution probability; x_4 (solid grey) and x_3 (grey dashed). Increasing a novel strategy's ability to contribute (a) towards group evasion increases overall anti-predator investment within the group. When $a = 0$ the strategic investment level of a novel strategy does not influence the probability of group escape; $g(z)$. When $a > 0$ we use $g(y, z)$. For comparison of results, parameter values are the same as in Fig 2.2a; chapter 2: $\gamma_C = 1.9, \gamma_D = 1.92, \epsilon = -0.75, n = 25, p_0 = 0.3, p_1 = 0.6, k_1 = 0.1$.

in section 2.3.1 of 2 (but with an argument of investment level in place of the probability to invest 1). As such, $\hat{z}_x^-$ (blue solid line) is indistinguishable from x_4 (solid grey) and $\hat{z}_x^+$ (blue dashed) is in near perfect agreement with x_3 (grey dashed); Fig 4.11. When $a > 0$ there is additional incentive for a novel strategy to invest. We find that the intermediate investment level defined by $\hat{z}_x^-$ increases with a. Concomitantly, we see that $\hat{z}_x^+$ decreases with a. This latter finding effectively lowers the threshold (initial investment level of group members) for the maximum investment level to be adopted.

It might appear that the behavioural outcomes in groups under the contribution investment and the contribution probability mechanisms are similar. The underlying interpretation of strategy, however, explains how the outcomes are different. While strategy is viewed probabilistically, with the precise same intermediate probabilities x_4 and $1 - x_4$, all individuals in the population respectively adopt either of the 1 or 0 investment levels. The resultant monomorphic mixed strategy population is equivalent to a polymorphic population in which a frequency x_4 invest 1 and $1 - x_4$ invest 0. The upshot is that an expected nx_4 group members choose investment level 1 while an expected $n(1 - x_4)$ choose 0. In contrast, viewing strategy as a continuous investment level, all individuals choose some intermediate strategy.

Here, we highlight these differences explained above by allowing the introduction of novel investment strategies, into the polymorphic population defined by the contribution probability mechanism. We show that the extremal strategies both converge to the unique intermediate level of investment, defined by the singular strategy of the contribution investment. Such a process has been termed "merging" [60]. In the polymorphic population with investment strategies

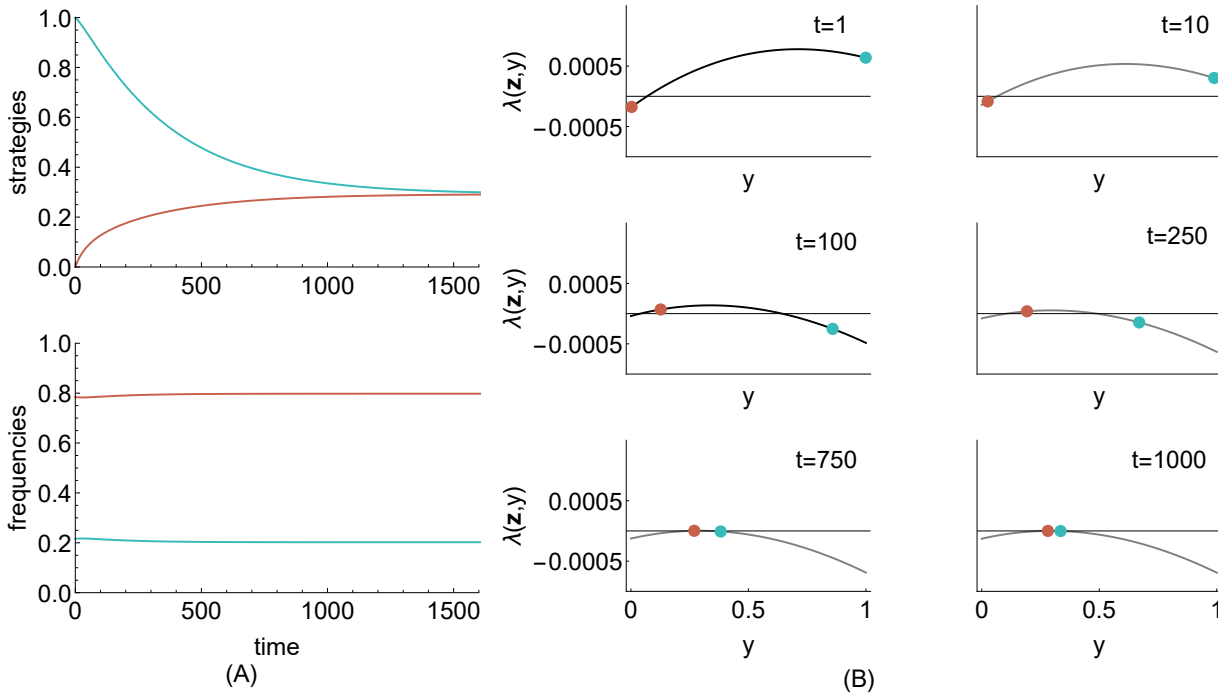


FIGURE 4.12: (A): The population is initially dimorphic with respective strategies $z_C = 1$ and $z_D = 0$ and frequencies x_4 and $1 - x_4$. Both strategies, in tandem, converge to the singular strategy of the contribution investment, z_x^-. Each strategy maintains a positive population frequency, up to the point where the strategies are equal, at which the frequencies remain constant. (B): The adaptive landscape. Novel strategies $y < z_C$ and $y > z_D$ invade the population: the directional changes in strategy are defined by the invasion fitness evaluated at z_C and z_D.

$\mathbf{z} = (z_C, z_D) = (1, 0)$ and frequencies $\mathbf{x} = (x_4, 1 - x_4)$, the invasion fitness of a novel strategy y is

$$\lambda(\mathbf{z}, y) = G(y; \mathbf{z}; \mathbf{x}) - \bar{G}) = G(y; z_C, z_D; x_4, 1 - x_4) - \bar{G}) \tag{4.32}$$

where $\bar{G} = x_4 G(z_C; \mathbf{z}; \mathbf{x}) + (1 - x_4) G(z_D; \mathbf{z}; \mathbf{x})$. The adaptive dynamics are given by the two selection gradients for the respective strategies z_C and z_D;

$$\dot{z}_C = \frac{\partial}{\partial y} \lambda(\mathbf{z}, y)|_{y=z_C} \tag{4.33a}$$

$$\dot{z}_D = \frac{\partial}{\partial y} \lambda(\mathbf{z}, y)|_{y=z_D} \tag{4.33b}$$

Both z_C (blue) and z_D (maroon) converge to the singular strategy of the contribution investment; (A) in Fig 4.12. The singular strategy z_x^- resides at a maxima of the adaptive landscape; panel "t=1000", and is an ESS once the population is monomorphic. While the population is dimorphic we see that the strategies of 1 and 0 lie at local minima of the invasion fitness; (B) of Fig 4.12. Changes in strategies are governed by $\dot{z}_C < 0$ and $\dot{z}_D > 0$, up to the point at which $z_C = z_D = z_x^-$. The population has transformed from a mixed strategy monomorphism (equivalent to a pure strategy polymorphism) to a monomorphism consisting of the precise same intermediate continuous strategy.

The results presented in this example demonstrate the differences between groups in which

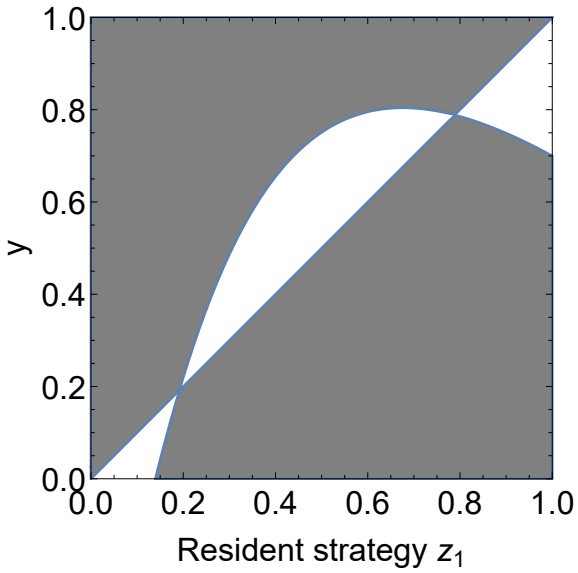


FIGURE 4.13: The PIP for the adaptive dynamics. Singular strategy $\hat{z}_x^-$ is CS but not an ESS, and $\hat{z}_x^+$ is neither an ESS nor CS.

individuals adopt a flexible level of investment into anti-predator activities (the contribution investment) and groups in which individuals adopt a fixed level of investment (equal to 1) with a flexible probability (the contribution probability). The adaptive dynamics of the continuous investment lead the population away from a polymorphic state of fixed extremal investment levels, to a monomorphic state of intermediate investment.

4.4.5 Case 5: A nonlinear evasion function, and a nonlinear choice of $\epsilon(z)$ or $\gamma(z)$

We finally move on to consider an evasion function as described by $h\big(r_g(y,\mathbf{z})\big)$ and nonlinear foraging or risk dilution functions. In general, deriving any meaningful expressions for the conditions of ESS and CS is not possible. We present a case whereby $\epsilon(z)$ is linear, and focus on the properties of $\gamma(z)$. It can be shown that the combination of such functional choices results in a combination of conditions for which a singular strategy may be CS but not an ESS.

Application: linear $\epsilon(z)$, $\gamma(z) = \gamma_D + \delta(k_0+1)z/(k_0+z)$ and $g(y;\mathbf{z};\mathbf{x}) = p_0 + p_1 r_g(y,\mathbf{z})/(k_1 + r_g(y,\mathbf{z}))$

In this example, we assume that a novel strategy does affect the average behaviour in the group (i.e., $a = 1$ in equation 4.29). With these payoffs, singular strategies are the roots of the a forth degree polynomial given in appendix C.2.4. We find at most two positive solutions $\hat{z}_x^-$ and $\hat{z}_x^+$. For a range of parameters we find $\hat{z}_x^-$ is CS but not an ESS; Fig 4.13. We also find cases where $\hat{z}_x^-$ is continuously stable (not visualised through PIPs). While $\hat{z}_x^-$ is CS but not an ESS, invasion of novel strategies leads the resident population of strategies to a global minimum of the adaptive landscape. The minima is referred to as a branching point of the adaptive dynamics [101]. At the branching point, two distinct strategies emerge and are able to co-exist in a so called protected polymorphism [79].

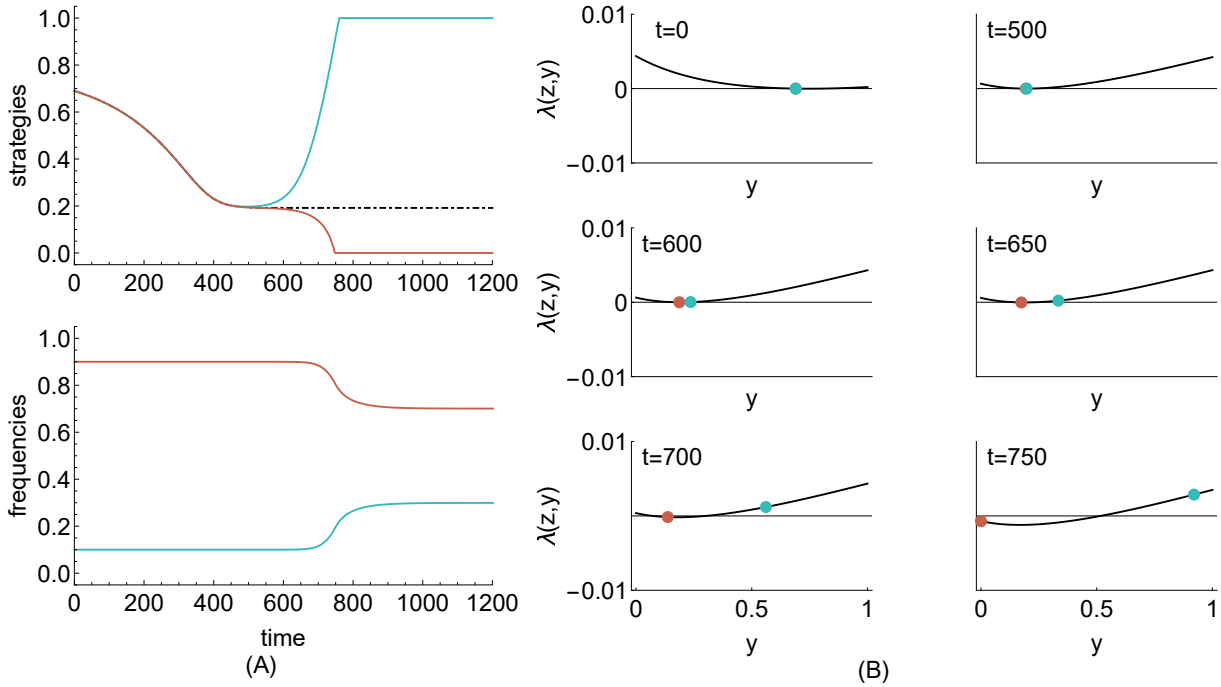


FIGURE 4.14: Invasion simulation. (A): Starting with z_1 (blue) and z_2 (maroon), the trajectories of the mean field canonical equation of adaptive dynamics (black dot-dashed curve) and the G-function are in perfect agreement up to the singular strategy $\hat{z}_x^-$. Once the population of strategies reaches $\hat{z}_x^-$ novel strategies are able to invade on either side and the population splits into two distinct strategy branches; z_1 adopts a high investment level while z_2 a low level. After branching the strategies are able to co-exist in a protected polymorphism. We see that z_1 and z_2 maintain non-zero frequencies. (B): The adaptive landscape. Between ''t=500" and"t=600" a gulf opens between z_1 and z_2. Each strategy climbs the adaptive landscape in direction proportional to increases in fitness: the z_1 increases in value while z_2 decreases. The result is bi-directional changes in strategy.

A numerical simulation illustrating how equations (4.8) emulate the branching process, is shown in Fig 4.14. We start with two similar strategies z_1 and z_2 able to co-exist in an approximately monomorphic population. Both strategies converge to the $\hat{z}_x^-$. Once the singular strategy is reached, we see that the strategies diverge into two respective branches; z_1 begins to increase in value and z_2 begins to decrease in value; panel (A) in Fig 4.14. After the two strategies are sufficiently distinct, the monomorphic invasion fitness (equation 4.7) fails. We apply $\lambda(\mathbf{z}, y)$ to emulate the branching process; panel (B) in Fig 4.14. Changes in strategy are governed by $\dot{z}_1 = \partial_y \lambda(\mathbf{z}, y)|_{y=z_1} > 0$ and $\dot{z}_2 = \partial_y \lambda(\mathbf{z}, y)|_{y=z_2} < 0$. Eventually, we see that both of the extremal strategies of 1 (maximum) or 0 (minimum) investment into anti-predator activities co-exist within the group.

Group size is generally considered as an important factor in explaining the anti-predation behaviour of its members. Here, bifurcation analysis indicates that group size (n) considerably affects the strategies of group members; Fig 4.15. Two transcritical bifurcations at $(n_-, 1)$ and $(n_+, 0)$ define the region for which $\hat{z}_x^- \in (0, 1)$; panel (A) of Fig 4.15. A bifurcation at n^* results in a change in the evolutionary stability of $\hat{z}_x^-$. In smaller sized groups ($n_- < n < n^*$) we see that $\hat{z}_x^-$ is continuously stable. Whereas, in larger groups ($n^* < n < n_+$), we find that $\hat{z}_x^-$ is a branching point. The stability change in the singular strategy can be explained by changes in the adaptive landscape. In cases where $n < n^*$, the singular strategy $\hat{z}_x^-$ is a local maxima of the adaptive

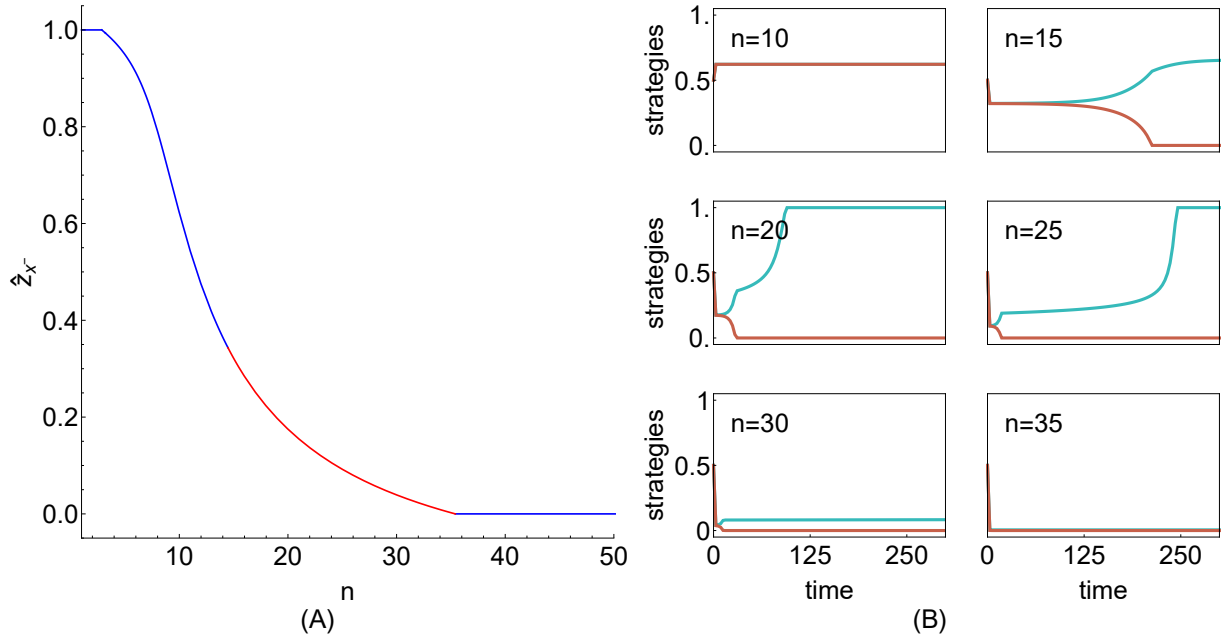


FIGURE 4.15: (A): Bifurcation diagram. At $(n_+, 0)$ and $(n_-, 1)$, $\hat{z}_x^-$ leaves phase space respectively passing through 0 and 1. For $n > n_+$, the 0 investment level is a CSS, and for $n < n_-$, 1 is a CSS. When present, the singular strategy loses evolutionary stability at $n^* \approx 14$. While $n_- < n < n^*$, $\hat{z}_x^-$ is a CSS. When $n^* < n < n_+$, $\hat{z}_x^-$ is a branching point. (B): Numerical simulations confirm the transition depicted in (A). In panel $n = 10$, the singular strategy is continuously stable and z_1 and z_2 converge to it. In panels $n = 15$ to $n = 30$ we see that z_1 and z_2 respectively diverge into two distinct solution branches.

landscape, at $\hat{z}_x^-$ becomes an inflection point of the adaptive landscape, and for $n^* < n$, $\hat{z}_x^-$ is a local minima; Fig 4.16.

We see that $\hat{z}_x^-$ decreases with n; (A) in Fig 4.15. Decreasing anti-predator investment with group size is to be expected: in smaller groups, the magnitude of the effect that unequal dilution of risk exerts on individual survivorship is more pronounced. It is then relatively more beneficial for group members to invest into anti-predator activities. As group size increases, each group member experiences a smaller probability of being targeted regardless of strategy. The unequal dilution effect exerts relatively smaller individual fitness benefits with increasing strategy investment. It is then relatively less beneficial to invest into anti-predator activities. What is not expected is the way in which the transition in investment occurs. As group size becomes larger than n^*, the group becomes behaviourally differentiated. There is no longer one unique strategy within the group which best balances the trade-off between foraging benefits and predation risk. Instead, group members are better off investing more, or less, than $\hat{z}_x^-$.

The main result of this example, illustrated in Fig 4.14 is a so-called ESS-coalition [80]. By definition of the ESS-coalition, an individual's payoff is the same when using either investment strategy of 0 or 1. However, the payoff an individual receives when using each respective strategy is comprised by a different combination of costs and benefits. The 0 investment strategy subjects a group member to the greatest foraging benefits. Whereas the 1 investment strategy subjects a group member to the lowest predation risk. The ESS-coalition relates directly to the results of chapter 2. In chapter 2 we find a monomorphic population in which all individuals invest 1 with a probability

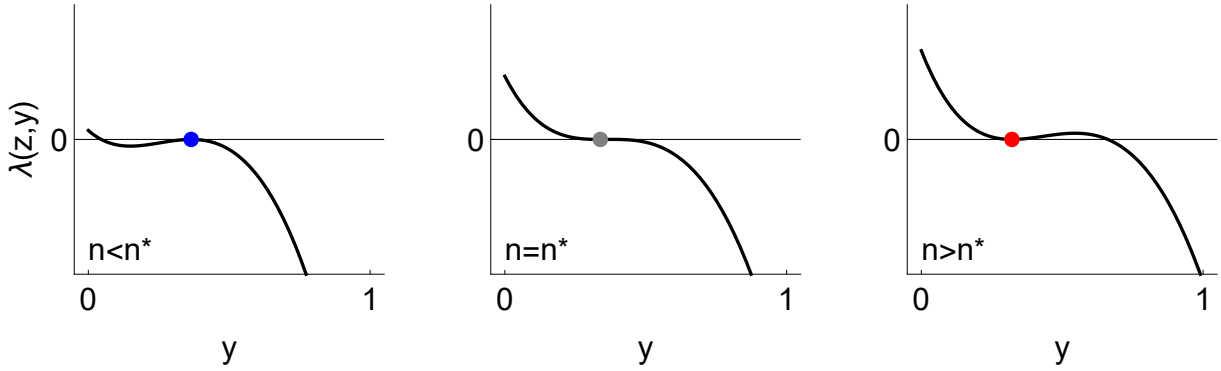


FIGURE 4.16: The adaptive landscape changes with group size, illustrating how $\hat{z}_x^-$ becomes a branching point. In each panel $\hat{z}_x^-$ is shown by the solid point. We see that while $n < n^*$, the singular strategy is a local maxima of the adaptive landscape and is therefore an ESS (blue). When $n = n^*$, the singular strategy is an inflection point and is ESS-neutral (grey). When $n^* < n$, the singular strategy is a minima of the adaptive landscape and is not an ESS (red).

x_4, equally valid interpreted as a polymorphic population in which a frequency x_4 invest 1 and $1-x_4$ invest 0. The example shown in this section (see Fig 4.14) predicts a polymorphic population in which x_1 invest 1 and $1-x_1$ invest 0, with the frequency of x_1 determined by equation (4.8b). There is however a subtle difference in these outcomes. When an individual invests 1 with a probability x_4 and (0 with probability $1 - x_4$) it demonstrates a non-consistent propensity to prioritise anti-predator activities; it randomly chooses a behaviour with a fixed probability. Hence, the mixed strategy interpretation of chapter 2 results in a random proportion x_4 whom invest 1. Whereas, upon a branching point of the contribution investment strategy, a fixed proportion of individuals consistently invest 1 and a fixed proportion of individuals consistently invest 0. Our results clearly indicate that two distinct behavioural phenotypes can emerge from group interactions. However, depending on how we interpret strategy, each phenotype may be consistently or non-consistently expressed.

4.5 Discussion

In this chapter we have studied the continuous extension of the model presented in chapter 2. We have analysed how allowing gradual change in anti-predator investment can lead to outcomes that qualitatively differ from discrete anti-predator strategies analysed therein. We have also shown when the two interpretations of strategy lead to the same (qualitative) outcomes. We focus on situations whereby group members are subjected to a foraging-predation trade-off; case 4 of Table 2.1. Considering the investment strategy, we develop a series of models which capture different qualitative structures of this trade-off. Each distinct structure is defined by functional choices which reflect the relationship between individual strategy and survivorship. We show that the choices of function govern the dynamical outcomes of the investment strategy, and thereby de-macerate the qualitative similarities and differences of the contribution probability and the contribution investment. To analyse changes in the investment strategy we apply the G-function, pioneered by Brown and Vincent [73].

Our results are broadly defined by our assumptions of how strategies within a group affect survivorship. Cases are considered where (i) the probability that the group eludes the predator is independent of the behaviours within the group. We also consider (ii) that the behaviours within the group contribute to some probability of a failed attack, i.e., the group evades. The probability of group evasion reflects a production function of a public good; individual survivorship probability (through failed attack) depends on the respective amounts of anti-predator investment made by group members, but is equal amongst all. In case (i) we find that all group members adopt the same investment level into anti-predator activities. Contrastingly, in case (ii) we find that it is possible for the individual behaviours within the group to become differentiated. In all examples, we contrast these findings to the discrete version of the game.

The first example which we present (section 4.4.1) adopts a linear structure to reflect case 4 of Table 2.1. Biologically, one expects such a trade-off in situations where group members' intake rates are described by a Holling type I functional response, e.g., decrease linearly with investment into vigilance. Increasing investment into anti-predator activities is known to improve the ability of a group member to appropriately respond to a predator. For example, the more effort into anti-predator vigilance an individual invests, the sooner it detects and reacts to a predator [255]. It is difficult to establish a linear relationship between an individual's anti-predator investment and its per-capita dilution of risk [145]. Yet, it is in general impossible to argue that any relationship universally describes the two quantities. If $\epsilon(y)$ is characterised by a linear form then our results demonstrate that the behaviour of all individuals in the group can be uniquely determined by a convergence unstable singular strategy. We provide conditions on individual dilution of risk and foraging gains which uniquely determine their anti-predator investment behaviours. Our results here indicate that investment into anti-predator behaviours decrease with increasing group size (through an increase in the basin of attraction of the zero investment strategy). This finding is generally supported throughout the wider theoretical and empirical literature on vigilance behaviours in groups of prey [18, 49, 55, 136, 191].

Interestingly, the results of this example mirror the results found in chapter 2. We see a qualitative equivalence between the outcomes of the continuous and mixed strategy dynamics because the invasion fitness of a mixed strategy (contribution probability) is linear in y, and the G-function adopts a linear choice of $\gamma(y)$ and $\epsilon(y)$. The only difference in the results presented here and those of chapter 2 occurs due to differences in the selection gradients of each interpretation of strategy (see appendix C.1). We note that in most parameter regimes the singular strategy of the contribution investment mechanism closely resembles the singular strategy of the contribution probability mechanism (see $x_3^{(p_1=0)}$ defined in equation (2.13) of section 2.3.1). As such, in most cases we would expect the behaviours in groups to be the same, regardless of whether we interpret strategy as continuous investment into anti-predator activities or the probability to invest all into anti-predator activities.

The second example presented in section 4.4.2 considers the case where an individual's foraging gains follow a convex decreasing relationship with investment into anti-predator activities. We keep the assumption that individual risk dilution decreases linearly with investment. In this case, each additional amount of anti-predator investment is less costly with respect to foraging gains than the last. We see qualitatively similar results compared to the example in section 4.4.1: all group

members invest either maximum or minimum effort into the anti-predator strategy, depending on the location of a convergence unstable singular strategy in phase space. We shown that the results of this example are symptomatic of our choice of $\gamma(y)$, and that a more general consideration of a nonlinear relationship between foraging and strategy investment can allow for a range of behaviours to characterise a group. For example, while foraging benefits are concave decreasing with strategic investment, it is possible for a continuously stable strategic investment level. In this case, all group members would choose some intermediate amount of effort into anti-predator activities.

In the third example (section 4.4.3) we consider that as an individual increases its investment into the anti-predator strategy, its relative conspicuousness decreases in a saturating manner. Decreases in foraging gains with strategic investment are assumed linear. Here, we find that a unique intermediate level of strategic investment (a convergence stable, evolutionary stable interior singular strategy) can potentially characterise the anti-predator behaviours of all group members. The results of this example reflect the expectations that selection pressure on anti-predator behaviours decreases with increases in group size. We find that the intermediate investment level adopted by the group decreases with group size. Anti-predator vigilance is theoretically predicted to follow a similar pattern [32, 49]. In line with our results, numerous observational studies have shown that investment into vigilance decreases with increases in group size [127, 191]. We also reflect on how other parameters affect the investment strategy adopted by group members. For example, by increasing the magnitude of the parameter ϵ we increase an individual's relative fitness benefits through its experience of an unequal risk dilution, as that individual increases its strategy investment. As such, selection favours a greater investment into anti-predator activities, and the entire group becomes characterised by a behavioural phenotype with greater preference to prioritise anti-predator activities.

The results presented in section 4.4.3 contrast the results presented in chapter 2. To approximate flexibility in behaviour, the model of chapter 2 presents an interpretation of a mixed strategy: an individual invests maximum effort (contributes) with a probability x (i.e., plays strategy C) and invests no effort (does not contribute) with a probability $1-x$ (i.e., plays D). While the mixed strategy $x_3^{(p_1=0)}$ (see equation (2.13), section 2.3.1) is not convergence stable, the only possible outcomes of the model in chapter 2 are pure strategy monomorphisms. All individuals invest all or nothing into anti-predator activities. On the other hand, while the adaptive dynamics of the investment strategy yield a convergence stable, evolutionary singular strategy within $(0,1)$ we see that all individuals find some intermediate level of investment into anti-predator activities. The adaptive dynamics of the contribution probability and the contribution investment are considerably different.

In the example of section 4.4.4 we move to consider a probability of group evasion which depends on the behaviours within the group. We maintain the simplest of assumptions that changes in individual foraging benefits and dilution risk are linear with strategic investment. Results show that an intermediate level of anti-predator investment (a continuously stable interior singular strategy) can describe the behaviours within the group. This finding directly relates to the results of section 2.3.1, chapter 2. In this payoff regime, the findings of section 2.3.1 are interpreted as all group members choosing the maximum investment strategy with a probability $x_4 \in (0,1)$; Fig 2.2a. We see that the singular strategy of the contribution investment is in near perfect agreement with x_4 when

the novel strategy is unable to affect the probability of group evasion; Fig 4.11. Such similarities in singular strategy might lead one to conclude similarities between the outcomes of this chapter and the discrete game of chapter 2. The underlying interpretation of strategy explains why this is not so. The outcome of the discrete strategy game is equivalent to a polymorphic population in which a frequency x_4 invest maximum, and $1 - x_4$ invest minimum effort. Here, we show that this polymorphic population is structurally unstable. By allowing novel investment strategies to enter a group composed of the two extremal investment strategies, we see that all group members end up choosing a unique intermediate investment level. This behavioural transition indicates why it is important to consider a flexible behaviour, rather than a fixed behaviour played with a flexible probability. Provided that the degree of the investment strategy can continuously vary, our findings show that anti-predator behaviours cannot be restricted to a discrete choice of options. Indeed, individuals within groups do flexibly alter their responses to predators depending on environment. For example, the proportion of time grouping individuals spend vigilant is known to vary according to perceived predator threat level [18, 288], as well as climactic conditions (see chapter 3). In predator inspection, the proximity to a predator which group members approach can change with environmental riskiness [289]. Some grouping species fine tune their nearest neighbour distances depending on foraging demands and predation risk [162].

In section 4.4.5 we provide an example in which a group becomes behaviourally differentiated. We see that some group members show a propensity to invest all into anti-predator activities. Other group members exhibit a propensity to invest nothing. Interestingly, in this example we see that the two factors which determine individual survivorship; the probability of group evasion and the probability of targeting given failed evasion, can define a critical group size at which the behavioural split occurs; Fig 4.15. This critical group size reflects an inflection point of the adaptive landscape; Fig 4.16. Groups smaller than this critical size exhibit all group members choosing the same intermediate investment level (a continuously stable singular strategy; a local maxima of the adaptive landscape), groups larger exhibit two distinct phenotypes.

The split in behavioural types of section 4.4.5 results from a branching point of the contribution investment strategy. An alternative way to generate a mixture of behaviours is the mixed strategy of chapter 2. The mixed strategy results in an expected proportion x_4 of the group to invest all, and $1-x_4$ to invest nothing. The results which we present here therefore indicate that there are two ways in which a group may come to exhibit two distinct anti-predator behavioural phenotypes. Either (i) a branching point of the adaptive dynamics of the investment strategy leads to a diversification from an initially homogenous group, to a mix where a fixed proportion consistently invest more and a fixed proportion consistently invest less. Or (ii) a convergence stable interior fixed point of the contribution probability strategy leads an initially homogenous group to a mix where all individuals randomly choose to invest all or nothing, with a fixed probability between zero and one. The former case generates consistent differences in individual behaviours, the latter case yields non-consistent differences in individual behaviours.

Concluding remarks

In this research we show that grouping individuals can exhibit a variety of highly flexible behavioural responses when subjected to a foraging predation trade-off. We demonstrate that different behavioural outcomes emerging from groups can be explained by a simple mathematical analysis of how an individual's behaviour affects its dilution risk and its foraging benefits. We show that different structures between risk dilution and foraging gains can directly explain the different behaviours within the group. Depending on the trade-off structure, we identify cases in which all group members choose minimal or maximal levels of effort in their anti-predator strategies, where all group members choose an intermediate level of effort in their anti-predator behaviour, and when the trade-off results in two distinct anti-predator behaviours both existing within the group. Each distinct behavioural outcome is explained in simple mathematical terms. Given that a wide range of group living species have to balance predation risk with foraging needs, the results of this research contribute to the understanding of anti-predator behavioural decisions in many grouping species.

Chapter 5

Conclusion

The results presented in this thesis concern the questions of how grouping animals subjected to predation adjust their choices of anti-predator behaviours. In the research, it is demonstrated that the three principal anti-predation effects of collective vigilance, risk dilution and confusion of the predator, play a considerable role in influencing these responses. Two alternative approaches to modelling the anti-predator behaviours within groups are applied. Analytical results concerning deterministic evolutionary game theory, for which group members can choose from a discrete set of behaviours, are given. Also through an analytical approach, the current research presents results which characterise the choices of grouping individuals that can flexibly choose from a continuum of behaviours. In both respective approaches, to understand the behaviours of group members, a core assumption of the present research is that the principal anti-predation effects may act complementary, co-dependently or compensatory when exerting pressures on the expression of phenotypes within prey groups.

In the first part of this thesis, grouping individuals are given a choice between contributing towards overall group safety, or not. Chapter 2 presents an array of qualitatively distinct behavioural dynamics from interactions within groups. Each distinct behavioural outcome depends on both the influence and the magnitude, of each principal anti-predation effect. By analytically investigating how anti-predation effects concomitantly influence the behavioural choice of a group member, this research provides new dynamical insights into the behaviours emanating from groups. No prior theoretical synthesis has incorporated all three anti-predation effects. The results presented here contribute to the existing knowledge of anti-predator behaviours within groups.

A particular contribution of chapter 2 is that the research provides a novel way of modelling an individual's risk dilution to depend on its strategic behaviour. From the hypothesis of unequal risk dilution, a trade-off (Fig 2.1) in individual predation risk and non-predation grouping benefits naturally emerges. To our knowledge this trade-off has not previously been theoretically explored. Through analysing qualitatively different structures of the trade-off, this research contributes to the understanding of how the behaviours of grouping individuals change. In particular, the research presented offers explanations as to how individual behaviours might change in different ways with factors previously assumed to have specific directional effects. Discrete game theoretical models studying anti-predator behaviours in groups tend to find a negative group size effect with either the frequency (of) or the probability that an individual contributes towards overall group safety (sometimes referred to as "cooperative" behaviour [63, 251]). The results of chapter 2 demonstrate that the negative group size effect is not universal but rather emblematic of a specific structure

in the trade-off identified (see Cases of Fig 2.1). Moreover, the research presented here provides results which demonstrate a less documented positive group size effect [136, 191].

The significance of considering behavioural mutations has also been highlighted. Random behavioural decisions are often overlooked in discrete strategy games (for example see [49, 55, 251]. Generally, individuals (and certainly animals) possess propensities to randomly explore their behavioural options [21, 40]. Such random effects are likely to affect the behavioural dynamics. The results presented in chapter 2 show that when considering small rates of behavioural exploration, one finds qualitatively similar results to the case of pure behavioural imitation. Conversely, when an individual has a high probability to randomly choose its behaviour, the population is always characterised by a more equitable split between behavioural phenotypes. A possible issue when not considering behavioural exploration (i.e., under pure imitation) is that in the absence of at least one exemplar of a certain behaviour (i.e., the population is at an absorbing state $x = 0$ or $x = 1$), there is no mechanism for behaviour to spread. Through considering exploration and a well mixed population, the results presented here ensure that group members can always observe, and subsequently choose a better performing behaviour. Through allowing recurrent behavioural mutations to affect anti-predator behaviour in groups, the present research provides a more biologically realistic contribution to models within the field.

Overall chapter 2 provides a new theoretical grounding which presents and explains how interactive pressures exerted by risk dilution, vigilance and confusion effects, result in specific behaviours emerging from the group. The model presented provides a mathematically concise explanation for these findings.

Chapter 3 qualitatively compares the theoretical results of chapter 2 to a real group living species subject to the principal anti-predation effects. Contributions of this section are made by providing qualitative evidence which supports the theoretical core of this research: that the three principal anti-predation effects are mutually interconnected. The data analysis presented in chapter 3 highlights the importance of this rational by demonstrating that theoretical predictions of behavioural phenotypes in groups are only appropriate when considering multiple anti-predation effects. Theoretical predictions corresponding to reduced models are made and compared to the full model (including all three anti-predation effects), and the observational data. The research presented here demonstrates that only when considering at lest two of the three principal anti-predation effects can grouping behaviours be predicted. This component of the research illustrates the relative importance of each anti-predation effect on group behaviour, and in particular, demonstrates the importance of an unequal dilution effect. A negative group size effect is theoretically predicted from the trade-off Fig 2.1 and empirically observed. This result supports the modelling approach, and justifies the use of ϵ.

Perhaps the most significant contribution of chapter 3 is that the research presented therein confirms the applicability of the theoretical framework of chapter 2. In this thesis, the theoretical models presented, are, unlike previous models e.g., Pulliam [130], not species specific, and instead intend to be broad enough to capture a wide range of group living species. The analysis presented in chapter 3 shows that when applied to an abstract choice or prey group, the theoretical results presented in this thesis are reflected in wild groups.

In most grouping animals, anti-predator behaviours are not limited to a discrete and finite set.

More realistically, group members can demonstrate flexible and continuously adjustable responses to predators [32, 46, 49]. Most previous models studying flexible anti-predator behaviours in groups of prey have applied to a static concept of an ESS to a continuous strategy [35]. A static interpretation leads to questions of whether the theoretically predicted ESS behaviours are actually attainable (and are often not reflected by data – see Pulliam [49] for example). Other models studying anti-predator behaviour in prey groups have applied a discrete strategy space, i.e., the mixed strategy, to approximate flexible behaviour [59]. This approach leads to the question of how well a flexible behaviour can be approximated by individuals probabilistically choosing from a set of discrete and fixed actions. In chapter 4 we present the adaptive dynamics of anti-predator behaviours within groups. We replace the static Maynard Smith [35] interpretation of an ESS with a dynamic interpretation. In the dynamic sense, behaviours within the group are occasionally challenged by novel behaviours. We also replace the notion of a flexible behaviour as a mixed strategy, and use a continuously adjustable strategy. The latter part of this research thesis answers the following key questions: how does an appropriately scaled anti-predator behaviour within groups change with time? Is the behaviour resistant to invasion? and can the behaviour invade similar behaviours? Analytic expressions are provided which quantify the behavioural dynamics.

Building on from theoretical construct of the first section of this thesis, chapter 4 identifies how a group member's investment into a flexible anti-predation behaviour affects both its foraging abilities and it's predation risk. Through analysing different structures of one case of this trade-off (functionally equivalent to Fig 2.1), this section of research contributes by providing analytical conditions for qualitatively distinct behavioural dynamics in groups. Comparisons are made between the resulting strategic choices when group members free to choose a continuous behaviour, and the case where group members face a discrete behavioural choice. Depending on the structure of the trade-off, equivalences and differences between the discrete and continuous strategy outcomes are identified and explained. The research explains when all group members choose the same discrete behaviour (pure strategy), when all group members choose a behaviour not accessible through a discrete consideration of strategy, and when a group becomes behaviourally differentiated; some group members choose one of the (pure) discrete behaviours, while other group members do not. The results of this research show that when individuals flexibly choose behaviours in response to predation, one can expect different behavioural choices within the group compared to when group members have an "either or" choice.

The research presented within this thesis is limited in the sense that we only consider anti-predator behaviours that are defined by how they affect predator detection and evasion. For instance, individuals may choose to prioritise an anti-predator strategy, or individuals may choose to prioritise foraging. There is a third option which we do not consider: individuals may choose to leave the group. In reality, groups may disband seasonally [18, 122] or due to conflicting behaviours within the group [4]. Future development of the present research would consider a so-called "loner" strategy [57]. Tentative investigation (not shown presently) with a similar structural trade-off described in this research, indicates that the dilution effect plays a dominant role in the behavioural dynamical outcomes. In the absence of a dilution effect, the three strategy analogue of the present research produces a rock-paper-scissors type game, characterised by a heteroclinic cycle. When all individuals choose to prioritise the anti-predator behaviour, all individuals receive greatest safety

(and thereby fitness). Yet, there is always an incentive for an individual to rely upon the anti-predator behaviours of others and increase feeding. As such, all group members choose to prioritise foraging. Eventually, no group members choose the anti-predator behaviour and all group members are subject to maximum predation risk. Individuals receive a more preferable risk profile by leaving the group, and therefore choose the loner strategy. Each individual in a collection of loners receives a less favourable risk profile than it would in a group where all choose anti-predator behaviour. The group reforms with all individuals choosing the pro-active anti-predator behaviour. The cycle repeats. Interestingly, the introduction of a dilution effect to such a game substantially changes the dynamics. With pure dilution, the strategy that prioritises foraging is always a viable behaviour, and when the magnitude of the dilution effect on fitness is small, the dynamics are characterised by limit cycles.

With the methods presented in this thesis, we aim to understand more about the behavioural decisions made by grouping animals when threatened with a predator. Increasing this understanding will have wide implications for models of predator-prey interactions and population dynamics.

Appendix A

A.1 The Kramers-Moyal Expansion

A Kramers-Moyal expansion is applied to approximate a continuos probability density function $p(x,t)$ for large N. Let $p(x,t)$ denote the probability density function of the frequency of the C. We rescale time such that $t = \tau/N$, set $x = mj/N$ so that $p(x,t) = NP^{\tau}(j)$ and

$$p\left(x, t+\frac{1}{N}\right) - p(x,t) = p\left(x-\frac{1}{N}, t\right)T^{+}(x-\frac{1}{N}) + \tag{A.1a}$$

$$p\left(x+\frac{1}{N}, t\right)T^{-}(x+\frac{1}{N}) - p(x,t)T^{-}(x) - p(x,t)T^{+}(x) \tag{A.1b}$$

The transition probabilities and probability densities are approximated by a taylor series in x and t;

$$p\left(x, t+\frac{1}{N}\right) \approx p(x,t) + \frac{\partial}{\partial t}p(x,t)\frac{1}{N} + \frac{1}{2!}\frac{\partial^2}{\partial t^2}p(x,t)\frac{1}{N^2} + ... \tag{A.2a}$$

$$p\left(x \pm \frac{1}{N}, t\right) \approx p(x,t) \pm \frac{\partial}{\partial x}p(x,t)\frac{1}{N} + \frac{1}{2!}\frac{\partial^2}{\partial x^2}p(x,t)\frac{1}{N^2} \tag{A.2b}$$

$$T^{\pm}(x \pm \frac{1}{N}) \approx T^{\pm}(x) \pm \frac{\partial}{\partial x}T^{\pm}(x)\frac{1}{N} + \frac{1}{2!}\frac{\partial^2}{\partial x^2}T^{\pm}(x)\frac{1}{N^2} \pm ... \tag{A.2c}$$

Substituting the expressions of lines A.2a, A.2b, A.2c into equation A.1 we obtain

$$\left(p(x,t) + \frac{\partial}{\partial t}p(x,t)\frac{1}{N} + \frac{1}{2!}\frac{\partial^2}{\partial t^2}p(x,t)\frac{1}{N^2} + ...\right) - p(x,t) \approx$$
$$\left(p(x,t) - \frac{\partial}{\partial x}p(x,t)\frac{1}{N} + \frac{1}{2!}\frac{\partial^2}{\partial x^2}p(x,t)\frac{1}{N^2}\right)\left(T^{+}(x) - \frac{\partial}{\partial x}T^{+}(x)\frac{1}{N} + \frac{1}{2!}\frac{\partial^2}{\partial x^2}T^{+}(x)\frac{1}{N^2}\right) +$$
$$\left(p(x,t) + \frac{\partial}{\partial x}p(x,t)\frac{1}{N} + \frac{1}{2!}\frac{\partial^2}{\partial x^2}p(x,t)\frac{1}{N^2}\right)\left(T^{-}(x) + \frac{\partial}{\partial x}T^{-}(x)\frac{1}{N} + \frac{1}{2!}\frac{\partial^2}{\partial x^2}T^{-}(x)\frac{1}{N^2}\right) -$$
$$p(x,t)T^{-}(x) - p(x,t)T^{+}(x).$$

Truncating at terms of order higher than $1/N^2$ in the above expression we find

$$\frac{\partial}{\partial t}p(x,t)\frac{1}{N} \approx$$
$$\frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\frac{T^+(x)+T^-(x)}{N^2}\right)p(x,t)+\frac{1}{2}\left(T^+(x)+T^-(x)\right)\frac{\partial^2}{\partial x^2}\left(\frac{1}{N^2}p(x,t)\right)+$$
$$\frac{\partial}{\partial x}\left(\frac{1}{N}p(x,t)\right)\frac{\partial}{\partial x}\left(\frac{T^+(x)+T^-(x)}{N}\right)+ \tag{A.4a}$$
$$\frac{\partial}{\partial x}\left(\frac{T^-(x)-T^+(x)}{N}\right)p(x,t)+\left(T^-(x)-T^+(x)\right)\frac{\partial}{\partial x}\left(\frac{1}{N}p(x,t)\right) \tag{A.4b}$$

and applying the Chain rule to lines A.4a and A.4b we find that

$$\begin{aligned}\frac{1}{2}\frac{\partial^2}{\partial x^2}\left(p(x,t)\frac{T^+(x)+T^-(x)}{N^2}\right) =&\frac{1}{2}p(x,t)\frac{\partial^2}{\partial x^2}\left(\frac{T^+(x)+T^-(x)}{N^2}\right)+\\ &\frac{1}{2}\left(T^+(x)+T^-(x)\right)\frac{\partial}{\partial x^2}\left(\frac{1}{N^2}p(x,t)\right)+\\ &\frac{\partial}{\partial x}\left(\frac{1}{N}p(x,t)\right)\frac{\partial}{\partial x}\left(\frac{T^+(x)+T^-(x)}{N}\right)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial x}\left(p(x,t)\frac{T^-(x)-T^+(x)}{N}\right) =&p(x,t)\frac{\partial}{\partial x}\left(\frac{T^-(x)-T^+(x)}{N}\right)+\\ &\left(T^-(x)-T^+(x)\right)\frac{\partial}{\partial x}\left(\frac{1}{N}p(x,t)\right).\end{aligned}$$

Equation A.1 becomes

$$\frac{\partial}{\partial t}p(x,t)\frac{1}{N} \approx \frac{\partial}{\partial x}\left(p(x,t)\frac{T^-(x)-T^+(x)}{N}\right)+\frac{1}{2}\frac{\partial^2}{\partial x^2}\left(p(x,t)\frac{T^+(x)+T^-(x)}{N^2}\right) \tag{A.7}$$

which can be written as the diffusion approximation of the one step process;

$$\frac{\partial}{\partial t}p(x,t) = -\frac{\partial}{\partial x}\left[A(x,t)p(x,t)\right]+\frac{1}{2}\frac{\partial^2}{\partial x^2}\left[B^2(x,t)p(x,t)\right] \tag{A.8}$$

with drift and diffusion terms $A(x,t)$ and $B^2(x,t)$ respectively given by

$$A(x,t) = T^+(x)-T^-(x)$$
$$B^2(x,t) = \frac{1}{N}(T^+(x)+T^-(x))$$

A.2 The van Kampen approach

The Van Kampen expansion is an intuitive way of deriving the Fokker-Planck equation in that it indicates that the deterministic drift coefficient should be scaled differently from the fluctuating diffusion coefficient. Essentially the average movement should be of greater significance and of a

greater order than the random fluctuations. As such, we introduce the van Kampen ansatz [84]

$$i = Nx(\tau) + \sqrt{N}\xi \tag{A.10}$$

so that $i/N = x(\tau) + N^{-\frac{1}{2}}\xi$. Here $x(\tau)$ is the non-fluctuating varibale of the system and is of order N. The variable ξ is the fluctuating varibale with order $N^{\frac{1}{2}}$. By introducing the shift operators $E^{\pm}f(n) = f(n \pm 1)$ and writing t rather than τ the master equation can be written

$$\frac{dP^t(i)}{dt} = (1 - \mathbb{E}^+)P^t(i)T^+(i) + (1 - \mathbb{E}^-)P^t(i)T^-(i) \tag{A.11}$$

Writing the distribution of $\xi(t)$ according to the probability density function $p(\xi, t) = P^t(Nx(t) + \sqrt{N}\xi)$, it follows that we have

$$\frac{dP^t(i)}{dt} = \frac{\partial p(\xi,t)}{\partial \xi}\frac{d\xi}{dt} + N^{\frac{1}{2}}\frac{\partial p(\xi,t)}{\partial \xi}\frac{d\xi}{dt} = \frac{\partial p(\xi,t)}{dt} - N^{\frac{1}{2}}\frac{\partial p(\xi,t)}{\partial \xi}\frac{dx(t)}{dt} \tag{A.12}$$

The master equation reads

$$\begin{aligned}\frac{\partial p(\xi,t)}{dt} - N^{\frac{1}{2}}\frac{\partial p(\xi,t)}{\partial \xi}\frac{dx(t)}{dt} =& (1 - \mathbb{E}^+)p(\xi,t)T^+(Nx(t) + N^{\frac{1}{2}}\xi) + \\ & (1 - \mathbb{E}^-)p(\xi,t)T^-(Nx(t) + N^{\frac{1}{2}}\xi)\end{aligned} \tag{A.13a}$$

The transition rates can be expanded as a taylor series of $y = N^{\frac{1}{2}}\xi$ according to

$$T^{\pm}(xN + N^{\frac{1}{2}}\xi) \approx NT^{\pm}(x) + \frac{\partial}{\partial x}T^{\pm}(x)\xi\sqrt{N} + \cdots \tag{A.14a}$$

and the shift operators by

$$\mathbb{E}^{\pm} - 1 = \pm\frac{1}{\sqrt{N}}\frac{\partial}{\partial \xi} + \frac{1}{2N}\frac{\partial^2}{\partial \xi^2} + \cdots \tag{A.15}$$

It follows that by substituting the A.15 and A.14a into equation A.13a we have

$$\frac{\partial p(\xi,t)}{dt} - N^{\frac{1}{2}}\frac{\partial p(\xi,t)}{\partial \xi}\frac{dx(t)}{dt} = \tag{A.16a}$$

$$\begin{aligned}& p(\xi,t)\left(-\frac{1}{\sqrt{N}}\frac{\partial}{\partial \xi} + \frac{1}{2N}\frac{\partial}{\partial \xi^2} - \cdots\right)\left(NT^+(x) + \frac{\partial}{\partial x}T^+(x)\xi\sqrt{N} + \cdots\right) + \\ & p(\xi,t)\left(\frac{1}{\sqrt{N}}\frac{\partial}{\partial \xi} + \frac{1}{2N}\frac{\partial}{\partial \xi^2} + \cdots\right)\left(NT^-(x) + \frac{\partial}{\partial x}T^-(x)\xi\sqrt{N} + \cdots\right)\end{aligned} \tag{A.16b}$$

The terms of order $N^{\frac{1}{2}}$ give the macroscopic (deterministic) equation for the system:

$$\frac{dx(t)}{dt} = T^+(x) - T^-(x) \tag{A.17}$$

Collecting terms in order of N^0 we obtain

$$\frac{\partial p(\xi,t)}{\partial \xi} = -\frac{\partial}{\partial x}\left(T^+(x) - T^-(x)\right)\frac{\partial}{\partial \xi}\xi p(\xi,t) + \frac{1}{2}\left(T^+(x) + T^-(x)\right)\frac{\partial^2}{\partial \xi^2}p(\xi,t) \tag{A.18}$$

By substituting $z = N^{-\frac{1}{2}}\xi + x(t)$ we obtain the Fokker-Planck equation

$$\frac{\partial}{\partial t}p(x,t) = -\frac{\partial}{\partial x}\big(A(x,t)p(x,t)\big) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\big(B^2(x,t)p(x,t)\big) \tag{A.19}$$

A.3 Model Analysis

A.3.1 Investigating state shifts with parameters for $\delta, \epsilon < 0$

As we have demonstrated in the main text the movement of the key determining values – the values of n which determine the positions of the transcritical bifurcations in the system – that enable ecological statements to be made from the mathematics. Therefore in this and the subsequent section we have determined the sign of the partial derivatives of n_1 and n_2 with respect to the physical parameters.

We find that for n_1

$$\frac{\partial n_1}{\partial \epsilon} = \frac{\gamma_C(1-p_0)}{\delta} < 0 \ , \ \frac{\partial n_1}{\partial p_0} = -\frac{(\delta + \epsilon\gamma_C)}{\delta} < 0 \ , \ \frac{\partial n_1}{\partial \gamma_C} = -\frac{\epsilon\gamma_D(1-p_0)}{\delta^2} > 0 \tag{A.20}$$

For n_2 we find that

$$\frac{\partial n_2}{\partial \gamma_C} = -\frac{\epsilon\gamma_D\big((1+k)(1-p_0)-p_1\big)}{\delta^2(1+\epsilon)(1+k)} = -\frac{\overbrace{\epsilon\gamma_D}^{-ve}\big(\overbrace{1-p_0-p_1+k-kp_0}^{+ve}\big)}{\underbrace{\delta^2(1+\epsilon)(1+k)}_{+ve}} > 0$$

$$\frac{\partial n_2}{\partial \epsilon} = \frac{\gamma_D\big((1+k)(1-p_0)-p_1\big)}{\delta(1+\epsilon)^2(1+k))} = \frac{\overbrace{\gamma_D}^{+ve}\big(\overbrace{1-p_0-p_1+k-kp_0}^{+ve}\big)}{\underbrace{\delta(1+\epsilon)^2(1+k)}_{-ve}} < 0$$

$$\frac{\partial n_2}{\partial p_1} = -\frac{\delta+\epsilon\gamma_C}{\delta(1+\epsilon)(1+k)} < 0$$

$$\frac{\partial n_2}{\partial p_0} = -\frac{\delta+\epsilon\gamma_C}{\delta(1+\epsilon)} < 0$$

$$\frac{\partial n_2}{\partial k} = \frac{(\delta+\epsilon\gamma_C)p_1}{\delta(1+\epsilon)(1+k)^2} = \frac{\overbrace{(\delta+\epsilon\gamma_C)p_1}^{-ve}}{\underbrace{\delta(1+\epsilon)(1+k)^2}_{-ve}} > 0 \tag{A.21a}$$

Investigating state shifts with model parameters for $\delta, \epsilon > 0$

Since the signs of δ and ϵ are reversed, the values n_1 and n_2 move in the opposite direction of travel with other parameters as compared with above:

$$\frac{\partial n_1}{\partial \epsilon} > 0 \ , \ \frac{\partial n_1}{\partial p_0} < 0 \ , \ \frac{\partial n_1}{\partial \gamma_C} < 0 \ \text{(equivalently } \frac{\partial n_1}{\partial \gamma_D} > 0\text{)} \tag{A.22}$$

For n_2 changes with all parameters follow;

$$\frac{\partial n_2}{\partial p_1} < 0 \ , \ \frac{\partial n_2}{\partial p_0} < 0 \ , \ \frac{\partial n_2}{\partial k} > 0 \ , \ \frac{\partial n_2}{\partial \epsilon} > 0 \text{ and } \frac{\partial n_2}{\partial \gamma_D} > 0 \text{ (equivalently } \frac{\partial n_2}{\partial \gamma_C} < 0) \tag{A.23}$$

A.3.2 Considering different the evasion function $g(n,x) = \frac{p_0 n}{k_1+n} + \frac{p_1 x}{k_2+x}$

We solve the new dynamical system

$$\begin{aligned}\dot{x} = x(1-x)\Bigg\{ &\frac{(\delta+\epsilon\beta_{s_1})\big[nk_2(1-p_0)+k_1k_2-x(n(p_1+p_0-1)+k_1(1-p_1))\big]}{n(k_1+n)(k_2+x)(1+\epsilon x)} - \\ &\frac{\delta\big[n(k_1k_2+nk_2)+nx(k_1+\epsilon k_1k_2+n+\epsilon nk_2)+\epsilon nx^2(k_1+n)\big]}{n(k_1+n)(k_2+x)(1+\epsilon x)}\Bigg\} \\ &= 0\end{aligned} \tag{A.24a}$$

to once again find 4 fixed points for the system; $x_1 = 0$, $x_2 = 1$ and two quadratic solutions x_3 and x_4. For ease of reading we write x_3 and x_4 such that

$$\begin{aligned}x_3 =& \frac{1}{2\delta\epsilon n(k_1+n)}\Big\{A + \sqrt{B+C^2}\Big\} \\ x_4 =& \frac{1}{2\delta\epsilon n(k_1+n)}\Big\{A - \sqrt{B+C^2}\Big\}\end{aligned}$$

where

$$\begin{aligned}A =& (\delta+\epsilon\gamma_C)(k_1(1-n-p_1-\epsilon k_2 n)-n(n-1+\epsilon k_2 n+p_0+p_1))+ \\ & \epsilon\gamma_C(k_1(n+\epsilon k_2 n)+n(n+\epsilon k_2 n)) \\ B =& 4\delta\epsilon k_2 n(n+k_1)\big[(\delta+\epsilon\gamma_C)((1-n)(n+k_1)-np_0)+\epsilon\gamma_C(n(n+k_1))\big] \\ C =& (\delta+\gamma_C)(k_1(1-n-\epsilon k_2 n-p_1)-n(n-1+\epsilon k_2 n+p_0+p_1))+ \\ & \epsilon\gamma_C(k_1(n+\epsilon k_2 n)+n(n+\epsilon k_2 n))\end{aligned}$$

We work on the case $\delta, \epsilon < 0$ to find new critical group sizes that determine the stability of $x^* = x_1, x_2, x_3, x_4$ (for the case $\delta, \epsilon > 0$ we simply find a reversal in the sign of the one-dimensional Jacobian).

Three new critical group sizes are found with the new system;

$$\begin{aligned}n_1 =& \frac{(\delta+\epsilon\gamma_C)(1-p_0)+k_1\gamma_C+\sqrt{(\delta+\epsilon\gamma_C)^2(1-2p_0+p_0^2)+(\delta+\epsilon\gamma_C)(2k_1\delta+2k_1\delta p_0)+(k_1\delta)^2}}{2\delta} \\ n_2 =& \frac{(\delta+\epsilon\gamma_C)((1+k_2)(1-k_1-p_0)-p_1)+\epsilon k_1(1+k_2)\gamma_D+\sqrt{(\delta+\epsilon\gamma_C)^2\Gamma+(\delta+\epsilon\gamma_C)\Psi-\Theta}}{2\delta(1+\epsilon)(1+k_2)}\end{aligned}$$

where

$$\begin{aligned}\Gamma &= (k_1^2(1+k_2)^2+2k_1(1+k_2)(1+k_2+p_0+k_2p_0-p_1)+((1+k_2)(-1+p_0)+p_1)^2) \\ \Psi &= 2\gamma_D\epsilon k_1(1+k_2)((1+k_2)(1+k_1+p_0)-p_1) \\ \Theta &= \epsilon^2 k_1^2(1+k_2)^2\gamma_D^2.\end{aligned}$$

and n_3 which is omitted but available upon request.

We find that x_1 is stable if and only if

$$\begin{aligned}\lambda_1 =& \frac{(\delta+\epsilon\gamma_C)((-1+n)(k_1+n)+np_0)-n(k_1+n)\epsilon\gamma_C}{n(k_1+n)} < 0 \iff \\ & n^2\delta + n((p_0-1)(\delta+\gamma_C)+k_1\delta) - k_1(\delta+\gamma_C) < 0 \iff \\ & n > n_1\end{aligned}$$

The equilibrium x_2 is stable if and only if

$$\begin{aligned}\lambda_2 =& \frac{\gamma_D(1+k_2)((k_1+n)(n-1+\epsilon n)+np_0)+\gamma_D(k_1+n)p_1}{(1+\epsilon)(1+k_2)n(k_1+n)} - \\ & \frac{\gamma_C(1+\epsilon)[(1+k_2)((n-1)(k_1+n)+np_0)+(k_1+n)p_1]}{(1+\epsilon)(1+k_2)n(k_1+n)} < 0 \iff \\ & -n^2(\delta(1+\epsilon+k_2+\epsilon k_2)) + n((\delta+\epsilon\gamma_C)(1-p_0+k_2-k_2p_0-p_1) - \delta(k_1+\epsilon k_1+k_1k_2-k_1p_1)) \\ & +(\delta+\epsilon\gamma_C)(k_1+k_1k_2-k_1p_1) < 0\end{aligned}$$

Using the quadratic formula where by definition

$a = \delta(1+\epsilon)(1+k_2) < 0$ and $c = (\delta+\epsilon\gamma_C)(k_1+k_1k_2-k_1p_1) < 0$ and $n > 0$ we must have

$b = ((\delta+\epsilon\gamma_C)(1-p_0+k_2-k_2p_0-p_1) - \delta(k_1+\epsilon k_1+k_1k_2-k_1p_1)) > 0.$

Then

$an^2+bn+c = 0 \iff n = \dfrac{-b\pm\sqrt{b^2-4ac}}{2a}$ $\therefore$ $\lambda_2 < 0 \iff an^2+bn+c < 0 \iff$

$n < \dfrac{-b+\sqrt{b^2-4ac}}{2a}$ for positive values of n. Thus it follows that

$n < n_2$

(A.29a)

A.3.3 Investigating state shifts with model parameters for $\delta, \epsilon < 0$

We find that as we vary our model parameters, the values n_1 and n_2 which determine the behaviour of a given sized group change.

We find that n_1 changes with γ_C, γ_D, ϵ, p_0 and k_1. This is found by taking the partial derivative of n_1 with respect to each parameter; Recall that $\delta, \epsilon < 0$ and $p_0, p_1, k_2 \in (0,1)$ and $k_1 > 0$.

For $\partial_\epsilon n_1$ we see:

$$\frac{\partial n_1}{\partial\epsilon} = \frac{2\gamma_C(1-p_0)}{4\delta} + \frac{2\gamma_C\left[\delta((p_0-1)^2+k_1(p_0+1))+\epsilon\gamma_C(p_0-1)^2\right]}{4\delta^2\sqrt{\frac{(\delta(1+k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}}} < 0 \tag{A.30a}$$

where $\dfrac{2\gamma_C(1-p_0)}{4\delta} < 0$ and

$2\gamma_C\left[\delta((p_0-1)^2+k_1(p_0+1))+\epsilon\gamma_C(p_0-1)^2\right] < 0$

which makes the overall expression less than 0 (A.30b)

We find for $\partial_{p_0} n_1$:

$$\frac{\partial n_1}{\partial p_0} = -\frac{(\delta+\epsilon\gamma_C)}{2\delta} + \frac{2(\delta+\epsilon\gamma_C)(\delta(k_1-1+p_0)+\epsilon\gamma_C(p_0-1))}{4\delta^2\sqrt{\frac{(\delta(1+k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}}} < 0 \tag{A.31a}$$

where $-\frac{(\delta+\epsilon\gamma_C)}{2\delta} < 0$ and

$$2(\delta+\epsilon\gamma_C)(\delta(k_1-1+p_0)+\epsilon\gamma_C(p_0-1)) = 2\overbrace{((p_0-1)(\delta+\epsilon\gamma_C)-\delta k_1)}^{+ve}\overbrace{(\delta+\epsilon\gamma_C)}^{-ve} < 0$$

which makes the overall expression less than 0 (A.31b)

We find for $\partial_{k_1} n_1$:

$$\frac{\partial n_1}{\partial k_1} = \frac{1}{2}\left[-1+\frac{2\delta[(\delta+\epsilon\gamma_C)p_0+\delta(1+k_1)+\epsilon\gamma_C]}{\delta^2\sqrt{\frac{(\delta(1+k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}}}\right] > 0 \tag{A.32a}$$

Since expression A.32a may be written as

$$\frac{1}{2}\left[-1+\frac{\delta(1+k_1+p_0)+\epsilon\gamma_C(1+p_0)}{\delta\sqrt{\frac{(\delta(1+k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}}}\right] \tag{A.33}$$

we simply consider the second term in the square brackets:

$$\frac{\frac{\delta(1+k_1+p_0)}{\delta}+\frac{\epsilon\gamma_C(1+p_0)}{\delta}}{\sqrt{\frac{(\delta(1+k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}}} \tag{A.34}$$

where

$$\overbrace{\frac{\delta(1+k_1+p_0)}{\delta}}^{+ve}+\overbrace{\frac{\epsilon\gamma_C(1+p_0)}{\delta}}^{+ve} > 0$$

Squaring both the numerator (Num^2) and denominator (Den^2) we see that

$$\begin{aligned}\text{Num}^2 =& \left[\frac{\delta(1+k_1+p_0)}{\delta}+\frac{\epsilon\gamma_C(1+p_0)}{\delta}\right]^2 \\ =& \frac{1}{\delta^2}\left[\underbrace{\delta^2(1+k_1+p_0)^2}_{A}+\underbrace{2\delta\epsilon\gamma_C(1+p_0)(1+k_1+p_0)}_{B}+\underbrace{\epsilon^2\gamma_C^2(1+p_0)}_{C}\right]\end{aligned} \tag{A.35a}$$

and

$$\begin{aligned}\text{Den}^2 =& \frac{(\delta(1+k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2} = \\ & \frac{1}{\delta^2}\left[\underbrace{\delta^2(1+k_1-p_0)^2}_{A'}+\underbrace{2\delta\epsilon\gamma_C((1+p_0)(1+k_1+p_0)-2(1+k_1)-4p_0)}_{B'}+\underbrace{\epsilon^2\gamma_C^2(1-p_0)}_{C'}\right]\end{aligned} \tag{A.36a}$$

Comparing the terms in lines A.35a and A.36a term by term it is obvious that $A > A', B > B', C > C'$ and therefore

$$\left[\frac{\delta(1+k_1+p_0)}{\delta}+\frac{\epsilon\gamma_C(1+p_0)}{\delta}\right]^2 > \frac{(\delta(1+k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}$$

meaning that $\dfrac{\delta(1+k_1+p_0)+\epsilon\gamma_C(1+p_0)}{\delta\sqrt{\frac{(\delta(1+k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}}} > 1$

Therefore, the expression A.32a must be greater than 0 meaning that the expression A.34 is necessarily positive.

$$\frac{\partial n_1}{\partial\gamma_C} = \frac{1}{2\delta^3}\Bigg[\delta^2(1-k_1-p_0+\epsilon(1-p_0))+\delta(\delta(k_1+p_0-1)+\epsilon\gamma_C(p_0-1))+ \frac{\epsilon\gamma_D\big[-\delta((p_0-1)^2+k_1(1+p_0))-\epsilon\gamma_C(p_0-1)^2\big]}{\sqrt{\frac{\delta(-1-k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}}}\Bigg] > 0 \tag{A.38a}$$

Looking at the first term in the brackets of A.38a we have

$$\delta^2(1-k_1-p_0+\epsilon(1-p_0))+\delta(\delta(k_1+p_0-1)+\epsilon\gamma_C(p_0-1)) =$$

$$\delta^2(1-k_1-p_0+\epsilon(1-p_0)+k_1+p_0-1)+\delta\epsilon\gamma_C(p_0-1) =$$

$$\delta^2\epsilon(1-p_0)-\delta\epsilon\gamma_C(1-p_0) = \overbrace{(1-p_0)}^{+ve}\overbrace{\delta(\delta\epsilon-\epsilon\gamma_C)}^{-ve} < 0$$

meaning that $\dfrac{\delta^2(1-k_1-p_0+\epsilon(1-p_0))+\delta(\delta(k_1+p_0-1)+\epsilon\gamma_C(p_0-1)}{2\delta^3} > 0$

Looking at the second term within the brackets of A.38a we have

$$\overbrace{\epsilon\gamma_D}^{-ve}\big[\overbrace{-\delta((p_0-1)^2+k_1(1+p_0))}^{+ve}-\overbrace{(\epsilon\gamma_C(p_0-1)^2)}^{-ve}\big] < 0$$

and since $2\delta^3\sqrt{\dfrac{\delta(-1-k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}} < 0$

we have that $\dfrac{\epsilon\gamma_D\big[-\delta((p_0-1)^2+k_1(1+p_0))-\epsilon\gamma_C(p_0-1)^2\big]}{2\delta^3\sqrt{\frac{\delta(-1-k_1)-\epsilon\gamma_C)^2-2(\delta+\epsilon\gamma_C)(\delta(1-k_1)+\epsilon\gamma_C)p_0+(\delta+\epsilon\gamma_C)^2p_0^2}{\delta^2}}} > 0$

which makes the overall expression of A.38a greater than 0

We find that n_2 changes with all parameters. The unwieldy nature of expressions do not enrich the content of this thesis, and we show only the signs of the partial derivatives. Upon request a written up pdf document is available to read further calculations, and (or) also a mathematica file is free upon request. To illustrate the unwieldy nature of these partials, the least complex to analyse is

$$\frac{\partial n_2}{\partial p_1} = \frac{1}{4\delta^2(1+\epsilon)^2(1+k_2)^2}\Bigg\{ -2\delta(1+\epsilon)(\delta+\epsilon\gamma_C)(1+k_2) - \frac{\delta(1+\epsilon)(1+k_2)}{\sqrt{(\delta+\epsilon\gamma_C)^2\Gamma+(\delta+\epsilon\gamma_C)\Psi-\Theta}} \times \frac{2(\delta+\epsilon\gamma_C)\big[\delta((1+k_2)(1-p_0+k_1(1+\epsilon))-p_1)+\epsilon\gamma_C((1+k_2)(1-p_0)-p_1)\big]}{\sqrt{(\delta+\epsilon\gamma_C)^2\Gamma+(\delta+\epsilon\gamma_C)\Psi-\Theta}}\Bigg\} \tag{A.39a}$$

Looking at the first term in the large curly brackets we find

$$\overbrace{(-2\delta)}^{+ve}\overbrace{(1+\epsilon)}^{+ve}\overbrace{(\delta+\epsilon\gamma_C)}^{-ve}\overbrace{(1+k_2)}^{+ve} < 0$$

Looking at the second term inside the curly brackets we find

$\overbrace{2(\delta+\epsilon\gamma_C)}^{-ve}\overbrace{\delta(1+\epsilon)(1+k_2)}^{-ve} > 0$ and inside the square brackets the term

$$\delta((1+k_2)(1-p_0+k_1(1+\epsilon))-p_1) = \delta\big[\overbrace{k_1(\epsilon(1+k_2)+(1+k_2))}^{+ve}+\overbrace{(1-p_0-p_1)}^{+ve}+\overbrace{k_2(1-p_0)}^{+ve}\big]$$
$$< 0$$

and $\epsilon\gamma_C((1+k_2)(1-p_0)-p_1) = \overbrace{\epsilon\gamma_C}^{-ve}(\overbrace{(1-p_0-p_1)}^{+ve}+\overbrace{k_2(1-p_0)}^{+ve}) < 0$

Therefore the whole expression becomes

$$\overbrace{\frac{1}{4\delta^2(1+\epsilon)^2(1+k_2)^2}}^{+ve}\Bigg\{\overbrace{-2\delta(1+\epsilon)(\delta+\epsilon\gamma_C)(1+k_2)}^{-ve} - \frac{\overbrace{2(\delta+\epsilon\gamma_C)\delta(1+\epsilon)(1+k_2)}^{+ve}\Big[\big[\overbrace{\delta((1+k_2)(1-p_0+k_1(1+\epsilon))-p_1)}^{-ve}+\overbrace{\epsilon\gamma_C((1+k_2)(1-p_0)-p_1)}^{-ve}\big]\Big]}{\sqrt{(\delta+\epsilon\gamma_C)^2\Gamma+(\delta+\epsilon\gamma_C)\Psi-\Theta}}\Bigg\}$$
$$< 0$$

Seek the accompanying files for details of the following partial derivatives

$$\frac{\partial n_2}{\partial p_0} < 0, \frac{\partial n_2}{\partial k_1} > 0, \frac{\partial n_2}{\partial \epsilon} < 0, \frac{\partial n_2}{\partial k_2} > 0, \frac{\partial n_2}{\partial \gamma_D} < 0 \tag{A.41}$$

Stability for $\delta, \epsilon < 0$

The fixed point x_3 is never stable when $\delta, \epsilon < 0$.

We find that x_4 is stable if and only if $n_3 < n < n_1$, or $n_2 < n < n_1$ depending on our parameter values.

A.4 Considering a multiplicative evasion function: $\frac{p_0 n}{k_1+n}(1+\frac{p_1 x}{k_2+x})$

By applying this evasion function, the dynamical system becomes

$$\dot{x} = x(1-x)\Bigg\{\frac{\gamma_C\Big(n+\epsilon\big(\frac{np_0p_1x}{(k_1+n)(k_2+x)}+\frac{np_0}{k_1+n}+nx-1\big)+\frac{np_0p_1x}{(k_1+n)(k_2+x)}+\frac{np_0}{k_1+n}-1\Big)}{n(1+\epsilon x)}+ \\ \frac{\gamma_D\big(\gamma_D-(\epsilon nx+\frac{np_0p_1x}{(k_1+n)(\mathrm{k2}+x)}+\frac{np_0}{k_1+n}+n)\big)}{n(1+\epsilon x)}\Bigg\} \tag{A.42a}$$

which we solve to find the quadratic solutions x_3 and x_4;

$$x_3 = \frac{\gamma_C(k_1+n)(e(k_2n-1)+n-1)-\gamma_C(\epsilon+1)np_0p_1+\gamma_D(k_1+n)(ek_2n+n-1)+\gamma_D np_0p_1}{2\epsilon n\delta(k_1+n)}+ \\ \frac{\sqrt{((k_1+n)(-n\delta(ek_2+1)+\gamma_C\epsilon+\delta)-np_0p_1(\gamma_C\epsilon+\delta))^2+\mathcal{K}}}{2\epsilon n\delta(k_1+n)}$$

and

$$x_4 = \frac{\gamma_C(k_1+n)(e(k_2n-1)+n-1)-\gamma_C(\epsilon+1)np_0p_1+\gamma_D(k_1+n)(\epsilon k_2n+n-1)+\gamma_D np_0p_1}{2\epsilon n\delta(k_1+n)}- \\ \frac{\sqrt{((k_1+n)(-n\delta(ek_2+1)+\gamma_C\epsilon+\delta)-np_0p_1(\gamma_C\epsilon+\delta))^2+\mathcal{K}}}{2\epsilon n\delta(k_1+n)} \tag{A.44a}$$

where

$$\mathcal{K} = 4ek_2n\delta(k_1+n)^2(\gamma_C(e-n+1)+\gamma_D(n-1)). \tag{A.45}$$

The critical group sizes n_1 and n_2 become;

$$n_1 = \frac{1}{2}\Bigg(+\frac{\gamma_D(k_1+p_0-1)-\gamma_C(\epsilon(p_0-1)+k_1+p_0-1)}{\delta}+ \\ \sqrt{\frac{-2p_0(\gamma_C\epsilon+\delta)(\gamma_C(\epsilon-k_1+1)+\gamma_D(k_1-1))+\mathcal{H}+p_0^2(\gamma_C\epsilon+\delta)^2}{\delta^2}}\Bigg)$$

and

$$n_2 = \frac{1}{2}\left(a_0-\frac{1}{\sqrt{a_1-a_2+a_3}}\right) \tag{A.47}$$

where

$$\mathcal{H} = (-\gamma_C(\epsilon+k_1+1)+\gamma_Dk_1+\gamma_D)^2 \tag{A.48}$$

and

$$a_0 = \frac{\gamma_D((\epsilon+1)k_1(k_2+1)+k_2(p_0-1)+p_0p_1+p_0-1)}{(\epsilon+1)(k_2+1)\delta}$$

$$a_1 = \frac{\gamma_C^2(\epsilon+1)^2(k_1^2(k_2+1)^2+2k_1(k_2+1)(k_2p_0+k_2+p_0p_1+p_0+1)+(k_2(p_0-1)+p_0p_1+p_0-1)^2)}{(\epsilon+1)^2(k_2+1)^2\delta^2}$$

$$a_2 = \frac{2\gamma_C\gamma_D(\epsilon+1)}{(\epsilon+1)^2(k_2+1)^2\delta^2}\left(\frac{(\epsilon+1)k_1^2(k_2+1)^2+(\epsilon+2)k_1(k_2+1)(k_2p_0+k_2+p_0p_1+p_0+1)}{(\epsilon+1)^2(k_2+1)^2\delta^2}+\frac{(k_2(p_0-1)+p_0p_1+p_0-1)^2}{(\epsilon+1)^2(k_2+1)^2\delta^2}\right)$$

$$a_3 = \frac{\gamma_D^2\Big((\epsilon+1)^2k_1^2(k_2+1)^2+2(\epsilon+1)k_1(k_2+1)(k_2p_0+k_2+p_0p_1+p_0+1)}{(\epsilon+1)^2(k_2+1)^2\delta^2}+\frac{(k_2(p_0-1)+p_0p_1+p_0-1)^2\Big)}{(\epsilon+1)^2(k_2+1)^2\delta^2}$$

The value n_3 is not written algebraically here due to it's length and complexity, but can be found in a separate Mathematica file upon request. Due to the length of these expressions we do not give all of calculations of the partial derivatives (but are available on request), however, we do show a structural equivalence to the other models. As can be seen in that n_1 clearly does not depend on p_1, k_2 and thus selfish groups are unaffected by the cooperative aspect of evasion. Whereas n_2 clearly does have a dependence on p_1, k_2 and is modulated by cooperative behaviour in the evasion function.

$$\begin{aligned}&\frac{\partial n_1}{\partial p_0}<0\ \ \frac{\partial n_1}{\partial k_1}>0\ \ \frac{\partial n_1}{\partial \epsilon}<0\ \ \frac{\partial n_1}{\partial \gamma_D}<0\\ &\frac{\partial n_2}{\partial p_0}<0\ \ \frac{\partial n_2}{\partial p_1}<0\ \ \frac{\partial n_2}{\partial k_1}>0\ \ \frac{\partial n_2}{\partial \epsilon}<0\ \ \frac{\partial n_2}{\partial k_2}>0\ \ \frac{\partial n_2}{\partial \gamma_D}<0\end{aligned} \tag{A.50a}$$

A.4.1 Stability for $\delta, \epsilon < 0$

The fixed point x_3 is never stable when $\delta, \epsilon < 0$.

We find that x_4 is stable if and only if $n_3 < n < n_1$, or $n_2 < n < n_1$ depending on our parameter values.

Appendix B

B.1 Parameterising the evasion fucnction $g(n, x)$

We make a series of assumptions supported by empirical data in order to parametrise the predator evasion function used in equation 3.6. First, the effect of group size on the probability of predator detection is saturating in redshank, and plateaus for flocks of around 20 birds [25]. To capture this we choose $k_1 = 10$ so that as a flock reaches a size of 20 birds, there are little increases in the probability of evasion with further increases in flock size.

We apply the critical group sizes of n_1 and n_2 given in section 2.3.1 chapter 2, as functions of model parameters. Using these critical group sizes we use the model predicted vigilance composition of a flock and infer key parameters using previous research on redshank attacks. The theoretical model predicts that a flock of size than n_2 will be entirely vigilant, and flocks of size greater than n_1 will be entirely non-vigilant. As $n_1 = n_1(d)$ is a function of distance, and $n_2 = n_2(d, n)$ is a function of distance and flock size, for any distance to cover, we can predict the vigilant state of the flock.

The model results predict that for $d \to 40, n \to 50$, that $\max\{n_1(40)\} = 10$ and that $\max\{n_2(40, 50)\} = 7$. We assume that when $d \to 40, n \to 50$ (since a flock of size 50 is greater than $n_1(40) = 10$), there will be no vigilant birds ($x = 0$). In this case, the probability of evasion for a theoretical non-vigilant flock of size 50 birds at 40 meters from cover, is 0.93 [207]. Therefore we use Equation 3.6 with $x = 0, n = 50$ and $d = 40$ such that

$$\frac{A_1 50}{10 + 50} = 0.93 \tag{B.1}$$

suggesting that $A_1 = 1.1$.

Second, we assume that if $d \to 40, n \to 50$ and $x = 1$ then the probability of evasion is near to 1. This follows from observed studies in which an exponential fit to attack data gives a flock of size 50 birds at 40 meters from cover and our assumption that vigilant flocks have a greater chance of evading the predator than non-vigilant flocks. We make the simplifying assumption that

$$\frac{A_1 n}{k_1 + n} + \frac{A_2(n)}{k_2 + 1} = 1 \tag{B.2}$$

We take $A_2(n) = (1 - \Lambda(n))(k_2 + 1)$ where $\Lambda(n)$ monotonically decreases with n. This captures the observation that as group size increases confusion and group size effects become more influential in evasion meaning that additional anti-predation gains from the frequency of vigilant individuals decrease [18, 151]. That is, as n becomes larger the significance of the effects on evasion from $p_1(n)$

decrease. A way in which to model this without over-complicating the parameter fit is to apply

$$\Lambda(n) = A_1(\frac{n}{\kappa_1 + n\kappa_2}) \tag{B.3}$$

meaning that $A_2(n) = (1 - A_1(\frac{n}{\kappa_1+n\kappa_2}))(k_2 + 1)$.

Third, we assume that if birds are sufficiently away from cover (i.e. over 30 meters) [25, 207] then the chance of evasion plateaus. In this scenario, because $\min\{n_2(30, 1)\} = 8$ we assume that a solitary bird is vigilant with probability one ($x = 1$). At a sufficient distance from cover we expect the probability of evasion for a solitary vigilant bird be [207]

$$\frac{A_1}{11} + \left(1 - A_1\left[\frac{1}{\kappa_1 + \kappa_2}\right]\right) = 0.6 \tag{B.4}$$

such that $\kappa_1 + \kappa_2 = 2.2$

Fourth, we assume that when $d \to 0$, $n = 1$ and then any individual is vigilant with probability one ($x = 1$) because of such close proximity to cover. To this end, we expect the probability of evasion for a solitary vigilant bird at cover to reflect previous research [207]: we propose a probability of evasion of 0.46 should reflect a solitary vigilant bird at cover. As such

$$\frac{A_1(1 - \lambda_1)}{11} + \left(1 - A_1\left[\frac{1}{\kappa_1 + \kappa_2}\right]\right))(1 - \lambda_2) = 0.46 \tag{B.5}$$

Fifth, as $d \to 0$ and $n \to 1$ if a solitary bird is non-vigilant, then it's only chance of survival is by a failed attack (possibly due to confusion effects). Non-vigilant birds are on average 0.07 seconds slower to react than vigilant birds. Given that sparrowhawks fly at a speed of $25m/s$ we find that this translates to allow a predator about 1.75m closer to non-vigilant redshank as compared to vigilant birds [255]. Vigilant birds have a 1.75m distance advantage over non-vigilant ones, which between 0 and 1m to cover, we assume can make non-vigilant birds up to 7 times more likely to be targeted. Therefore, we assume that solitary non-vigilant birds approximately have a $\frac{0.46}{7} \approx 0.07$ probability of evasion at cover. Therefore when $n \to 1, d \to 0, x \to 0$ we have

$$\frac{1.1(1 - \lambda_1)}{11} = 0.07 \tag{B.6}$$

meaning $\lambda_1 = 0.30$.

We obtain λ_2 by substituting back into Equation B.5. As the probability of evasion for a solitary vigilant redshank at cover should match observed empirical data on attack success rates, we find

$$\frac{1.1(1 - 0.3)}{11} + (1 - 1.1(\frac{1}{2.2}))(1 - \lambda_2) = 0.46 \tag{B.7}$$

again using data and functional forms from [207]. It follows that $\lambda_2 = 0.22$

To fit κ_1 and κ_2 we examine the changing influence that vigilance has upon evasion as flock size increases. This approach is inferred via observations, but the precise nature of how the frequency of vigilant birds in a flock affects evasion as flock size varies will be somewhat ambiguous. We use the fact that at cover if a solitary bird is not vigilant then it has a $\frac{1.1(1-0.3)}{11} = 0.07$ probability of escape. If the bird is vigilant this increases to 0.46. Using $n_2(1, n)$ from the critical group sizes

of section 2.3.1 we find $\min\{n_2(1,20)\} = 25$. Thus, for a flock of size $n = 25 < n_2$ we predict all birds to be vigilant with probability one, when the flock is located at predatory concealing cover. Using observational attack success data we assume a vigilant flock of size 25 birds a 0.65 chance of evasion at cover. Conversely, if the flock is non-vigilant, because detection asymptotes with flocks of size 20, we only expect the flock to have a probability 0.63 of escape. Therefore we require a mapping such that

$$\Lambda(1) = 1.1(\frac{1}{\kappa_1 + 1\kappa_2})(1 - 0.22) = 0.46 - \frac{1.1(1-0.3)}{11} \tag{B.8a}$$

$$\Lambda(50) = 1.1(\frac{25}{\kappa_1 + 25\kappa_2})(1 - 0.22) = 0.65 - \frac{1.1(1-0.3)}{36} \tag{B.8b}$$

$$\iff \tag{B.8c}$$

$$0.45 = \frac{1}{\kappa_1 + 1\kappa_2} \tag{B.8d}$$

$$0.63 = \frac{25}{\kappa_1 + 25\kappa_2} \tag{B.8e}$$

$$\iff \tag{B.8f}$$

$$\kappa_2 = 1.56 \approx 1.6 \tag{B.8g}$$

It follows that $\kappa_1 = 2$ and $\kappa_2 = 1.6$.

Finally to fit the parameter k_2 we backward infer the response of an expected mixed vigilance flock for a given attack. The mixed population state x_4 is stable and physical for $n \in [n_2, n_1]$, which occurs for certain values of d and n. We find that $\min\{n_1(10)\} = 32$ and that $\max\{n_2(10,25)\} = 18$ for $\bar{d} = 10$ and $\bar{n} = 25$. In this case, our model predicts the flock to have an expected composition defined by x_4. Using the parameter values of $\bar{n} = 25, \bar{d} = 10$ we apply x_4 to the evasion function such that

$$P(\bar{d}, \bar{n}, x) = \frac{A_1(1 - \lambda_1 e^{\alpha_1 \bar{d}})\bar{n}}{k_1 + \bar{n}} + \frac{A_2(n)(1 - \lambda_2 e^{\alpha_2 \bar{d}})x_4^*}{k_2 + x_4^*} = 0.72 \tag{B.9}$$

where 0.72 shows an estimate of the probability of evasion found in observed studies for a flock of size 25 at 10 meters to cover [207]. Our evasion function should reflect the actual response for a flock and therefore reflect observational data. Using x_4 as a function of k_2 we apply $P(10, 15, x_4(k_2)^*) = 0.72$ [207] to obtain $k_2 = 0.24$.

B.2 The Beta regression model

The beta regression is based on a parametrisation of the beta density using

$$f(Y|a,b) = \frac{Y^{(a-1)}(1-Y)^{(b-1)}}{\mathcal{B}(a,b)} \quad , \quad 0 < Y < 1 \tag{B.10}$$

where the beta function is defined by $\mathcal{B}(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$, $\Gamma(.)$ is the gamma function, and a and b are shape parameters that enter the model according to the transformations:

$$a = u \cdot \phi \text{ and } b = (1-u) \cdot \phi$$

Using g_{L1} as a link function for u_i, the specification of the shape parameters becomes, $E[Y] = u_i = g_{L1}^{-1}(\mathbf{X}\beta)$ where $\mathbf{X}$ are the predictor variables and β are coefficients associated with each predictor. One also has that $\text{var}[Y] = u(1-u)/(1+\phi)$. Here, ϕ is the precision parameter; for fixed u, the larger ϕ the smaller the variance of Y, so ϕ^{-1} is a dispersion parameter.

B.3 Negative Binomial regression

The negative binomial distribution can be derived as a Poisson distribution with gamma distributed noise variable, where the dependent follows a Poisson distribution: $Y \sim \text{Poisson}(\theta)$. To accommodate for overdispersion one assumes that $\theta \sim \text{gamma}(r, p/(1-p)$ where r and p are shape parameters. The distribution function is defined by

$$f(Y|r,p) = \frac{\Gamma(Y+r)}{Y!\Gamma(r)} \cdot (1-p)^r p^Y = \text{NB}(r,p) \tag{B.11}$$

with the mean and shape parameters given by the parametrisation:

$$r = \phi \text{ and } u = \frac{\phi(1-p)}{p} \tag{B.12}$$

such that $E[Y] = u$ and $\text{Var}[Y] = u(1+u/\phi)$.

Appendix C

C.1 A comparison of selection gradients

In a monomorphic population with strategy z, the selection gradient for the contribution probability strategy (equation 4.10 in text) is

$$\delta g(z) + \delta(1 - g(z))\left(\frac{-\epsilon}{n(1+\epsilon z)}\right). \tag{C.1}$$

The selection gradient for the contribution investment depends on cases of $g(y,z)$ and $\gamma(y)$. If $g(y,z)$ is independent of y (i.e., a constant or function of the resident strategy alone), then the selection gradient for the investment strategy is

$$\gamma'(y)\big(g(z) + (1-g(z))\mathcal{D}(y;z;1)\big) + \gamma(y,z)(1-g(z))\frac{\partial}{\partial y}\mathcal{D}(y;z;1). \tag{C.2}$$

When the novel strategy affects the probability of group evasion we find the selection gradient for the contribution investment is

$$\gamma'(y)\big(g(z) + (1-g(z))\mathcal{D}(y;z;1)\big) \tag{C.3a}$$

$$+\gamma(y)\big(\frac{\partial}{\partial y}\mathcal{D}(y;z;1)(1-g(y,z)) + (1-\mathcal{D}(y;z;1))\frac{\partial}{\partial y}g(y,z)\big) \tag{C.3b}$$

For comparison of the above selection gradients, recall that $\mathcal{D}(y;z;1) = -\epsilon(y)/(n(1+\epsilon(z))$ in equations (C.2-C.3).

C.2 General Model analysis: constant evasion and linear γ and ϵ

With the G-function payoff of equation 4.4a, (asymptotic) stability of the ecological dynamics is given by the sign of

$$\begin{aligned} f'(x_1) = (1-x_1)\frac{\partial}{\partial x_1}\Bigg(& x_1\big(\gamma(z_1)-\gamma(z_2)\big) \\ & - \frac{x_1(1-p_0)}{n(1+x_1\epsilon(z_1)+(1-x_1)\epsilon(z_2))}\Big(\big(\gamma(z_1)-\gamma(z_2)\big)+\epsilon(z_1)\gamma(z_1)-\epsilon(z_2)\gamma(z_2)\Big)\Bigg) \\ & + \frac{\partial}{\partial x_1}(1-x_1)\Bigg(x_1\big(\gamma(z_1)-\gamma(z_2)\big) \\ & - \frac{x_1(1-p_0)}{n(1+x_1\epsilon(z_1)+(1-x_1)\epsilon(z_2))}\Big(\big(\gamma(z_1)-\gamma(z_2)\big)+\epsilon(z_1)\gamma(z_1)-\epsilon(z_2)\gamma(z_2)\Big)\Bigg). \end{aligned} \tag{C.4a}$$

If equation C.4a is negative the population-dynamic fixed point x_1 is stable.

The monomorphic invasion fitness given by equation (4.7) is positive if

$$\big(\gamma(z_1)-\gamma(y)\big) - \frac{(1-p_0)}{n(1+\epsilon(z_1))}\Big\{\big(\gamma(z_1)-\gamma(y)\big)+\epsilon(z_1)\gamma(z_1)-\epsilon(y)\gamma(y)\Big\} < 0. \tag{C.5}$$

We find that the requirement for evolutionary stability of a singular strategy $\hat{z}_x$ to be

$$\epsilon''(z) > \frac{\gamma''(z)(1+\epsilon(z))(n-1+p_0) - 2\epsilon'(z)\gamma'(z)(1-p_0)}{\gamma(z)(1-p_0)} \tag{C.6}$$

and convergence stability requires that

$$\epsilon''(z) > \epsilon'(z)\left(\frac{\epsilon'(z)}{1+\epsilon(z)} - \frac{\gamma'(z)}{\gamma(z)}\right) + \frac{\gamma''(z)(1+\epsilon(z))(n-1+p_0)}{2\gamma(z)(1-p_0)} \tag{C.7}$$

when$z = \hat{z}_x$

C.2.1 Model 1 analysis

The physicality of the candidate singular strategy is conditional on values of the group size n. For $1 > z_{x_1} > 0$ either of the following cases must be satisfied:

$$\begin{aligned} &\text{Case A}: \text{ If } \epsilon < \frac{\delta}{\gamma_D} \text{ then either } n < \frac{(\delta+\epsilon\gamma_D)(1-p_0)}{\delta} := n_1 \text{ or } n > \frac{(1-p_0)(\delta(2\epsilon-1)-\epsilon\gamma_D)}{\delta(\epsilon-1)} := n_2 \\ &\text{Case B}: \text{ If } \epsilon > \frac{\delta}{\gamma_D} \text{ then either } n < n_2 \text{ or } n > n_1 \end{aligned} \tag{C.8a}$$

depending on the value of p_0.

C.2.2 Model 2 analysis

Ths singular solutions are given by

$$\begin{aligned}\hat{z}_x^+ =& \frac{\epsilon k\big(\delta n - 2(\delta-\gamma_D)(1-p_0)\big)}{2\epsilon(\delta-\gamma_D)(1-p_0)} + \\ & \frac{\sqrt{\delta\epsilon k\big(4(\delta-\gamma_D)(p_0-1)^2(\epsilon k-1)+4(\delta-\gamma_D)(p_0-1)(\epsilon k-1)n+\delta\epsilon k n^2\big)}}{2\epsilon(\delta-\gamma_D)(1-p_0)} \\ \hat{z}_x^- =& \frac{\epsilon k\big(\delta n - 2(\delta-\gamma_D)(1-p_0)\big)}{2\epsilon(\delta-\gamma_D)(1-p_0)} - \\ & \frac{\sqrt{\delta\epsilon k\big(4(\delta-\gamma_D)(p_0-1)^2(\epsilon k-1)+4(\delta-\gamma_D)(p_0-1)(\epsilon k-1)n+\delta\epsilon k n^2\big)}}{2\epsilon(\delta-\gamma_D)(1-p_0)}\end{aligned}$$

The physicality of the solutions depend on two critical group sizes:

$$\begin{aligned}n_1 &= \frac{(1-p_0)(\delta+\epsilon\gamma_D k)}{\delta} \\ n_2 &= \frac{(1-p_0)(\delta(\epsilon-k+2\epsilon k)-\epsilon\gamma_D(1+k)^2)}{\delta(\epsilon-1)k}\end{aligned} \tag{C.10a}$$

If $0<\delta<\gamma_D$ then for the singular solution to be physical (in $(0,1)$), the condition of $2<n<n_2$ must be satisfied. Whereas when $(1+k)\gamma_D>\delta>\gamma_D$ the condition of $n_1<n<n_2$ is required.

The ESS condition (C.6) becomes

$$\begin{aligned}&\frac{2\delta k}{(k+y)^3} < \frac{2(1-p_0)\delta k\epsilon}{(k+y)^2(1-\epsilon y)(p_0+n-1)} && \iff \\ &\frac{1}{k+y} < \frac{\epsilon(1-p_0)}{(1-\epsilon y)(p_0+n-1)} \iff y > \frac{n+p_0(1+\epsilon k)-\epsilon k-1}{n\epsilon}\end{aligned}$$

Hence, when physical, a singular strategy $\hat{z}_x$ is evolutionary stable if and only if

$$\frac{n+p_0(1+\epsilon k)-\epsilon k-1}{n\epsilon} < \hat{z}_x < 1 \tag{C.12}$$

Convergence stablility is determined condition (C.7), which is checked directly using

$$\gamma'(y) < \frac{(1-p_0)\gamma(y)\epsilon'(y)^2-(1-p_0-n)\big(1+\epsilon(y)\big)^2\gamma''(y)}{(1-p_0)\big(1+\epsilon(y)\big)\epsilon'(y)} \tag{C.13}$$

at $y=\hat{z}_x$.

C.2.3 Model 3 analysis

We use the Sturm sequence $q_0(z_1), q_1(z_1), q_2(z_1)$, where $q_0(z_1) := q(z_1), q_1(z_1) := q'(z_1)$ and $q_i(z_1) := -q_{i-2}(z_1)/q_{i-1}(z_1)$ to deduce the number of roots of (4.27) in the open unit interval. One finds

$$q_1(z_1) = \delta n(2k(n-1+p_0) + \epsilon k(1+k)(n-2+2p_0) + 2(1+\epsilon+\epsilon k)(n-1+p_0)z_1) \tag{C.14a}$$

$$q_2(z_1) = \frac{1}{4}\epsilon k(k+1)n\left(4\gamma_D(1-p_0) + \delta k\left(\frac{\epsilon(k+1)n^2}{(\epsilon k+\epsilon+1)(n+p_0-1)} + 4p_0\right) - 4\delta k\right) \tag{C.14b}$$

$$q_3(z_1) = 0. \tag{C.14c}$$

We denote the sign sequence $\Sigma_l = \{q_i(l)\}$ for $i = 1,2,3$ and $l = 0,1$. Let α_l be the number of sign changes of Σ_l. Direct calculation show that for all possible parameter regimes we find $\max\{\alpha_l\} = 1, l = 0,1$. The number of roots between zero and one is given by $\alpha_0 - \alpha_1$. Hence, there can be only one singular strategy physical.

Singular strategies are the roots of equation (4.27):

$$\begin{aligned}
\hat{z}_x^+ =& \frac{\delta k\big(2(p_0+n-1) - \epsilon(2p_0+n-2)\big)}{2\delta(\epsilon-1)(p_0+n-1)} + \\
& \frac{\sqrt{\delta\epsilon k\big(4(\epsilon-1)(p_0-1)n(\delta k+\gamma_D) + 4(\epsilon-1)(p_0-1)^2(\delta k+\gamma_D) + \delta\epsilon k n^2\big)}}{2\delta(\epsilon-1)(p_0+n-1)} \\
\hat{z}_x^- =& \frac{\delta k\big(2(p_0+n-1) - \epsilon(2p_0+n-2)\big)}{2\delta(\epsilon-1)(p_0+n-1)} - \\
& \frac{\sqrt{\delta\epsilon k\big(4(\epsilon-1)(p_0-1)n(\delta k+\gamma_D) + 4(\epsilon-1)(p_0-1)^2(\delta k+\gamma_D) + \delta\epsilon k n^2\big)}}{2\delta(\epsilon-1)(p_0+n-1)}.
\end{aligned}$$

Direct calculation shows that $\hat{z}_x^-$ is always negative; analysis in text shows that $\hat{z}_x^+$ is the only possible solution in $(0,1)$.

Two values of n condition the singular strategies physicality. The singular strategy is positive if $n < n_1$, but is only lies between zero and one when $n \in (\max\{2, n_2\}, n_1)$ where n_1 and n_2 are given by

$$\begin{aligned}
n_1 &= \frac{(1-p_0)(\delta k + \epsilon\gamma_D)}{\delta k} \\
n_2 &= \frac{(p_0-1)(\gamma_D k - \epsilon + \delta(\epsilon + 2\epsilon k - (1+k)^2)}{\delta(1+k-\epsilon)(1+k)}
\end{aligned} \tag{C.16a}$$

C.2.4 Model 5 analysis

Singular strategies are the roots of

$$\mathcal{C}_5 = w_0 + w_1 z_1 + w_2 z_1^2 + w_3 + z_1^3 + w_4 z_1^4 : \tag{C.17a}$$

$$w_0 = k_0 k_1 \Bigg(- \gamma_C (1 + k_0) k_1 n (n - 1 + p_0) + \tag{C.17b}$$

$$\gamma_D k_1 n \big(n - 1 + p_0 + k_0(\epsilon + n - 1 + p_0 - \epsilon p_0)\big) - \gamma_D k_0 p_1 \Bigg) \tag{C.17c}$$

$$w_1 = k_0 k_1 \Bigg(\gamma_D \Bigg((k_0 + 1) n^2 (\epsilon k_1 + 2) + 2n(p_0 - 1)(\epsilon k_0 (k_1 - 1) + k_0 + 1) + \tag{C.17d}$$

$$n p_1 (-\epsilon k_0 + k_0 + 1) + p_1 (-\epsilon k_0 + k_0 - 1) \Bigg) - \tag{C.17e}$$

$$\gamma_C (k_0 + 1)(n(\epsilon k_1 (n + 2p_0 - 2) + 2(n + p_0 - 1) + p_1) + p_1) \Bigg) \tag{C.17f}$$

$$w_2 = \gamma_D k_0 \Bigg(n(p_0 - 1)\left(\epsilon k_0 (4k_1 - 1) + \epsilon k_1^2 + k_0 + 1\right) + (k_0 + 1) n^2 (2\epsilon k_1 + 1) + \tag{C.17g}$$

$$n p_1 (\epsilon k_0 (2k_1 - 1) + k_0 + 1) + k_1 p_1 (\epsilon (k_0 - 1) + 1) \Bigg) - \tag{C.17h}$$

$$\gamma_C (k_0 + 1)(k_0 n (2\epsilon k_1 (n + 2p_0 - 2) + \tag{C.17i}$$

$$n + p_0 - 1) + k_0 p_1 (e(2k_1 n + k_1) + n) + k_1 (\epsilon k_1 n (p_0 - 1) + p_1)) \tag{C.17j}$$

$$w_3 = \epsilon \Bigg(- \gamma_C (1 + k_0) \Bigg(k_0 n (n + 2(p_0 + p_1 - 1)) + k_1 (p_1 + n(2p_0 + p_1 - 2)) \Bigg) + \tag{C.17k}$$

$$\gamma_D k_0 (n^2 + k_0 n (n + 2(p_0 + p_1 - 1)) + k_1 (p_1 + n(2p_0 + p_1 - 2))) \Bigg) \tag{C.17l}$$

$$w_4 = \epsilon (\gamma_D k_0 - \gamma_C (1 + k_0)) n (p_0 + p_1 - 1) \tag{C.17m}$$

Bibliography

[1] W. Cresswell and D. Whitfield. "How starvation risk in Redshanks Tringa totanus results in predation mortality from Sparrowhawks Accipiter nisus." *Ibis.* 150 (2008), pp. 209–218.

[2] S. Lima and L. Dill. "Behavioral Decisions Made under the Risk of Predation: A Review and Prospectus". *Can. J. Zool.* 68 (Apr. 1990), pp. 619–640. DOI: 10.1139/z90-092.

[3] T. Day, P. A. Abrams, and J. M. Chase. "The role of size-specific predation in the evolution and diversification of prey life histories". *Evolution* 56 (2002).

[4] J. Krause and G. Ruxton. *Living In Groups.* Oxford, U.K: Oxford University Press, 2002.

[5] R. F. Storms, C. Carere, and F. Zoratto. "Complex patterns of collective escape in starling flocks under predation". *Behav. Ecol. Sociobiol.* 73.10 (2019). DOI: https://doi.org/10.1007/s00265-018-2609-0.

[6] A. J. W. Ward, J. E. Herbert-Read, D. J. T. Sumpter, and J. Krause. "Fast and accurate decisions through collective vigilance in fish shoals". *Proc. Natl. Acad. Sci.* 108.6 (2011), pp. 2312–2315.

[7] P. A. Bednekoff. "Mutualism among safe, selfish sentinels: a dynamic game". *Am. Nat.* 150.3 (1997), pp. 373–392. DOI: 10.1086/286070..

[8] T. H. Clutton-Brock, M. J. O'Riain, P. N. M. Brotherton, D. Gaynor, R. Kansky, A. S. Griffin, and M. Manser. "Selfish sentinels in cooperative mammals". *Science* 284 (1999), pp. 1640–1644.

[9] D. T. Blumstein. "The Evolution, Function, and Meaning of Marmot Alarm Communication". *Adv. Study Behav.* 37 (2007), pp. 371–401.

[10] D. J. T. Sumpter. *Collective Animal Behaviour.* Princeton University Press, 2010.

[11] C. C. Ioannou, V. Guttal, and I. Couzin. "Predatory fish select for coordinated collective motion in virtual prey". *Science* 337 (2012), pp. 1212–1215.

[12] R. A Fisher. *The genetical theory of natural selection.* Oxford, Uk: Oxford Clarendon Press, 1930, p. 302. URL: https://www.biodiversitylibrary.org/item/69976.

[13] S. Wright. "Evolution In Mendelian Populations". *Genetics* 16.2 (1931), pp. 97–159. ISSN: 0016-6731. eprint: https://www.genetics.org/content/16/2/97.full.pdf. URL: https://www.genetics.org/content/16/2/97.

[14] J. B. S. Haldane. *The causes of evolution.* London: Longmans, 1932.

[15] G. Roberts. "A real-time response of vigilance to changes in group size". *Anim. Behav.* 50.5 (1994), pp. 1371–1374.

[16] S. L. Lima. "Distance to cover, visual obstructions, and vigilance in house sparrows". *Behaviour* 102.3 (1987), pp. 231–237.

[17] E. D. Brodie (Jr), D. R. Formanowicz, and E. D. Brodie III. "Predator avoidance and antipredator mechanisms: distinct pathways to survival". *Ethol. Ecol. Evol.* 3 (1991), pp. 73–77.

[18] W. Cresswell. "Flocking is an effective anti-predation strategy in redshanks, Tringa totanus." *Anim. Behav.* 47.2 (1994), pp. 433–442.

[19] H. A. Orr. "Fitness and its role in evolutionary genetics". *Nat. Rev. Genet.* 10.8 (2009), pp. 531–539. DOI: 10.1038/nrg2603.

[20] R. Barrett and H. Hoekstra. "Molecular spandrels: Tests of adaptation at the genetic level". *Nat. Rev. Genet.* 12 (Nov. 2011), pp. 767–80. DOI: 10.1038/nrg3015.

[21] T. W. Fawcett, S. Hamblin, and L-A Giraldeau. "Exposing the behavioural gambit: the evolution of learning and decision rules". *Behav. Ecol.* 24 (2012), pp. 2–11.

[22] B. Adriaenssens and J. I. Johnsson. "Natural selection, plasticity and the emergence of a behavoural syndrome in the wild". *Ecol. Lett.* (2013), pp. 47–55.

[23] W. Cresswell and J. L. Quinn. "Contrasting risks from different predators change the overall nonlethal effects of predation risk". *Behav. Ecol.* 24 (2013), pp. 871–876.

[24] F. Zoratto, M. Leonardo, O. Ludovica, P. Rianne, M. Eens, D. Santucci, E. Alleva, and C. Carere. "Behavioural response of European starlings exposed to video playback of conspecific flocks: Effect of social context and predator threat." *Behav. Process.* 103 (2014), pp. 269–277.

[25] W. Cresswell and J. Quinn. "Predicting the optimal prey group size from predator hunting behaviour". *J. Anim. Ecol.* 80.2 (2011), pp. 310–319.

[26] C. L. Devereux, M. J. Whittingham, E. Fernandez-Juricic, A. Vickery, and J. R. Krebs. "Predator detection and avoidance by starlings under differing scenarios of predation risk." *Behav. Ecol.* 17.2 (2005), pp. 302–309.

[27] M. David, M. Salignon, and M-J. Perrot-Minnot. "Shaping the antipredator strategy: flexibility, consistency, and behaviorual correlations under varying predation threat". *Behav. Ecol.* 25.5 (2014), pp. 1148–1156.

[28] T. J. Dewitt, A. Sih, and J. A. Hucko. "Trait compensation and cospecialization in a freshwater snail: size, shape and antipredator behaviour". *Anim. Behav.* 58.2 (1999), pp. 397–407. ISSN: 0003-3472. DOI: https://doi.org/10.1006/anbe.1999.1158. URL: http://www.sciencedirect.com/science/article/pii/S0003347299911582.

[29] P. A. Bednekoff and G. E. Woolfenden. "Florida scrub-jays (Aphelocoma coerulescens) are sentinels more when well-fed (even with no kin nearby)". *Ethology* 109.11 (2003), pp. 895–903.

[30] D. Mikolajewski, M Block, J. Rolff, F. Johansson, A. Beckerman, and R. Stoks. "Predator-driven trait diversification in a dragonfly genus: Covariation in behavioral and morphological antipredator defense". *Evol.; Int. J. Org. Evol.* 64 (Nov. 2010), pp. 3327–25. DOI: 10.1111/j.1558-5646.2010.01078.x.

[31] M. M. Dehn. "Vigilance for predators: detection and dilution effects". *Behav. Ecol. Sociobiol.* 26.5 (1990), pp. 337–342.

[32] J. M. McNamara and A. I. Houston. "Evolutionary stable levels of vigilance as a function of group size". *Anim. Behav.* 43.4 (1992), pp. 641–658.

[33] J. Weibul. *Evolutionary Game Theory*. Cambridge MA: The MIT Press, 2002.

[34] J. Hofbauer and K. Sigmund. "Evolutionary Game Dynamics". *Bull Amer Math Soc* 40.4 (2003), pp. 479–519.

[35] J. Maynard Smith. *Evolution and the Theory of Games*. Cambridge, U.K.: Cambridge University Press, 1982.

[36] J. von Neumann and O. Morgenstern. *The Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.

[37] J. F. Nash. "Equilibrium points in n-person games". *Proc. Natl. Academy Sci.* 36.1 (1950), pp. 48–49. ISSN: 0027-8424. DOI: 10.1073/pnas.36.1.48. eprint: https://www.pnas.org/content/36/1/48.full.pdf. URL: https://www.pnas.org/content/36/1/48.

[38] W. D. Hamilton. "Extraordinary sex ratios. A sex-ratio theory for sex linkage and inbreeding has new implications in cytogenetics and entomology". *Science* 156.3774 (1967), pp. 477–488. DOI: doi:10.1126/science.156.3774.477..

[39] J. Maynard Smith and G. R. Price. *The logic of animal conflict*. Vol. 246. 15. 1973.

[40] A. Traulsen, C. Hauert, H. De Silva, M. A. Nowak, and K. Sigmund. "Exploration dynamics in evolutionary games". *Proc. Natl. Acad. Sci.* 106 (Feb. 2009), pp. 709–12. DOI: 10.1073/pnas.0808450106.

[41] J. Hofbauer and K. Sigmund. *The Theory of Evolution and Dynamical Systems*. Cambridge, U.K.: Cambridge University Press, 1988.

[42] D. T. Bishop and C. Cannings. "A generalized war of attrition". *J. Theor. Biol.* 70.1 (1978), pp. 85–124. ISSN: 0022-5193. DOI: https://doi.org/10.1016/0022-5193(78)90304-1. URL: https://www.sciencedirect.com/science/article/pii/0022519378903041.

[43] R. Cressman and Y. Tao. "The replicator equation and other game dynamics". *Proc. Natl. Acad. Sci.* 111 (2014), pp. 10810–10817.

[44] P. D. Taylor and L. B. Jonker. "Evolutionarily Stable Strategies and Game Dynamics". *Math. Biosci.* 40.1-2 (1978), pp. 145–156.

[45] G. Palm. "Evolutionary stable strategies and game dynamics for n-person games". *J. Math. Biol.* 19 (1984), pp. 329–334. DOI: https://doi.org/10.1007/BF00277103.

[46] C. J. Proctor, M. Broom, and G. D. Ruxton. "Modelling antipredator vigilance and flight response in group foragers when warning signals are ambiguous". *J. Theor. Biol.* 244.4 (2001), pp. 409–417.

[47] C. S. Gokhale and A. Traulsen. "Evolutionary Multiplayer Games". *Dyn. Games Appl.* 4 (2014), pp. 468–488.

[48] J. W. Baron and T. Galla. "How successful are mutants in multiplayer games with fluctuating environments? Sojourn times, fixation and optimal switching". *R. Soc. open Sci.* (2018), p. 5172176.

[49] H. R. Pulliam. "On the advantages of flocking". *J. Theor. Biol.* 38.2 (1973), pp. 419–422.

[50] M. Mangel and C. W. Clark. *Dynamic Modelling in Behav. Ecol.* Princeton: Princeton University Press, 1988.

[51] U. Motro. "Co-operation and defection: Playing the field and the ESS". *J. Theor. Biol.* 151.2 (1991), pp. 145–154. ISSN: 0022-5193. DOI: https://doi.org/10.1016/S0022-5193(05)80358-3. URL: http://www.sciencedirect.com/science/article/pii/S0022519305803583.

[52] M. Bukowski and J. Miekisz. "Evolutionary and asymptotic stability in symmetric multiplayer games". *Int. J. Game Theor.* 33 (2004), pp. 41–54. DOI: https://doi.org/10.1007/s001820400183.

[53] M. Broom and J. Rychtář. "Nonlinear and Multiplayer Evolutionary Games". In: *Advances in Dynamic and Evolutionary Games*. Ed. by F. Thuijsman and F. Wagener. Ann. Int. Soc. Dyn. Games. Birkhäuser, Cham, Apr. 2016, pp. 95–115. ISBN: 978-3-319-28012-7. DOI: 10.1007/978-3-319-28014-1_5.

[54] C. Hauert, F. Michor, M. A. Nowak, and M. Dobeli. "Synergy and discounting of cooperation in social dilemmas". *J. Theor. Biol.* 239.2 (2006), pp. 195–202.

[55] J. Pena and G. Noldeke. "Group size effects in social evolution". *J. Theor. Biol.* 457 (2018), pp. 211–220.

[56] M. Archetti. "How to Analyze Models of Nonlinear Public Goods". *Games* 9.2 (2018). ISSN: 2073-4336. DOI: 10.3390/g9020017. URL: https://www.mdpi.com/2073-4336/9/2/17.

[57] M. Doebeli and C. Hauert. "Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game". *Ecol. Lett.* 8.7 (2005), pp. 748–766. DOI: https://doi.org/10.1111/j.1461-0248.2005.00773.x.

[58] M. Archetti. "Cooperation as a volunteer's dilemma and the strategy of conflict in public goods games". *J. Evol. Biol.* 22.11 (2009), pp. 2192–2200.

[59] M. Archetti. "A Strategy to increase cooperation in the volunteer's dilemma: reducing vigilance improves alarm calls". *Evolution* 65.3 (2011), pp. 885–892. DOI: https://doi.org/10.1111/j.1558-5646.2010.01176.x.

[60] T. Sasaki and I. Okada. "Cheating is evolutionarily assimilated with cooperation in the continuous snowdrift game". *Biosystems* 131 (2015), pp. 51–59. DOI: https://doi.org/10.1016/j.biosystems.2015.04.002.

[61] M. Doebeli, C. Hauert, and T. Killingback. "The Evolutionary Origin of Cooperators and Defectors". *Science* 306 (Oct. 2004), pp. 859–62. DOI: 10.1126/science.1101456.

[62] S. L. Lima. "Vigilance While Feeding and its Relation to the Risk of Predation". *J. Theor. Biol.* 124.3 (1987), pp. 303–316.

[63] E. Sirot. "Negotiation may lead selfish individuals to cooperate: the example of the collective vigilance game". *Proc. R. Soc. B.* 279.1739 (2012), pp. 2862–2867.

[64] P. Schuster and K. Sigmund. "Replicator Dynamics". *J. Theor. Biol.* 100.3 (1983), pp. 533–538.

[65] M. A. Nowak and K. Sigmund. "Evolutionary Dynamics of Biological Games". *Science* 303.5659 (2004), pp. 793–799.

[66] T. Antal, M. A. Nowak, and A. Traulsen. "Strategy abundance in 2×2 games for arbitrary mutation rates". *J. Theor. Biol.* 257.2 (2009), pp. 340–344. ISSN: 0022-5193. DOI: https://doi.org/10.1016/j.jtbi.2008.11.023. URL: http://www.sciencedirect.com/science/article/pii/S0022519308006127.

[67] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge, U.K.: Cambridge University Press, 1998.

[68] F. Dercole and S. Rinaldi. *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications*. Princeton: Princeton University Press, 2008.

[69] M. A. Nowak, N. L. Komarova, and P. Niyogi. "Evolution of Universal Grammar". *Science* 291.5501 (2001), pp. 114–118. DOI: 10.1126/science.291.5501.114.

[70] H. Kauhanen. "Replicator-mutator dynamics of linguistic convergence and divergence". *R. Soc. Open Sci.* 7.11 (2020), p. 201682. DOI: 10.1098/rsos.201682.

[71] K. Özcimder, B. Dey, A. Franci, R. Lazier, D. Trueman, and N. E. Leonard. "Social decision-making driven by artistic explore–exploit tension". *Interdiscip. Sci. Rev.* 44.1 (2019), pp. 55–81. DOI: 10.1080/03080188.2018.1544806.

[72] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. 2nd ed. New York: Springer-Verlag, 1990.

[73] T. L. Vincent and J. S. Brown. *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics*. Cambridge, U.K: Cambride University Press., 2005.

[74] Y. Cohen, T. L. Vincent, and J. S. Brown. "A G-function approach to fitness minima, fitness maxima, evolutionary stable strategies and adaptive landscapes". *Evol. Ecol. Research* 1.8 (1999), pp. 923–942.

[75] T. Vincent, M. Van, and B Goh. "Ecological Stability, Evolutionary Stability and the Ess Maximum Principle". *Evolutionary Ecology* 10 (Nov. 1996), pp. 567–591. DOI: 10.1007/BF01237708.

[76] P. Marrow, U. Dieckmann, and R. Law. "Evolutionary dynamics of predator-prey systems: an ecological perspective". *J. Math. Biol.* 34 (1996), pp. 556–578.

[77] G. Meszena, E. Kisdi, U. Dieckmann, S. A. H. Geritz, and J. A. J. Metz. "Evolutionary Optimisation Models and Matrix Games in the Unified Persepctive of Adaptive Dynamics". *International Institute for Applied Systems Analysis* (2001).

[78] B. J. McGill and J. Brown. "Evolutionary Game Theory and Adaptive Dynamics of Continuous Traits". *Annual Review of Ecology, Evolution and Systematics* 38 (2001), pp. 403–435.

[79] F. Dercole and S. A. H. Geritz. "Unfolding the resident–invader dynamics of similar strategies". *J. Theor. Biol.* 394 (2016), pp. 231–254.

[80] J. Apaloo, J. Brown, and T. L. Vincent. "Evolutionary game theory: ESS, convergence stability, and NIS". *Evo. Ecol. Research* 11 (2009), pp. 489–515.

[81] J. S. Brown and T. L. Vincent. "Evolution of cooperation with shared costs and benefits". *Proc. R. Soc. B.* 275.1646 (2008), pp. 1985–1994.

[82] V. Krivan and R. Cressman. "On evolutionary stability in prey–predator models with fast behavioral dynamics". *Evol. Ecol. Res.* 11 (Feb. 2009).

[83] H. Riskin. *The Fokker-Plank Equation*. New York: Springer-Verlag, 1989.

[84] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Holland: Elsevier, 1992.

[85] C. W. Gardiner. *Handbook of Stochastic Methods*. Berlin: Springer, 1985.

[86] X. Liu, Q. Pan, and M. He. "Stochastic dynamics in the fitness-based process which can be on behalf of the standard Moran, local and Wright-Fisher processes". *J. Theor. Biol.* 460 (2019).

[87] A. Traulsen and M. A. Nowak. "Evolution of cooperation by multi-level selection." *Proc. Natl. Acad. Sci.* 103.29 (2006), pp. 10952–10955.

[88] B. Wu, A. Traulsen, and C. S. Gokhale. "Dynamic Properties of Evolutionary Multi-Player Games in Finite Populations". *Games* 4.29 (2013), pp. 182–199.

[89] B. Wu, P. M. Altrock, L. Wang, and A. Traulsen. "Universality of weak selection". *Phys. Rev. E* 82 (2010), p. 046106.

[90] A. Traulsen, J. M. Pacheco, and M. A. Nowak. "Pairwise comparison and selection temperature in evolutionary game dynamics". *J. Theor. Biol.* 246 (2007), pp. 522–529.

[91] J. M. Pacheco, F. C. Santos, M. O. Souza, and B. Skyrms. "Evolutionary dynamics of collective action in N-person stag hunt dilemmas". *Proc. R. Soc. B.* 276.1655 (2008).

[92] C. Taylor, D. Fudenberg, A. Sasaki, and M. A. Nowak. "Evolutionary dynamics in finite populations". *Bull. Math. Biol.* 66.62 (2004), pp. 1621–1644.

[93] A. Traulsen, J. C. Claussen, and C. Hauert. "Coevolutionary Dynamics: From Finite to Infinite Populations". *Phys. Rev. Lett.* 95 (2005).

[94] D. Fudenberg, L. Imhof, M. A. Nowak, and C. Taylor. "Stochastic Evolution as a Generalized Moran Process". Preprint. Harvard, 2017.

[95] D. Rossa and F. Dercole.and F. Vicini. "Extreme Selection Unifies Evolutionary Game Dynamics in Finite and Infinite Populations". *Bull. Math. Biol.* 79 (2017), pp. 1070–1099.

[96] A. Traulsen, J. C. Claussen, and C. Hauert. "Coevolutionary dynamics in large, but finite populations". *Phys. Rev. E* 74 (2006).

[97] A. Traulsen, M. A. Nowak, and J. M. Pacheco. "Stochastic dynamics of invasion and fixation". *Phys. Rev. E* 74 (1 July 2006), p. 011909. DOI: 10.1103/PhysRevE.74.011909. URL: https://link.aps.org/doi/10.1103/PhysRevE.74.011909.

[98] J. C. Claussen. "Discrete stochastic processes, replicator and Fokker-Planck equations of coevolutionary dynamics in finite and infinite populations". *Banach Center Publications* 80 (Apr. 2008). DOI: 10.4064/bc80-0-1.

[99] U. Dieckmann and R. Law. "The dynamical theory of coevolution: a derivation from stochastic ecological processes". *J. Math. Biol.* 34 (1996), pp. 579–612.

[100] J. A. J. Metz, S. Geritz, G. Meszéna, F. Jacobs, and J. V. Heerwaarden. "Adaptive Dynamics: A Geometrical Study of the Consequences of Nearly Faithful Reproduction". In: 1995, pp. 183–231.

[101] S. A. H. Geritz, E. Kisdl, G. Meszena, and J. A. J Metz. "Evolutionary singular strategies and the adaptive growth and branching of the evolutionary tree." *Evolutionary Ecology* 12 (1998), pp. 35–57.

[102] P. Abrams. ""Adaptive Dynamics" vs. "adatpive dynamics"". *J. Evol. Biol.* 18 (2005), pp. 1162–1165.

[103] M. Broom and J. Rychtář. "The Evolution of a Kleptoparasitic System under Adaptive Dynamics". *J. Math. Biol.* 54 (Mar. 2007), pp. 151–77. DOI: 10.1007/s00285-006-0005-2.

[104] J. A. J. Metz, K. Stankova, and J. Johansson. "The canonical equation of adaptive dunamics for life histories: from fitness-returns to selection gradients and Pontryagin's maximum principle". *J. Math. Biol.* 72 (2018), pp. 1125–1152.

[105] B. Allen, M. A. Nowak, and U. Dieckmann. "Adaptive dynamics with interaction structure". *Am. Nat.* 181.6 (2013), e139–63. DOI: 10.1086/670192. URL: https://app.dimensions.ai/details/publication/pub.1058853989%20and%20http://pdfs.semanticscholar.org/0725/e4a3afebce9fe1fe85d1f7d039b756866568.pdf.

[106] D. Waxman and S. Gavrikets. "20 Questions on Adaptive Dynamics". *J. Evol. Biol.* 18.5 (2005), pp. 1139–1154. DOI: https://doi.org/10.1111/j.1420-9101.2005.00948.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1420-9101.2005.00948.x.

[107] J. A. J. Metz. *Adaptive dynamics*. IIASA Interim Report. IIASA, Laxenburg, Austria, 2012. URL: http://pure.iiasa.ac.at/id/eprint/10223/.

[108] N. G. van Kampen. *Fundamental problems in statistical mechanics of irreversible processes*. Ed. by E. G. D. Cohen. Amsterdam, North Holland, 1962.

[109] F. B. Christiansen. "On Conditions for Evolutionary Stability for a Continuously Varying Character". *Am. Nat.* 138.1 (1991), pp. 37–50. DOI: 10.1086/285203. eprint: https://doi.org/10.1086/285203. URL: https://doi.org/10.1086/285203.

[110] P.A. Abrams and Y. Harada H. Matsudai. "Evolutionarily unstable fitness maxima and stable fitness minima of continuous traits". *Evol. Ecol.* 7 (1993).

[111] J. Hofbauer and K. Sigmund. "Adaptive dynamics and evolutionary stability". *Appl. Math. Lett.* 3.4 (1990), pp. 75–79. ISSN: 0893-9659. DOI: https://doi.org/10.1016/0893-9659(90)90051-C. URL: http://www.sciencedirect.com/science/article/pii/089396599090051C.

[112] M. Dobeli. *Adaptive Diversification*. Princeton, U.S.A: Princeton University Press, 2011.

[113] F. Dercole, U. Dieckmann, M. Obersteiner, and S. Rinaldi. "Adaptive dynamics and technological change". *Technovation* 28.6 (2008), pp. 335–348. ISSN: 0166-4972. DOI: https://doi.org/10.1016/j.technovation.2007.11.004. URL: http://www.sciencedirect.com/science/article/pii/S0166497207001496.

[114] K. Paverin. "Adaptive dynamics of cooperation may prevent the coexistence of defectors and cooperators and even cause extinction". *Proc. R. Soc. B* 277 (2010).

[115] G. F. Turner and T. J. Pitcher. "Attack Abatement: A Model for Group Protection by Combined Avoidance and Dilution". *Am. Nat.* 128 (1986), pp. 228–240.

[116] S. L. Lima and A. P. Zoller. "Anti-predatory vigilance and the limits to collective detection: visual and spatial seperation between foragers." *Behav. Ecol. Sociobiol.* 38.5 (1996), pp. 355–363.

[117] L. Landeau and J. Terborgh. "Oddity and the 'confusion effect' in predation". *Anim. Behav.* 34.5 (1986), pp. 1372–1380.

[118] W. D. Hamilton. "Geometry of the Selfish Herd". *J. Theor. Biol.* 31.2 (1971), pp. 295–311.

[119] G. Beauchamp. *Social Predation: How group living benefits predators and prey*. London, U.K: Elsevier, 2014.

[120] W. D. Hamilton. "The evolution of altruistic behaviour". *Am. Nat.* 9 (1963), pp. 354–356.

[121] A. I. Bijleveld, E. O. Folmer, and T. Piersma. "Experimental evidence for cryptic interference among socially foraging shorebirds". *Behav. Ecol.* 23.4 (2012), pp. 806–814.

[122] G. Beauchamp. "Foraging speed in staging flocks of semipalmated sandpipers: evidence for scramble competition". *Oecologia.* 169.4 (2012), pp. 975–980.

[123] J. Minderman, J. Lind, and W. Cresswell. "Behaviourally mediated indirect effects: interference competition increases predation mortality in foraging redshanks". *J. Anim. Ecol.* 75.3 (2006), pp. 713–723.

[124] G. Beauchamp. "How does food density influence vigilance in birds and mammals?" *Anim. Behav.* 78.2 (2009), pp. 223–231.

[125] C. Leohle. "Social Barriers to Pathogen Transmission in Wild Animal Populations". *Ecology* 76 (1996), pp. 326–335.

[126] I. M. Cote and R. Poulin. "Parasitism and group size in social animals: a meta-analysis". *Behav. Ecol.* 6 (1995), pp. 159–165.

[127] M. A. Elgar. "Predator Vigilance And Group Size In Mammals And Birds: A Critical Review Of The Empirical Evidence." *Biol. Rev.* 64.1 (1989), pp. 13–33.

[128] W. R. Siegfired and L. G. Underhill. "Flocking as an anti-predator strategy in doves". *Anim. Behav.* 23.3 (1975), pp. 504–508.

[129] L. A. Ebensperger and P. K. Wallen. "Grouping increases the ability of the social rodent, *Octodon degus* to detect predators when using exposed microhabitats." *Oikos.* 98.3 (2002), pp. 491–497.

[130] H. R. Pulliam, G. H. Pyke, and T. Caraco. "The scanning behavior of juncos: A game-theoretical approach". *J. Theor. Biol.* 95.1 (1982), pp. 89–103.

[131] R. Cressman and J. Garay. "The effects of opportunistic and intentional predators on the herding behavior of prey". *Ecology* 92.2 (2011), pp. 432–440. DOI: 10.1890/10-0199.1.

[132] G. V. N. Powell. "Experimental analysis of the social value of flocking by Starlings (Sturmus Vulgaris) in relation to predation and foraging." *Anim. Behav.* (1974).

[133] H. Poysa. "Group foraging, distance to cover and vigilance in the teal *Anas crecca*." *Anim. Behav.* 48.4 (1994), pp. 921–928.

[134] C. C. Ioannou, F. Bartumeus, J. Krause, and G. D. Ruxton. "Unified effects of aggregation reveal larger prey groups take longer to find". *Proc. R. Soc. B* 278.1720 (2011), pp. 2985–2990. DOI: 10.1098/rspb.2011.0003. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspb.2011.0003. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2011.0003.

[135] C. J. Barnard. "Flock feeding and time budgets in the House sparrow *Passer Domesticus L.*" *Anim. Behav.* 28.1 (1980), pp. 295–309.

[136] G. Roberts. "Why individual vigilance declines as group size increases". *Anim. Behav.* 51 (1996), pp. 1077–1086.

[137] G. Beauchamp and G. D. Ruxton. "Changes in Vigilance with Group Size under Scramble Competition." *Am. Nat.* 161.4 (2003). PMID: 12776891, pp. 672–675. DOI: 10.1086/368225. eprint: https://doi.org/10.1086/368225. URL: https://doi.org/10.1086/368225.

[138] J. Krause and J. J. Godin. "Predator preferences for attacking particular prey group sizes: consequences for predator hunting success and predation risk". *Anim. Behav.* 50.2 (1995), pp. 465–473.

[139] W. Foster and J. Treherne. "Evidence for the dilution effect in the selfish herd from fish predation on a marine insect". *Nature* 293 (1981), pp. 466–467. DOI: https://doi.org/10.1038/293466a0.

[140] I Vine. "Risk of Visual Detection and Pursuit by a Predator and the Selective Advantage of Flocking Behaviour". *J. Theor. Biol.* 30.2 (1971), pp. 405–422.

[141] M. Treisman. "Predation and the evolution of gregariousness. I. Models for concealment and evasion." *Anim. Behav.* 23.4 (1975), pp. 779–800.

[142] A. L. Jackson, S. Brown, T. N. Sherratt, and G. D. Ruxton. "The effects of group size, shape and composition on ease of detection of cryptic prey". *Behaviour* 142 (2005), pp. 811–826. URL: http://eprints.gla.ac.uk/10373/.

[143] J. Lehtonen and K. Jaatinen. *Behav. Ecol. Sociobiol.* 70 (2016), pp. 449–458. DOI: https://doi.org/10.1007/s00265-016-2075-5.

[144] F. J. Wrona and R. W. J. Dixon. "Group Size and Predation Risk: A Field Analysis of Encounter and Dilution Effects". *Am. Nat.* 137.2 (1991), pp. 186–201. DOI: 10.1086/285153.

[145] P. A. Bednekoff and S. L. Lima. "Re-examining safety in numbers: interactions between risk dilution and collective detection depend upon predator targeting behaviour". *Proc. Biol. Sci.* 265.1409 (1998), pp. 2021–2026. DOI: 10.1098/rspb.1998.0535. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspb.1998.0535. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.1998.0535.

[146] G. Beauchamp. "Disentangling the various mechanisms that account for the decline in vigilance with group size". *Behav. Process.* 136 (2017), pp. 59–63. ISSN: 0376-6357. DOI: https://doi.org/10.1016/j.beproc.2017.01.014. URL: http://www.sciencedirect.com/science/article/pii/S0376635716302327.

[147] E. Sorato, P. R. Gulletta, S. C. Griffitha, and A. F. Russell. "Effects of predation risk on foraging behaviour and group size: adaptations in a social cooperative species". *Anim. Behav.* 84 (2012), pp. 823–834.

[148] C. Ioannou, I. Ramnarine, and C. Torney. "High-predation habitats affect the social dynamics of collective exploration in a shoaling fish". *Sci. Adv.* 3 (May 2017), e1602682. DOI: 10.1126/sciadv.1602682.

[149] J. Lazarus. "Vigilance, flock size and domain of danger in the white-fronted goose." *Wildfowl* 29 (1978), pp. 135–145.

[150] G. Beauchamp. "Should Vigilance always decrease with group size?" *Anim. Behav.* 51 (2001), pp. 47–52. DOI: 10.1007/s002650100413.

[151] W. Cresswell and J. Quinn. "Attack frequency, attack success and choice of prey group size for two predators with contrasting hunting strategies". *Anim. Behav.* 80.4 (2010), pp. 643–648.

[152] C. L. Devereux, E. Fernandez-Juricic, J. R. Krebs, and M. J. Whittingham. "Habitat affects escape behaviour and alarm calling in Common Starlings Sturnus vulgaris". *Ibis* 150 (2008), pp. 191–198.

[153] M. S. Botham, C. J. Kerfoot, V. Louca, and J. Krause. "Predator choice in the field; grouping guppies, *Poecilia reticulata* receive more attacks". *Behav. Ecol. Sociobiol.* 59 (2 2005), pp. 181–184.

[154] L. Morrell, G. D. Ruxton, and R. James. "Spatial positioning in the selfish herd". *Behav. Ecol.* 21 (Mar. 2010), pp. 1367–1373. DOI: 10.1093/beheco/arq157.

[155] A. J. Wood and G. J. Ackland. "Evolving the Selfish herd: emergence of distinct aggregating strategies in an individual-based model." *Proc. Biol. Sci.* (2007).

[156] R. S. Olson, D. Knoester, and C. Adami. "Critical interplay between density-dependent predation and evolution of the selfish herd". In: Amsterdam, The Netherlands, July 2013, pp. 247–254. DOI: 10.1145/2463372.2463394.

[157] H. S. Kimbell and L. J. Morrell. "Selfish herds' of guppies follow complex movement rules, but not when information is limited". *Proc Biol Sci* 282 (2015).

[158] M. Haynes and C. Moore-Crawford. *Birds of a feather*. 1996. URL: https://terpconnect.umd.edu/~wrstrick/secu/ansc455/lab6.htm.

[159] J. K. Parrish. "Re-examining the selfish herd: are central fish safer?" *Anim. Behav.* 38.6 (1989), pp. 1048–1053. ISSN: 0003-3472. DOI: https://doi.org/10.1016/S0003-3472(89)80143-5. URL: https://www.sciencedirect.com/science/article/pii/S0003347289801435.

[160] D. H. Brunton. "Impacts of Predators: Center Nests Are Less Successful than Edge Nests in a Large Nesting Colony of Least Terns". *The Condor* 99 (2 1997), pp. 372–380.

[161] T. Stankowich. "Marginal predation methodologies and the importance of predator preferences". *Anim. Behav.* 66.3 (2003), pp. 589–599. ISSN: 0003-3472. DOI: https://doi.org/10.1006/anbe.2003.2232. URL: https://www.sciencedirect.com/science/article/pii/S0003347203922329.

[162] W. Cresswell and J. Quinn. "Testing domains of danger in the selfish herd: sparrowhawks target widely spaced redshanks in flocks". *Proc. R. Soc. B.* (2006).

[163] S. V. Viscido and D. S. Wethey. "Quantitative analysis of fiddler crab flock movement: evidence for 'selfish herd' behaviour". *Anim. Behav.* 63.4 (2002), pp. 735–741. ISSN: 0003-3472. DOI: https://doi.org/10.1006/anbe.2001.1935. URL: https://www.sciencedirect.com/science/article/pii/S0003347201919359.

[164] L. J. Morrell, G. D. Ruxton, and R. James. "The temporal selfish herd: predation risk while aggregations form". *Proc. R. Soc. B.*. 278 (2010), pp. 605–612.

[165] R. James, P.G. Bennett, and J. Krause. "Geometry for mutualistic and selfish herds: the limited domain of danger". *J. Theor. Biol.* 228.1 (2004), pp. 107–113. ISSN: 0022-5193. DOI: https://doi.org/10.1016/j.jtbi.2003.12.005. URL: http://www.sciencedirect.com/science/article/pii/S0022519303004600.

[166] S. D. Algar, T. Stemler, and M. Small. "The active selfish herd". *J. Theor. Biol.* 471 (2019), pp. 82–90. ISSN: 0022-5193.

[167] T. C. Reluga and S. V. Viscido. "Simulated evolution of selfish herd behavior". *J. Theor. Biol.* 234.2 (2005), pp. 213–225. ISSN: 0022-5193. DOI: https://doi.org/10.1016/j.jtbi.2004.11.035. URL: http://www.sciencedirect.com/science/article/pii/S0022519304005715.

[168] L. M. Bolt. "Predator Confusion Hypothesis". In: *Encyclopedia of Evolutionary Psychological Science*. Ed. by Viviana Weekes-Shackelford, Todd K. Shackelford, and Viviana A. Weekes-Shackelford. Springer International Publishing, 2016, pp. 1–5. ISBN: 978-3-319-16999-6. DOI: 10.1007/978-3-319-16999-6_1516-1. URL: https://doi.org/10.1007/978-3-319-16999-6_1516-1.

[169] J. M. Jeschke and R. Tollrian. "Effects of predator confusion on functional responses". *Oikos.* 111.3 (2005), pp. 547–555.

[170] C. C. Ioannou, C. R. Tosh, L. Neville, and J. Krause. "The confusion effect-from neural networks to reduced predation risk". *Behav. Ecol.* 19.1 (2008), pp. 126–130.

[171] G. D. Ruxton, A. L. Jackson, and C. R. Tosh. "Confusion of predators does not rely on specialist coordinated behaviour". *Behav. Ecol.* 18.3 (2007), pp. 590–596.

[172] C. R. Tosh, A. L. Jackson, and G. D. Ruxton. "The Confusion Effects in Predatory Neural Networks." *Am. Nat.* (2006).

[173] D. C. Krakauer. *Groups confuse predators by exploiting perceptual bottlenecks: a connectionist model of the confusion effect*. 1995. DOI: https://doi.org/10.1007/BF00177338.

[174] F. Zoratto, D. Santucci, and E. Alleva. "Theories commonly adopted to explain the antipredatory benefits of the group life: the case of starling (Sturnus vulgaris)". *Rend. Fis. Acc. Lincei* 20 (2009), pp. 163–176. DOI: https://doi.org/10.1007/s12210-009-0042-z.

[175] R. S. Olson, A. Hintze, F. C. Dyer, D. B. Knoester, and C. Adami. "Predator confusion is sufficient to evolve swarming behaviour". *J. R. Soc. Interface* 10 (2013).

[176] J. K. Parrish. "Comparison of the hunting behavior of four piscine predators attacking schooling prey." *Ethology* 95 (1993), pp. 233–246.

[177] T. J. Pitcher and J. K. Parrish. "Functions of Shoaling Behaviour in Teleosts". In: *The Behaviour of Teleost Fishes*. Ed. by T. J. Picther. Springer, Boston, MA, 1986, pp. 294–337.

[178] W. L. Romey and A. R. Lamb. "Flash Expansion Threshold in Whirligig Swarms". *PLOS ONE* 10.8 (Aug. 2015), pp. 1–12. DOI: 10.1371/journal.pone.0136467. URL: https://doi.org/10.1371/journal.pone.0136467.

[179] C. C. Ioannou, L. Morrell, G. D. Ruxton, and J. Krause. "The Effect of Prey Density on Predators: Conspicuousness and Attack Success Are Sensitive to Spatial Scale". *Am. Nat.* 173.4 (2009), pp. 499–506.

[180] C. C. Ioannou and J. Krause. "Searching for prey: the effects of group size and number". *Anim. Behav.* 75.4 (2008), pp. 1383–1388. DOI: https://doi.org/10.1016/j.anbehav.2007.09.012.

[181] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. "Empirical investigation of starling flocks: a benchmark study in collective animal behaviour". *Anim. Behav.* 76.1 (2008), pp. 201–215. ISSN: 0003-3472.

[182] C. Schradin. "Confusion Effects in a Reptilian and a Primate Predator". *Ethology* 108.8 (2000), pp. 691–670.

[183] J. M. Jeschke and R. Tollrian. "Prey swarming: Which predators become confused and why?" *Anim. Behav.* 74.3 (2007), pp. 387–393.

[184] O. Ohguchi. "Experiments on the selection against colour oddity in water fleas by three-spined sticklebacks". *Zeitschrift fur Tierpsychologie* 47 (1978), pp. 254–267.

[185] R. A. Fuller, S. Bearhop, N. B. Metcalfe, and T. Piersma. "The effect of group size on vigilance in Ruddy Turnstones "Arenaria interpres" varies with foraging habitat". *Ibis* 155.2 (2013), pp. 246–257.

[186] L. A. Ebensperger, M. J. Hurtado, and R. Ramos-Jiliberto. "Vigilance and Collective detection of Predators in Degus (Octodon degus)." *Ethology* 112.9 (2006), pp. 879–887.

[187] J. W. Popp. "Effects of food-handling time on scanning rates among American goldfinches". *Auk* 105 (1988), pp. 384–385.

[188] S. L. Lima. "Back to the basics of anti-predatory vigilance: The group size effect". *Anim. Behav.* 49.1 (1995), pp. 11–20.

[189] O. Pays, P. Blanchard, M. Valeix, S. Chamaille-Jammes, P. Duncan, S. Periquet, M. Lombard, G. Ncube, T. Tarakini, E. Makuwe, and H. Fritz. "Detecting predators and locating competitors while foraging: an experimental study of a medium-sized herbivore in an African savanna". *Oecologia* 169.2 (2012), pp. 419–430.

[190] H. Fritz, M. Guillemain, and D. Durant. "The cost of vigilance for intake rate in the mallard (Anas playrhynchos): an approach through foraging experiments." *Ethology Ecology & Evolution* 14 (2002), pp. 91–97.

[191] G. Beauchamp. "A Spatial model of producing and scrounging". *Anim. Behav.* 76.6 (2008), pp. 1935–1942.

[192] R. R. Repasky. "Using Vigilance Behavior to Test Whether Predation Promotes Habitat Partitioning". *Ecology* 77.6 (1996), pp. 1880–1887. DOI: https://doi.org/10.2307/2265791. URL: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/2265791.

[193] C. J. Barnard. "Factors affecting flock size mean and variance in a winter population of house sparrows". *Anim. Behav.* 74.1/2 (1980), pp. 114–127.

[194] T. Caraco. "Time Budgeting and group size: a test of theory". *Ecology* 60.3 (1979), pp. 618–627.

[195] C. Schutz and C. H. Schulze. "Scanning behaviour of foraging Ruffs "Philomachus pugnax" during spring migration: is flock size all that matters?" *J. Ornithol.* 152.3 (2011), pp. 609–616.

[196] J. Fryxell, A. Mosser, A. Sinclair, and C. Packer. "Group formation stabilizes predator-prey dynamics". *Nature* 449 (Nov. 2007), pp. 1041–3. DOI: 10.1038/nature06177.

[197] J.D. Goss-Custard. "The Energetics of Prey Selection by Redshank, Tringa totanus (L.), in Relation to Prey Density". *J Anim Ecol* (1977).

[198] A. Kvist and A. Lindstrom. "Gluttony in migratory waders - unprecedented energy assimilation rates in vertebrates." *Oikos.* 103.2 (2003), pp. 397–402.

[199] L. A. Ebensperger and M. J. Hurtado. "Seasonal changes in the time budget of degus, "Octodon degus". *Behaviour* 142 (2005), pp. 91–112.

[200] R. C. Ydenberg, C. V. J. Welham, R. Schmid-Hempel, P. Schmid-Hempel, and G. Beauchamp. "Time and energy constraints and the relationships between currencies in foraging theory." *Behav. Ecol.* 5.2 (1994), pp. 28–34.

[201] J. L. Quinn and W. Cresswell. "Escape response delays in wintering redshank, Tringa totanus, flocks: perceptual limits and economic decisions". *Anim. Behav.* 69.6 (2004), pp. 1285–1292.

[202] A. Sansom, W. Cresswell, J. Minderman, and J. Lind. "Vigilance benefits and competition costs in groups: do individual redshanks gain an overall foraging benefit?" *Anim. Behav.* 75.6 (2008), pp. 1869–1875.

[203] W. Cresswell. "The function of alarm calls in redshanks, *tringa totanus*". *Anim. Behav.* 47.3 (1994), pp. 736–738.

[204] M. Yause, J. L. Quinn, and W. Cresswell. "Multiple effects of weather on the starvation and predation risk trade-off in choice of feeding location in Redshank." *Functional Ecology* 17.6 (2003), pp. 727–736.

[205] W. Cresswell. "Surprise as a winter hunting strategy in Sparrowhawks *Accipiter nisus*, Peregrines *Falco peregrinus* and Merlins *F. colurnburiws*". *IBIS* 138.3 (1996), pp. 684–692.

[206] J. L. Quinn and W. Cresswell. "Faced with a choice, sparrowhawks more often attack the more vulnerable prey group". *Oikos.* 104.1 (2004), pp. 71–76.

[207] W. Cresswell, J. Lind, and J. L. Quinn. "Predator-hunting success and prey vulnerability: quantifying the spatial scale over which lethal and non-lethal effects of predation occur". *J. Anim. Ecol.* 79.3 (2010), pp. 556–562.

[208] R. Dobler and M. Kolliker. "Behavioural attainability of evolutionary stable strategies in repeated interactions". *Anim. Behav.* 77 (2009), pp. 1427–1434.

[209] A. Heino and B. Godo. "Fisheries-induced selection pressures in the context of sustainable fisheries". *Bull. Mar. Sci* 70 (2002), pp. 639–656.

[210] N. P. Lester, B. J. Shuter, P. Venturelli, and D. Nadeau. "Life-history plasticity and sustainable exploitation: a theory of growth compensation applied to walleye management". *Ecol. Appl.* 24 (2014), pp. 38–54.

[211] J. R. Stephens and D. W. Stephens. "The economic basis of cooperation: tradeoffs between selfishness and generosity". *Behav. Ecol.* 15 (2004), pp. 255–261.

[212] J. M. McNamara and A. I. Houston. "Phenotypic plasticity as a state dependent life-history decision". *Evo. Ecol.* 6 (1992), pp. 243–253.

[213] J. R. Stephens. "The Evolutionary Biology of Decision making". In: *Better Than Conscious?: Decision Making, the Human Mind, and Implications For Institutions.* MIT Press, 2008. Chap. 13, pp. 285–304.

[214] J. M. McNamara and A. I. Houston. "State-dependent life histories". *Nature* 380 (1996), pp. 215–220.

[215] P. A. Abrams and H. Mastuda. "The effect of adaptive change in the prey on the dynamics of an exploited predator population". *Can. J. Fish. Aquat. Sci.* 62 (2005), pp. 758–766.

[216] P. A. Abrams and H. Mastuda. "Effects of predator-prey interactions and adaptive change on sustainable yield". *Can. J. Fish. Aquat. Sci.* 61 (2004), pp. 175–184.

[217] T. Zentall. "Action imitation in birds". *Learning behavior* 32 (Mar. 2004), pp. 15–23. DOI: 10.3758/BF03196003.

[218] R. Dukas. *Evolutionary ecology of learning. From: Cognitive Ecology: the evolutionary ecology of information processing and decision making.* Chicago Press, 1998, pp. 129–174.

[219] E. C. Snell-Rood. "Selective Processes in Development: Implications for the Costs and Benefits of Phenotypic Plasticity". *Integr. Comp. Biol.* 52 (2012), pp. 31–42.

[220] D. H. Nussey, A. J. Wilson, and J. E. Brommer. "The evolutionary ecology of individual phenotypic plasticity in wild populations". *European Soc. Evol. Biol.* (2007), pp. 831–844.

[221] E. C. Snell-Rood. "An overview of the evolutionary causes and consequences of behavioural plasticity". *Anim. Behav.* 85 (2013), pp. 1004–1011.

[222] F. A. C. C. Chalub and M. O. Souza. "The continuous limit of the Moran process and the diffusion of mutant genes in infinite populations". *arXiv: Analysis of PDEs* (Mar. 2006).

[223] E. Fernandez-Juricic, R. Smith, and A. Kacelik. "Increasing the costs of conspecific scanning in socially foraging starlings affects vigilance and foraging behaviour". *Anim. Behav.* 69.1 (2005), pp. 73–81.

[224] J. R. Courter and G. Ritchison. "Alarm calls of tufted titmice convey information about predator size and threat". *Behav. Ecol.* 10 (2010), pp. 936–942.

[225] C. Hauert, J. Y. Wakano, and M. Doebeli. "Ecological public goods games: Cooperation and bifurcation". *Theor. Popul. Biol.* 73.2 (2008), pp. 257–263. ISSN: 0040-5809. DOI: https://doi.org/10.1016/j.tpb.2007.11.007. URL: https://www.sciencedirect.com/science/article/pii/S0040580907001293.

[226] C. Packer and P. Abrams. "Should Co-operative Groups be More Vigilant than Selfish Groups?" *J. Theor. Biol.* 142 (1990), pp. 341–357.

[227] S. L. Lima and P. A. Bednekoff. "On the perception of targeting by predators during attacks on socially feeding birds". *Anim. Behav.* 82.3 (2011), pp. 535–542. ISSN: 0003-3472. DOI: https://doi.org/10.1016/j.anbehav.2011.06.007. URL: https://www.sciencedirect.com/science/article/pii/S0003347211002491.

[228] G. Roberts. "Cooperation through interdependence". *Anim. Behav.* 70 (2004), pp. 901–908.

[229] S. A. West, A. S. Griffen, and A. Gardner. "Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection". *J. Evol. Biol.* 20 (2006), pp. 415–432.

[230] W. D. Hamilton. "The genetic evolution of social behaviour." *J. Theor. Biol.* 7.1 (1964), pp. 1–52.

[231] P. D. Taylor. "Inclusive fitness arguments in genetic models of behaviour". *J. Math. Biol.* 34 (1996), pp. 654–674.

[232] I. Krams. "Communication in crested tits and the risk of predation". *Anim. Behav.* 61.6 (2001), pp. 1065–1068.

[233] C. L. Devereux, J. Whittingham, M. J. Whittington, E. Fernandez-Juricic, J. A. Vickery, and J. R. Krebs. "Predator detection and avoidance by starlings under differing scenarios of predation risk." *Behav. Ecol.* 17.3 (2006), pp. 303–309.

[234] J-G. J. Godin and H. E. McDonough. "Predator preference for brightly colored males in the guppy: a viability cost for a sexually selected trait". *Behav. Ecol.* 14 (2002), pp. 194–200.

[235] B. L. Ivins and A. T. Smith. "Responses of pikas Ochotona princeps, Lagomorpha to naturally occurring terrestrial predators". *Behav. Sociobiol.* 13 (1983), pp. 277–285.

[236] T. H. Clutton-Brock, P. N. M. Brotherton, F. Russell, M. J. O'Riain, D. Gaynor, R. Kansky, A. Griffin, M. Manser, L. Sharpe, G. M. McIlrath, T. Small, A. Moss, and S. Monfort. "Cooperation, Control, and Concession in Meerkat Groups". *Science* 291 (2001), pp. 478–481.

[237] J. L. Barker, P. B. Barclay, and H. K. Reeve. "Within-group competition reduces cooperation and payoffs in human groups". *Behav. Ecol.* 23 (2012), pp. 735–741.

[238] D. T. Gillespie. *Markov Processes: An introduction for physical scientists*. London, U.K.: Academic Press Inc., 1992.

[239] M. Harper, D. Fryer, and A. Vlasic. "Stochastic Replicator Dynamics Subject to Markovian Switching". *J. Math. Anal. Appl.* 2 (2014), pp. 0-2. DOI: https://arxiv.org/pdf/1406.2405.pdf.

[240] M. Harper, D. Fryer, and A. Vlasic. "Mean Evolutionary Dynamics for Stochastically Switching Environments". *https://arxiv.org/abs/1306.2373* 2 (2013), pp. 0-2.

[241] M. Kennedy and R. D. Gray. "Can Ecological Theory Predict the Distribution of Foraging Animals? A Critical Analysis of Experiments on the Ideal Free Distribution". *Oikos.* 68.1 (1993), pp. 158–166.

[242] J. VanderMeer and B. J Ens. "Models of Interference and Their Consequences for the Spatial Distribution of Ideal and Free Predators". *J. Anim. Ecol.* 66.6 (2014), pp. 846–858.

[243] P. Michelena and J-L. Deneubourg. "How Group Size Affects Vigilance Dynamics and Time Allocation Patterns: The Key Role of Imitation and Tempo". *PloS ONE* 6 (2011), e18631. https://doi.org/10.1371/journal.pone.0018631.

[244] A. Traulsen and C. Hauert. "Stochastic Evolutionary Game Dynamics". In: *Reviews of Nonlinear Dynamics and Complexity*. John Wiley Sons, Ltd. Chap. 2, pp. 25–61. ISBN: 9783527628001. DOI: 10.1002/9783527628001.ch2.

[245] R. Bass. *Stochastic Processes*. Cambridge, U.K: Cambridge University Press, 2011.

[246] M. H. Duong and T. A. Han. "On Equilibrium Properties of the Replicator–Mutator Equation in Deterministic and Random Games". *Dyn. Games. Appl.* 10 (2020), pp. 641–663. DOI: https://doi.org/10.1007/s13235-019-00338-8.

[247] M. Coste, T. Lajous-Loaeza, H. Lombardi, and M-F. Roy. "Generalized Budan–Fourier theorem and virtual roots". *J. Complex.* 21.4 (2005). Festschrift for the 70th Birthday of Arnold Schonhage, pp. 479–486. ISSN: 0885-064X. DOI: https://doi.org/10.1016/j.jco.2004.11.003. URL: https://www.sciencedirect.com/science/article/pii/S0885064X05000075.

[248] Y. A. Kuznetsov. *Elements of Applied Bifurcation Theory*. New York: Springer-Verlag, 2004.

[249] L. Bates. "Imitation: what animal imitation tells us about animal cognition". *Wiley Interdisciplinary Reviews: Cognitive Science* 1.5 (2010), pp. 685–695.

[250] L. Real and T. Caraco. "Risk and foraging in stochastic environments." *Annu Rev Ecol Syst* 17.1 (1986), pp. 371–390.

[251] M. Archetti and I. Scheuring. "Coexistence of Cooperation and Defection in Public Goods Games". *Evolution* 65.4 (2011), pp. 1140–1368.

[252] G. Hardin. "The Tragedy of the Commons". *Science* 162.3859 (1968), pp. 1243–1248. ISSN: 0036-8075. DOI: 10.1126/science.162.3859.1243. eprint: http://science.sciencemag.org/content/162/3859/1243.full.pdf. URL: http://science.sciencemag.org/content/162/3859/1243.

[253] R. Bergmuller, R. A. Johnstone, A. F. Russell, and R. Bshary. "Integrating cooperative breeding into theoretical concepts of cooperation". *Behav. Process.* 76 (2007), pp. 61–72.

[254] W. Cresswell. "Escape responses by redshanks, *Tringa totanus*, on attack by avian predators". *Anim. Behav.* 46.3 (1993), pp. 609–611.

[255] G. Hilton, W. Cresswell, and G. D. Ruxton. "Intraflock variation in the speed of escape-flight response on attack by an avian predator." *Behav. Ecol.* 10 (1999), pp. 391–395.

[256] R. S. Olson, P. B. Haley, F. C. Dyer, and C. Adami. "Exploring the evolution of a trade-off between vigilance and foraging in group-living organisms." *R. Soc. Open Sci.* 2 (2015).

[257] P. A. Bednekoff and S. L. Lima. "Risk allocation and competition in foraging groups: reversed effects of competition if group size varies under risk of predation". *Proc. R. Soc. B.* 271.1547 (2004), pp. 1491–1496. DOI: 10.1098/rspb.2004.2739. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspb.2004.2739. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2004.2739.

[258] W. D. Hintz and D. G. Lonzarich. "Maximizing foraging success: the roles of group size, predation risk, competition, and ontogeny". *Ecosphere* 9.10 (2018), e02456. DOI: https://doi.org/10.1002/ecs2.2456. URL: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/ecs2.2456.

[259] Z. Barta and L-A. Giraldeau. "Breeding colonies as information centers: a reappraisal of information based hypothesis using the producer scrounger game". *Behav. Ecol.* 12 (2001), pp. 121–127.

[260] E. Ranta, H. Rita, and K. Lindstrom. "Competition Versus Cooperation: Success of Individuals Foraging Alone and in Groups". *Am. Nat.* 142 (1993), pp. 42–58.

[261] G. Zdzislaw and B. Hans-Joachim. "Behavioural responses to video playbacks by zebra finch males". *Behav. Process.* 74.1 (2007), pp. 21–26.

[262] J. M. McNamara and A. I. Houston. "The Common Currency for Behavioural Decisions." *Am. Nat.* 127.3 (1986), p. 358.

[263] S. Creel, P. Schuette, and D. Christianson. "Effects of predation risk on group size, vigilance, and foraging behavior in an African ungulate community". *Behav. Ecol.* 25 (4 2014), pp. 773–784. DOI: https://doi.org/10.1093/beheco/aru050.

[264] R. M. Sibly. "Optimal Group Size is Unstable". *Anim. Behav.* (1983).

[265] W. Cresswell. "Age-Dependent Choice of Redshank (Tringa totanus) Feeding Location: Profitability or Risk?" *J Anim Ecol* 63.3 (1994), pp. 589–600.

[266] W. Cresswell. Personal communication. 2015.

[267] W. Cresswell and D. P Whitfield. "The effects of raptor predation on wintering wader populations at the Tyninghame estuary, southeast Scotland". *Ibis.* 136 (1994), pp. 223–232.

[268] A. Brodin and C. W. Clark. "Long-term hoarding in the Paridae: a dynamic model." *Behav. Ecol.* 8.2 (1997), pp. 178–185.

[269] A. Kvist and A. Lindstrom. "Maximum Daily Energy Intake: It Takes Time to Lift the Metabolic Ceiling." *Physiol. Biochem. Zool.* 73.1 (2000), pp. 30–36.

[270] T. Caraco, S. Martindale, and T. S. Whittam. "An empirical demonstration of risk-sensitive foraging preferences". *Anim. Behav.* 28.3 (1980), pp. 820–830. ISSN: 0003-3472. DOI: https://doi.org/10.1016/S0003-3472(80)80142-4.

[271] A. L. Nevaia, T. A. Waite, and K. M. Passino. "State-dependent choice and ecological rationality." *Theor J Biol.* 247.3 (2007), pp. 471–479.

[272] J.R Speakman. "The energetics of wading birds". *PhD Thesis* (1984).

[273] L. Zwarts and A. Blomert. "Selectivity of Whimbrels feeding on fiddler crabs explained by component specific digestibilities". *Ardea* 78.1-2 (1990), pp. 193–208.

[274] J. L. Quinn and W. Cresswell. "Predator hunting behaviour and prey vulnerability". *J. Anim. Ecol.* 73 (2004), pp. 143–154.

[275] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2020. URL: https://www.R-project.org/.

[276] Stan Development Team. *Rstan: the R interface to Stan. R package*. 2020. URL: http://mc-stan.org.

[277] P-C. Bürkner. "brms: An R Package for Bayesian Multilevel Models Using Stan". *Jour. Stat. Software* 80.1 (2017), pp. 1–28. DOI: 10.18637/jss.v080.i01. URL: https://www.jstatsoft.org/index.php/jss/article/view/v080i01.

[278] J. C. Douma and J. T. Weedon. "Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression". *Methods Ecol. Evol.* 10.9 (2019), pp. 1412–1430. DOI: https://doi.org/10.1111/2041-210X.13234. eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.13234. URL: https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13234.

[279] R. McElreath. Oxon, U.K: CRC Press, 2020.

[280] J. Gabry and T. Mahr. *bayesplot: Plotting for Bayesian Models*. R package version 1.8.1. 2021. URL: https://mc-stan.org/bayesplot/.

[281] S. N. Wood, F. Scheipl, and J. J. Faraway. "Straightforward intermediate rank tensor product smoothing in mixed models". *Stat. Comput.* 23 (2013), pp. 341–360. DOI: 10.1007/s11222-012-9314-z.

[282] G. Beauchamp. "Relationship between Distance to Cover, Vigilance and Group Size in Staging Flocks of Semipalmated Sandpipers." *Ethology* 116.7 (2010), pp. 645–652.

[283] F. Dercole, U. Dieckmann, M. Obersteiner, and S. Rinaldi. "Adaptive dynamics and technological change". *Technovation* 28.6 (2008), pp. 335–348. DOI: https://doi.org/10.1016/j.technovation.2007.11.004..

[284] I. Coolen, A. J. W. Ward, P. J. B. Hart, and K. N. Laland. "Foraging nine-spined sticklebacks prefer to rely on public information over simpler social cues". *Behav. Ecol.* 16.5 (2005), pp. 865–870. DOI: https://doi.org/10.1093/beheco/ari064.

[285] I. Eshel. "Evolutionary and continuous stability". *J. Theor. Biol.* 103.1 (1983), pp. 99–111. ISSN: 0022-5193. DOI: https://doi.org/10.1016/0022-5193(83)90201-1. URL: https://www.sciencedirect.com/science/article/pii/0022519383902011.

[286] T. Day and P. D. Taylor. "Evolutionary dynamics and stability in discrete and continuous games". *Evol. Ecol. Research* 5 (2003), pp. 605–613.

[287] U. Dieckmann and J. A. J. Metz. "Surprising evolutionary predictions from enhanced ecological realism". *Theor. Pop. Biol.* 69.3 (2006), pp. 263–281. DOI: https://doi.org/10.1016/j.tpb.2005.12.001. URL: https://www.sciencedirect.com/science/article/pii/S0040580905001607.

[288] O. Pays, S. P. Blomberg, P-C Renaud, F-R Favreau, and P. J. Jarman. "How Unpredictable Is the Individual Scanning Process in Socially Foraging Mammals?" *Behav. Ecol. Sociobiol.* 64.3 (2010), pp. 443–454. ISSN: 03405443, 14320762. URL: http://www.jstor.org/stable/20640532.

[289] C. D. FitzGibbon. "The costs and benefits of predator inspection behaviour in Thomson's gazelles". *Behav. Ecol. Sociobiol.* 34 (1994), pp. 139–148. DOI: https://doi.org/10.1007/BF00164184.