Bee Determined: A Mathematical Analysis of Trapline Formation in Bees

by

Ferdinand Grünenwald
Bachelor of Arts and Sciences, Quest University Canada

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

Bee Determined: A Mathematical Analysis of Trapline Formation in Bees

by

Ferdinand Grünenwald
Bachelor of Arts and Sciences, Quest University Canada

Supervisory Committee

Dr. Mark A. Lewis, Supervisor
(Departments of Mathematics & Statistics & Biology, University of Victoria)

Dr. Eric Foxall , Outside Member
(Department of Mathematics, University of British Columbia)

ABSTRACT

Traplining is a behaviour where animals visit stationary, renewing food sources in a repetitive, non-random order, like a trapper visiting their traps. Reynolds, Lihoreau and Dubois et al. developed biologically plausible models for how traplining might emerge as a result of a simple iterative improvement foraging strategy of bumblebees [44, 31, 14]. While these models have been investigated extensively empirically through simulations, a theoretical understanding of their properties has not yet been determined, and these models have never been fit to data directly. We address both of these research gaps.

In Chapter 1 we provide a mathematically rigorous description of the model, framing it as a version of an edge-reinforced random walk. We show that the model outlined by Reynolds et al. can give rise to stable traplining behaviour where simulated bees visit the same set of flowers in the same order at all large times. In fact, under the additional assumption that simulated bees do not visit any flower twice within the same foraging excursion, also called a bout, we show that the process is guaranteed to converge to a stable trapline eventually. We argue further that the model bridges a gap between two seemingly competing hypotheses for how traplines form in bee foraging behaviour, the nearest neighbour-hypothesis and the order of encounter-hypothesis. The nearest neighbour hypothesis says that bees prefer to fly to the nearest neighbour, whereas the order of encounter hypothesis claims that bees simply retrace their steps. We argue that when bees imperfectly retrace their steps, only weakly reinforcing each step, the resulting trapline is more likely to be the nearest neighbour route and under strong reinforcement, the resulting trapline is more likely to resemble the order of encounters.

In Chapter 2, we fit these models to flower visitation sequence data. Using a simulation study, we verify that we are able to accurately retrieve model parameters and distinguish among several candidate models. We apply these methods to real flower visitation sequence data collected by [53]. We are able to show that bees tend to retrace their steps and seem to be able to remember entire bouts. We find that linear reinforcement better models bee learning than exponential reinforcement, indicating that bees might not converge to the stable traplines previously hypothesized.

Table of Contents

List of Tables

List of Figures

Nomenclature

α	The encounter factor $1 + \ell$
ℓ	The encounter increment $\alpha - 1$
$\mathbf{1}(A)$	Indicator function of the event A
$\Pr(\cdot)$	The probability of an event
a	The non-encounter increment $\beta - 1$
$E(\cdot)$	Expectation of a random variable
X_n	Random variable of a reinforced random walk at time n
$Y_{b,s}$	Random variable of a reinforced random walk during bout number $b(n) = b$ and $s(n) = s$
β	The non encounter factor $1 + a$
L_b	Length of the b-th bout
$\mathrm{W}_{i,j}^{(n)}$	Weight associated with the directed edge i to j at time n

ACKNOWLEDGEMENTS

First and foremost, I want to thank Dr. Mark Lewis for supervising my research and for accepting me as a student. I am incredibly grateful for the freedom and trust you have given me—support that allowed me to explore these topics so openly and creatively. I would also like to extend my gratitude to my committee member Dr. Eric Foxall, whose expertise was invaluable to the development of Chapter 1. I am deeply appreciative of Dr. Joseph Woodgate for so generously sharing the bee-movement data, which made it possible to test my ideas on real observations. It was a joy—and a rare privilege in mathematical biology—to develop theory and immediately implement it on data. I owe special thanks to my parents, Martin and Petra, for their financial support and for believing in me from day one, and to my brother Paul, for always being just a phone call away. Finally, I would like to thank my lab mates for their valuable ideas, encouragement, and company during lab meetings and conferences.

DEDICATION

To all the living things out there.

Introduction

Traplining is a behaviour characterized by repetitive visits to renewing and stationary resources [51]. Darwin was among the first to ever describe such behaviour in bees foraging on flowers [19]. Since then, it has been described in various vertebrates such as primates [35], bats [21], and pigeons [20]. Bumblebees are likely among the most well-studied animals that reportedly use traplining as a foraging strategy[24, 44, 53, 14, 38, 37, 36, 51]. Understanding subtleties in foraging behaviour might help in understanding why some pollinators thrive while others decline [42, 23]. This is of crucial importance as pollinators provide essential ecosystem services to agriculture, whose decline poses a real threat to food security across the globe [18]. In 2013, Reynolds et al proposed a simple and biologically plausible iterative improvement heuristic to model the formation of traplines in bumblebees [44]. The key idea is to model the bee's movement being governed by a kind of reinforcement learning algorithm, where the bee is more likely to repeat actions that lead to desirable outcomes and others are more likely to be avoided. While the empirical plausibility has been examined in great detail, a mathematically rigorous description of their model and its properties is still lacking to this day [30, 31, 3]. Simulations suggest that the model as described in [44] is able to qualitatively exhibit traplining behaviour, where simulated bees, though moving inherently stochastically, seem to converge to visiting a set of flowers in the same repetitive sequence over and over again, a fascinating phenomenon from also a mathematical perspective. We ask, is

the trapline heuristic algorithm guaranteed to converge to a traplining solution to the foraging problem at hand? And if it converges, what solutions does it converge to? Such questions commonly asked by mathematicians working on reinforced random walks [11, 40] and reinforcement learning [48]. This will be the focus of the Chapter 1.

From a more applied perspective, it is reasonable to ask whether data from bee movement patterns are consistent with established heuristic learning models associated with traplining behaviours. We demonstrate in a simulation study that it is possible to fit variations of the trapline heuristic model, as described by [44] and [14], to flower visitation sequence data, and precisely determine the true data-generating model. Using maximum likelihood estimation to investigate a hypothesis comes with the advantage that we obtain parameter estimates precisely quantifying trapline strength, and we can adopt a multiple working hypothesis framework to see which model best explains the observed data. We confront seven distinct models with data collected by [53] to answer the following questions:

1. Do bees tend to revisit previously chosen routes?

2. Do bees take into consideration overall route quality in determining which route to take?

3. Do bees converge to stable traplines under stable conditions?

Chapter 1

A mathematical analysis of trapline models

1.1 Introduction

Traplining is a taxonomically widespread behaviour where animals repeatedly visit feeding locations [21, 50, 28]. Darwin was one of the first to document this repetitive foraging behaviour, mentioning how bumblebees repeatedly follow the same foraging routes, buzzing from flower to flower in an almost predictable way [19]. Despite being observed this early on, precise models for how bees form these foraging routes they follow over and over again have been lacking from the literature until recently. As a consequence, demonstrating the existence of traplining behaviour has remained more vague than many scientists would like it to be [25]. However, over the past 15 years, advancements in harmonic radar technology have allowed researchers to track bees over various spatial scales and demonstrate more rigorously that bees do tend to follow learned routes, visiting, often artificial, feeder locations in a repetitive fashion [38, 28], [53, 14].

The question arises: How do bees choose their traplines? Based on optimal foraging theory, we should expect that bees optimize their routes to maximize their

energy intake per unit time [26]. In the context of traplines, behaving optimally would mean maximizing the ratio of nectar collected per distance travelled. If we assume that per bout, a fixed amount of nectar needs to be collected and each feeding location contains the same amount of resource, optimal foraging behaviour would involve finding the shortest round trip, visiting a sufficient number of feeders in the perfect order to minimize overall travel distance. This problem is known as the travelling salesperson problem and is generally known to be in the class of NP-hard problems, or put more simply, is notoriously difficult to solve [39]. Due to the high computational complexity of the task at hand, and the limited computational capacity of a bumblebee brain, it seems unlikely that bumblebees can find an optimal route in every environment. Nevertheless, we would expect that bees follow some heuristic that allows them to find at least approximate solutions to the problem of maximizing the amount of resources collected per energy spent to collect the resource.

Other social insects, specifically ants, have inspired the ant colony algorithm [12], a way to solve the travelling salesperson problem in a way that is very similar to the proposed trapline heuristic algorithm proposed by Reynolds et al [44]. In a nutshell, the Ant-Algorithm imitates ants creating and following pheromone trails that influence the probability with which the simulated ants move from one place to another. Even though this algorithm is ant-inspired, it is not a biologically correct representation of the foraging behaviour of ants. The algorithm endows the artificial ants with abilities real ants do not have, such as a precise memory of which locations they have already visited and a global pheromone update rule. Furthermore, since our focus here lies on traplining behaviour in bees, pheromone trails left in the air seem like an unlikely mechanism to explain traplining behaviour. So instead, we assume bees must rely on their memory to navigate and solve the foraging problem at hand. Modelling memory-informed movement of animals is

both challenging but also an important next step in the field of animal movement modelling. While models have evolved from simple diffusion models to include environmental covariates, modelling the effects of memory and learned movement patterns has only recently become a modelling challenge that researchers are now willing to tackle [52, 45, 6]. Deeper insights into how animals fall into habits open up new questions and possibilities from a diverse set of perspectives, ranging from behavioural ecology to management. In the context of behavioural ecology, bees and the mechanisms of how they form traplines have been a point of discussion for a long time. In the literature, we found two main hypotheses as to how bumble bees may develop their traplines.

1. The *order of encounter hypothesis* states that bees attempt to revisit feeders in the order they discovered them. This hypothesis is mainly supported by [25].

2. The *nearest neighbour hypothesis* states that bees attempt to form the nearest neighbour route, following the local heuristic of moving to the nearest feeder location at all times. This hypothesis is mainly supported by Anderson, who claims that in general, this heuristic yields a trapline that is very close to optimal with respect to the overall distance travelled to visit all flowers available [2].

Neither of these hypotheses could be thoroughly falsified to this day, and as such, both are plausible. In 2012, Lihoreau et al. suggested a simple iterative improvement heuristic that seems biologically plausible [30]. In a nutshell, this model assumes bees are reinforced random walkers. This means bees move from location to location randomly. The probabilities of moving from one place to another change over time as a direct result of the walker's experience. For example, when a bee encounters food at the destination it chose to move to, we might expect the probability to

repeat this action to increase.

The ordered pair of locations (s, a) is called a *transition vector*, *transition*, *directed edge* or simply an *edge* throughout this report. In the language of reinforcement-learning, we can also think of (s, a) as a *state-action pair*. In fact the models presented by Reynolds et al., Lihoreau et al. and Dubois et al., are reinforcement learning algorithms in disguise where each state-action pair is assigned a weight based on previous experiences that is then used to compute the probability of performing action a (transition to location a) when the agent (the bee) is in state s (currently located at location s).

Association of an edge to an experience can happen at different points in time in the foraging process. In the model presented by Lihoreau et al., this association takes place upon returning to the nest. In this model, an edge is associated with a positive experience when it was traversed during a bout with the highest route quality experienced thus far. This type of global reinforcement is called *route-based* reinforcement. On the other hand, Dubois et al. adopt a different approach. In their model, reinforcement is only applied locally to each edge individually. The edge is associated with a positive experience if food is found at the destination and associated with a negative experience otherwise. This continuous updating rule of edge weights is dubbed online reinforcement. Note, it is unclear which of the two strategies, route-based vs. online learning, would be more advantageous. While route-based learning directly attempts to optimize route quality, it does not distinguish between transitions that are truly beneficial to increasing the overall route quality and those that could have been eliminated to obtain an even higher route quality. This direct association is a strength of online learning; however, the disadvantage is that the highest quality route is not necessarily the route that follows the locally most beneficial transitions selected for by online learning. As a result, the ant colony optimization algorithm uses both route-based (global) reinforcement

and online (local) reinforcement. While using nature-inspired algorithms to solve complex optimization problems is a fascinating task, the purpose of this work is to develop biologically plausible models for learning in bumblebees, which is why we will focus on the models developed by Reynolds et al, Lihoreu et al. and Dubois et al., with the intention to encapsulate bumblebee learning behaviour in finding efficient foraging routes. Generally, online reinforcement is considered to be more parsimonious than route-based reinforcement, since experiences are immediately processed and associated with traversed edges and do not need to be remembered until the end of the bout to then be evaluated and associated with the remembered edges, as it would be the case for route-based reinforcement. Lihoreau et al.'s (2013) model is based on six assumptions, as they state them:

1. The bee can uniquely identify each flower; this is possible using information obtained from the visual context (landmarks, panoramas) and/or integrating information on how far the bee has moved in a certain direction to infer its relative position to other locations. This process is called path integration, and is a well-studied foundational mechanism for insect navigation [7, 9].

2. The bee has a finite probability of using a transition joining each pair of flowers.

3. The initial probability of using a transition (a directed edge between two locations) depends on the Euclidean distance between the two flowers [46, 49].

4. The bee computes the net length of the route travelled using optic flow (odometer).

5. Having completed a route passing through all the flowers at least once (and thus filled their crop capacity), the bee compares the net length of the current route to the net length of the shortest route experienced so far that passes through all the flowers. Here, we assume that passing through all flowers at least once fills up the bee's crop capacity.

6. If the new route is no longer than previously experienced routes, the probabilities of using the edges forming this new route in the next foraging bout are multiplied by a real number $\alpha > 1$ called the *learning factor*, then all probabilities are rescaled with respect to all flowers so that they sum to unity. Repeating the shortest route, therefore, reinforces it.

In a later iteration of a similar model, the reinforcement scheme was adapted to be slightly more parsimonious. In [14] the Dubois et al. define a trapline model that largely agrees with the model outlined above but differs in the timing of reinforcement. They take an *online* reinforcement approach rather than the route-based reinforcement approach from the model outlined above, meaning that weights are adapted on a step by step basis rather than at the end of the route.

While versions of this model have proven to be useful in modelling and showcasing trapline behaviour in honeybees and bumblebees [3, 29, 44], our theoretical understanding of this model is fairly limited. For example, we do not know when the model would predict traplining behavior to emerge. It is also unclear which of the two main hypotheses "order-of-encounter-hypothesis" or "nearest-neighbour-hypothesis" this model really supports.

This prompts our research question that guides the investigations laid out in this chapter.

Question. *What kinds of models can qualitatively model traplining behaviour? And which models give rise to "order of encounter" routes and which models give rise to "nearest neighbour" routes?*

Simulations run by Reynolds et al, Lihoreau et al. and Dubois et al. suggest that traplining is an emergent property of both the route-based and the online learning models discussed above. We further simplify models developed by Reynolds et al, Lihoreau et al. and Dubois et al. to a still reasonably complex model so as

to facilitate rigorous mathematical analysis. In particular, we restrict the bees' movement to always visit a new flower at each step during the bout. Since weight changes made in online reinforcement only take effect in the next bout, both route-based and online learning models reduce to a single simplified model. We are able to prove that, in fact, in the simplified mode, the simulated bees are guaranteed to display traplining behaviour in a very strong sense. More precisely, we show that in the simplified model, simulated bees will follow the same foraging route at all large times. We further show that the order of encounter traplines and nearest neighbour traplines emerge as the learning factor approaches infinity, or approaches unity from above.

These investigations present interesting mathematical challenges in the context of the theory of reinforced random walks. Reinforced random walks have been a thriving area of research since their introduction by Coppersmith and Diaconis (1987). For further reading on the topic of reinforced random processes, we point to [40, 11].

1.2 The asymptotic dynamics of the reinforced coin

1.2.1 Definition

A key aspect of all trap-line models is the transition probabilities that are modulated based on previous experiences of the bee. We call this process reinforcement. In a popular family of trapline models, the probability of a traversed edge is multiplied by a factor, called the learning factor, and is re-normalized so that all probabilities associated with edges leaving a location sum to unity. We call this reinforcement multiplicative or exponential.

A great toy example to understand the mechanics of multiplicative reinforcement and the dynamics that arise from it is the simplest possible example: a reinforced coin.

Imagine a coin where, after each coin flip, the side it landed on becomes more likely by a factor $\alpha > 1$. This can be interpreted in terms of weights $w_0^{(n)}, w_1^{(n)}$ associated with failures and successes, respectively. Then $w_0^{(n)} = w_0^{(0)}\alpha^{k_0^{(n)}}$ and $w_1^{(n)} = w_1^{(0)}\alpha^{k_1^{(n)}}$ where $k_0^{(n)}$, and $k_1^{(n)}$ are the number of failures and successes up to time n. The weights are rescaled to sum to one and obtain the probabilities to flip either side

$$\Pr(X_n = 1 \mid w_0^{(n)} = w, w_1^{(n)} = w') = \frac{w'}{w + w'}.$$

Probabilities for obtaining either side of the coin grow in the same way as in [44]. where the probability of choosing an edge, corresponding here to the outcome of a coin flip, is multiplied by a learning factor α and normalized to sum to unity. Note, when weights grow exponentially effectively, there is no difference between normalizing weights after every time step or letting them grow and deriving probabilities by rescaling.

Definition 1.1. Let a reinforced coin be a sequence of random variables $(u_n)_{n\in N}$ be i.i.d uniform on $[0, 1]$ and define p_n recursively by

$$p_t = \begin{cases} f_1(p_n) = \dfrac{p_{n-1}\alpha}{1 - p_{n-1} + \alpha p_{n-1}} & \text{if } u_{n-1} \leq p_{n-1} \\ f_0(p_n) = \dfrac{p_{n-1}}{1 - (1 - p_{n-1}) + \alpha(1 - p_{n-1})} & \text{if } u_{n-1} > p_{n-1} \end{cases} \tag{1.1}$$

The outcomes of reinforced coin flips are then given by $X_n = \mathbf{1}(u_n \leq p_n)$ where u_n is uniform on $[0, 1]$.

1.2.2 Lower bounds on the convergence probabilities

Question. *What is the probability* $P := \Pr(X_n = 1 \forall n \in \mathbb{N})$ *that all coin flips are successes?*

Define the conditional probability $q_n = \Pr(X_n = 1 \mid X_i = 1 \forall i < n)$. The probability that all of $X_n = 1$ is then given by the product of conditional probabilities $\prod_{n\in\mathbb{N}} q_n$. Recall, an infinite product $\prod_{n\in\mathbb{N}}(1 - a_n)$ with $a_n \geq 0$ converges if and only if $\sum_{n\in\mathbb{N}} a_n$ converges. So we can conveniently write q_n in terms of its counter probability and express this in terms of the odds ratio $O_n = q_n/(1 - q_n)$,

$$q_n = 1 - (1 - q_n) = 1 - \frac{1}{1 + O_n}.$$

Why is it convenient to write q_n in this form? We notice that the odds ratio O_n grows geometrically with the number of successful coin flips. Indeed, if $X_t = 1$ then

$$O_{t+1} = \frac{q_{n+1}}{1 - q_{n-1}} = \frac{q_n\alpha}{1 - q_n + \alpha q_n}\frac{1 - q_n + \alpha q_n}{1 - q_n} = \alpha\frac{q_n}{1 - q_n} = \alpha O_n. \tag{1.2}$$

Combining these observations, we have that the probability that all coin flips are successes is positive.

$$P = \Pr(X_n = 1 \forall n \in \mathbb{N}) = \prod_{t\in\mathbb{N}} qn_t = \prod_{t\in\mathbb{N}} \left(1 - \frac{1}{1 + \alpha^n O_0}\right),$$

which converges to a positive number if and only if $\alpha > 1$ and $O_0 > 0$.

Knowing that $P > 0$, we will now find an explicit positive lower bound using the commonly known inequality $\log(1 - x) > -\frac{x}{1-x}$ for $x \in (0, 1)$ [13]. Let $r_n = 1 - q_n$ Thus, we write the product as a sum

$$\prod_{n\in\mathbb{N}} (1 - r_n) = \exp\left(\sum_{n\in\mathbb{N}} \log(1 - r_n)\right) > \exp\left(-\sum_{n\in\mathbb{N}} \frac{r_n}{1 - r_n}\right) = \exp\left(-\sum_{n\in\mathbb{N}} O_n^{-1}\right).$$

Since $O_n = \alpha^n O_0 = \alpha^n p_0/(1-p_0)$,

$$P > \exp\left(\sum_{t\in\mathbb{N}} O_n^{-1}\right) = \exp\left(-\frac{1-p_0}{p_0}\frac{\alpha}{\alpha-1}\right).$$

Lemma 1.2. *The reinforced coin has a positive probability that all future coin flips are successes. This probability is bounded below by*

$$\Pr(X_m = 1 \forall m > n \mid p_n = p) \geq \exp\left(-\frac{1-p}{p}\frac{\alpha}{\alpha-1}\right).$$

Proof. This is a direct consequence of the above observations. □

Question. *What is the probability that the same side of the coin lands facing up at all large times?*

We argue that this probability must be equal to 1 by showing that the number of times the less likely side of the coin is flipped is finite with probability 1.

Note that the inequality established above equally transfers to the event that all coin flips are failures. And we have

$$\Pr(X_n = 0 \forall n \in \mathbb{N}) \geq \exp\left(-\frac{p_0}{1-p_0}\frac{\alpha}{\alpha-1}\right), \tag{1.3}$$

$$\Pr(X_t = 1 \forall t \in \mathbb{N}) \geq \exp\left(-\frac{1-p_0}{p_0}\frac{\alpha}{\alpha-1}\right), \tag{1.4}$$

which gives the probability that all coin-flips have the same outcome is bounded below by the probability of exclusively flipping the more likely side

$$\Pr(X_n = X_0 \forall n \in \mathbb{N}) \geq \max\left\{\exp\left(-\frac{1-p_0}{p_0}\frac{\alpha}{\alpha-1}\right), \exp\left(-\frac{p_0}{1-p_0}\frac{\alpha}{\alpha-1}\right)\right\} \tag{1.5}$$

$$\geq \exp\left(-\frac{\alpha}{\alpha-1}\right). \tag{1.6}$$

Notice that this lower bound no longer depends on p_0 and thus

$$\Pr(X_n = X_N \forall n \geq N) \geq \exp\left(-\frac{\alpha}{\alpha - 1}\right). \tag{1.7}$$

Thus, every time the outcome of the current coin flip is not the more likely side, there is still a chance of at least $\exp\left(-\frac{\alpha}{\alpha-1}\right) > 0$ that all future coin flips will yield the more likely side. Intuitively, this means that eventually, the event that all future coin flips agree must happen.

Theorem 1.3. *For $\alpha > 1$, the reinforced coin yields the same outcome at all large times.*

Proof. Let N count the number of times the coin does not land with the more likely side facing up. Let $q = \exp\left(-\frac{\alpha}{\alpha-1}\right)$. By equation 1.7, we know the probability of at least n failures is at most

$$\Pr(N \geq n) \leq (1 - q)^n.$$

Recognizing the geometric distribution, we can conclude that $E(N) \leq \exp\left(\frac{\alpha}{\alpha-1}\right) < \infty$ and $N < \infty$ almost surely, which is what we wanted to show. □

Remark 1.4. Notice that we expect the number of times consecutive coin flips yield different outcomes is at most $E[N] \leq q = \exp\left(\frac{\alpha}{\alpha-1}\right)$.

We would also like to mention that the inequality above generalizes readily to sequences of events

Lemma 1.5. *Let $(A_n)_{n\in\mathbb{N}}$ be an infinite sequence of events with conditional odds given by*

$$O(A_n) = \frac{\Pr(A_n \mid \bigcap_{m<n} A_m)}{1 - \Pr(A_n \mid \bigcap_{m<n} A_m)}.$$

Then the probability that all of A_n happen in sequence is bounded below by

$$\Pr(\bigcap_{n\in\mathbb{N}} A_n) \geq \exp\left(-\sum_{n\in\mathbb{N}} \frac{1}{O(A_n)}\right).$$

Proof. This is followed again by writing the conditional probabilities as their counter probabilities expressed through the odds ratio and applying the bounds of the logarithm. Let $r_n = 1 - \Pr(A_n \mid (A_m)_{m<n})$

$$\begin{aligned}
\Pr((A_n)_{n\in\mathbb{N}}) &= \prod_{n\in\mathbb{N}} \Pr(A_n \mid (A_m)_{m<n}) && (1.8)\\
&= \prod_{n\in\mathbb{N}} (1 - r_n) && (1.9)\\
&= \exp\left(\sum_{n\in\mathbb{N}} \log(1 - r_n)\right) && (1.10)\\
&> \exp\left(\sum_{n\in\mathbb{N}} \frac{r_n}{1 - r_n}\right) && (1.11)\\
&= \exp\left(\sum_{n\in\mathbb{N}} \frac{1 - \Pr(A_n \mid (A_m)_{m<n})}{\Pr(A_n \mid (A_m)_{m<n}}\right) && (1.12)\\
&= \exp\left(\sum_{n\in\mathbb{N}} O_n^{-1}\right). && (1.13)
\end{aligned}$$

□

Using lemma 1.5 gives another way to see that the reinforced coin yields the same outcome at all large times.

Let B_n be the event of having $n+1$ consecutive coin tosses that yield the same outcome. Let $B_\infty = \lim_{n\to\infty} B_n$ then

$$\Pr(B_\infty) \geq \exp\left(-\sum_{n\in\mathbb{N}} \frac{1}{O(B_n)}\right) = \exp\left(-1/O(B_0)\sum \alpha^{-n}\right) = \exp\left(-\frac{1}{O(B_0)}\frac{\alpha}{\alpha - 1}\right). \tag{1.14}$$

TO derive the first equality, we use the fact that for the reinforced coin $O_n = O_0\alpha^n$, which follows directly when considering the growth of weights that define the probabilities of the coin. More precisely, the weight associated with the side that has been flipped increases by a factor α. Certainly, we have a streak of length 1 and thus B_0 happens almost surely, which implies an infinite odds ratio $O(B_0) = \infty$. Using equation (1.14) we obtain that $\Pr(B_\infty) \geq 1$. While this approach might be more elegant in this case, the previous argument showing that the number of times the coin does not yield the more likely outcome generalizes more readily to walks on graphs.

1.2.3 The effect of the reinforcement factor on the distribution of long-term behaviours

Here, we attempt to answer the question:

Question. *What is the effect of varying α on the probability of converging to either side of the coin?*

Generally, which side will be flipped at all large times is random. The exact distribution as a function of p_0 and α is non-trivial. Here we predict the long-term behaviour of the coin for the limiting cases when $\alpha \to \infty$ and when $\alpha \downarrow 1$. Maybe, surprisingly so, we will find that the convergent side becomes more predictable as the reinforcement factor decreases to one.

Strong reinforcement

We begin by investigating what happens in the case that $\alpha \to \infty$. As one might guess, in this scenario, the long-term behaviour is entirely predicted by the first coin flip.

Theorem 1.6. *As $\alpha \to \infty$ we have* $\Pr(X_t = X_0, \forall t \in \mathbb{N}) \to 1$

Proof. Conditioning on the first coin flip and using equations 1.3, 1.4 we get

$$\Pr(X_t = X_0 \forall t \in \mathbb{N}) = (1-p_0)\Pr(X_t = 0 \forall t > 0 \mid X_0 = 0) + p_0 \Pr(X_t = 1 \forall t \in \mathbb{N} \mid X_0 = 1) \tag{1.15}$$

$$\geq (1-p_0)\exp\left(-\frac{1}{\alpha}\frac{p_0}{1-p_0}\frac{\alpha}{\alpha-1}\right) + p_0 \exp\left(-\frac{1}{\alpha}\frac{1-p_0}{p_0}\frac{\alpha}{\alpha-1}\right). \tag{1.16}$$

Observe that as $\alpha \to \infty$ we see that the exponential terms approach 1 and thus

$$\lim_{\alpha\to\infty} \Pr(X_t = X_0 \forall t \in \mathbb{N}) = (1-p_0) + p_0 = 1,$$

which is what we set out to prove. □

Weak reinforcement

For a reinforced coin with a reinforcement factor close to one $\alpha = 1 + \epsilon$, we will show that the convergent side $X_\infty := \lim_{t\to\infty} X_t$ is almost surely the side the coin is biased towards initially.

Theorem 1.7. *Let $X_\infty := \lim_{t\to\infty} X_t$. As α approaches 1 from the right*

$$\Pr(X_\infty = 1) \to \begin{cases} 1 & \text{if } p_0 > 1/2 \\ 1/2 & \text{if } p_0 = 1/2 \\ 0 & \text{if } p_0 < 1/2 \end{cases}$$

To build some intuition for why this might be the case, consider the SDE approximation of the discrete time and space process $(p_n)_{n\in\mathbb{N}} \in [0,1]$ as Brownian motion with drift. Here, p_n is the success probability of the stochastic process of the reinforced coin $(X_n)_{n\in\mathbb{N}}$.

We construct the corresponding continuous time jump process $(p_t)_{t\in\mathbb{R}}$ using its

exponential embedding. This process can be further approximated, as we let the step size approach 0, by Brownian motion that agrees with (p_t) in the first and second moments.

Now the intuition is that as α approaches 1, it takes more time steps for the p_t to change to a significant degree. This means that the movement of p_t is governed more by a law of large numbers, manifesting in a drift process, than by its variance, causing erratic movement manifesting in a diffusion process.

Also notice how step sizes decrease as $\alpha \to 1^+$. Explicitly, we compute the step size

$$\Delta(p_t) = \begin{cases} \Delta_1(p_t) = f_1(p_t) - p_t \text{ if } X_t = 1 \\ \Delta_0(p_t) = f_0(p_t) - p_t \text{ if } X_t = 0 \end{cases}$$

where we define

$$\Delta_1(p) = f_1(p) - p = \frac{p\alpha}{1 - p + \alpha p} - p \tag{1.17}$$

$$\Delta_0(p) = f_0(p) - p = \frac{p}{1 - (1-p) + \alpha(1-p)} - p. \tag{1.18}$$

Letting $\alpha \downarrow 1$ also has the consequence of $\Delta(p_t) \to 0$ bringing us on track to use an SDE approximation

Let $\alpha = 1 + \epsilon$ with $\epsilon > 0$. The first moment, mean change in position, is given by

$$a_\epsilon(p) = E[\Delta(p)] = p\Delta_1(p) + (1-p)\Delta_0(p) \tag{1.19}$$

We can approximate this using a Taylor expansion around $\alpha = 1$ to get

$$\Delta_1(p) = f_1(p) - p = \frac{p(1+\epsilon))}{1 - p + (1+\epsilon)p} - p \approx p(1-p)\epsilon + O(\epsilon^2), \tag{1.20}$$

$$\Delta_0(p) = f_0(p) - p = \frac{p}{1 - (1-p) + \alpha(1-p)} - p \approx -p(1-p)\epsilon + O(\epsilon^2). \tag{1.21}$$

Substituting into the equation for the drift term, we have

$$a_\epsilon(p) \approx p(p(1-p)\epsilon + O(\epsilon^2)) + (1-p)(-p(1-p)\epsilon + O(\epsilon^2)) \tag{1.22}$$

$$= \epsilon p(1-p)(2p-1) + O(\epsilon^2) \tag{1.23}$$

Now the second moment is

$$\sigma_\epsilon^2(p) = E(\Delta^2(p)) = p\Delta_1(p)^2 + (1-p)\Delta_0(p)^2. \tag{1.24}$$

Again, we use the Taylor expansion of $\Delta_0(p), \Delta_1(p)$ up to second order to get an approximation for small values of ϵ which simplifies to

$$\sigma_\epsilon^2(p) = p(p(1-p)\epsilon)^2 + (1-p)(-p(1-p)\epsilon)^2 + O(\epsilon^3) \tag{1.25}$$

$$= \epsilon^2(p(p(1-p))^2 + (1-p)(p(1-p))^2) + O(\epsilon^3) \tag{1.26}$$

$$= \epsilon^2(p(1-p))^2(p+1-p) + O(\epsilon^3) \tag{1.27}$$

$$= \epsilon^2(p(1-p))^2 + O(\epsilon^3) \tag{1.28}$$

Now taking the limit as $\Delta t \to 0$, we have

$$dp_t = \mu_\epsilon(p)dt + \sqrt{\sigma_\epsilon^2(p)}dW_t \tag{1.29}$$

$$= \epsilon p(1-p)(2p-1) + \epsilon p(1-p)dW_t \tag{1.30}$$

$$p_t(0) = p_0 \in [0,1] \tag{1.31}$$

Notice that both drift and diffusion are of order $O(\epsilon)$, meaning both would vanish when $\epsilon \to 0$. This means the exponential embedding of the process slows down too much and eventually, nothing happens; the process is stuck.

To get a more meaningful idea of the original jump process, we need to speed time up. Currently, as we shrink ϵ, the process also slows down by a factor of ϵ. To counteract that, consider the time scale $\tau = t/\epsilon$. We have that $dt = \frac{1}{\epsilon}d\tau$. So, using the time scaling property of Brownian motion, the corresponding SDE of the sped up process is

$$dp_\tau^\epsilon = \mu_\epsilon(p_\tau^\epsilon)\frac{1}{\epsilon}d\tau + \sqrt{\sigma_\epsilon^2(p_\tau^\epsilon)}dW_\tau \tag{1.32}$$

$$= p(1-p)(2p-1) + \sqrt{\epsilon}p(1-p)dW_\tau \tag{1.33}$$

$$p_\tau^\epsilon = p_0 \in (0,1) \tag{1.34}$$

Now, as $\epsilon \to 0$ we observe deterministic behavior $\frac{dp}{d\tau} = p(1-p)(2p-1)$ This ODE has two stable equilibria, one at $p = 0$ and the other at $p = 1$. An unstable equilibrium is found at $p = 0.5$, as we would expect.

While this SDE approximation gives good intuition for why the weakly reinforced coin converges to the biased side, theorems for these approximations only hold over finite time horizons [17]. So, for a more rigorous approach, we turn to the optional stopping theorem, which says that for a bounded discrete-time Martingale $(Y_n) \in \mathbb{N}_0 \cup \{\infty\}$ and a stopping time τ then $E[Y_0] = E[Y_\tau]$ based on the equality of expectations it is possible to obtain explicit expressions for the boundary "hitting probabilities" which are of interest to us here. So, in order to apply the optional stopping theorem, we need to transform the process of the reinforced random coin into a bounded Martingale on $\mathbb{N} \cup \{\infty\}$, which converges to one of the boundary points. This is done by calculating the function, which is defined more precisely in the proof to follow.

Theorem 1.8. *Assume $(X_t)_{t\in\mathbb{N}}$ is the reinforced coin as defined in definition 1.1.*

Let $X_\infty := \lim_{t\to\infty} X_t$. *As* α *approaches* 1 *from the right. Then*

$$\Pr(X_\infty = 1) = \begin{cases} 1 & \text{if } p_0 > 1/2, \\ 1/2 & \text{if } p_0 = 1/2, \\ 0 & \text{if } p_0 < 1/2, \end{cases}$$

where p_n *is defined as in Definition 1.1.*

Proof. Note that the case $\Pr(X_\infty = 1) = 1/2$ when p_0 follows by symmetry. Furthermore, as a consequence of Theorem 1.3 it suffices to show that $p_n \to \mathbf{1}_{p_0>1/2}$ as $\alpha \downarrow 1$ with probability 1.

Let $S_n := S_n^+ - S_n^-$ be the effective number of successes where

$$S_n^+ := \sum_{i<n} \mathbf{1}_{X_i=1}, \quad S_n^- := \sum_{i<n} \mathbf{1}_{X_i=0}$$

are the respective numbers of successes and failures up to but not including time n. Equivalently to Definition 1.1

$$p_n = \frac{p_0 \alpha^{S_n}}{p_0 \alpha^{S_n} + (1 - p_0)}$$

which is easily derived when using weights. we say $w_0^{(n)}$ and $w_q^{(n)}$ are weights for failure and success on the n-th coinflip so that $p_n = \frac{w_1^{(n)}}{w_0^{(n)}+w_1^{(n)}}$. For $n = 0$ we can assume without loss of generality that $p_0 = w_1^{(0)}$ and $w_0^{(0)} = 1 - p_0$. Then the reinforcement condition from Definition 1.1 implies that

$$p_{n+1}/p_n = \begin{cases} \alpha & \text{if } X_n = 1, \\ 1/\alpha & \text{if } X_n = 0. \end{cases}$$

Thus $w_0^{(n)} = (1-p_0)\alpha^{S_n^-}$ and $w_1^{(n)} = p_0\alpha^{S_n^+}$. This gives us that

$$p_n = \frac{p_0\alpha^{S_n^+}}{w_1^{(0)}\alpha^{S_n^+} + w_0^{(0)}\alpha^{S_n^-}} = \frac{p_0\alpha^{S_n}}{p_0\alpha^{S_n} + (1-p_0)},$$

as desired.

For a given p_0, we have $\{p_n\}_{n\in\mathbb{N}} \subseteq \{x_k\}_{k\in\mathbb{N}}$ with

$$x_k := \frac{p_0\alpha^k}{p_0\alpha^k + (1-p_0)}.$$

We define the natural scale function $y : \mathbb{Z} \to \mathbb{R}$ so that if $Y_n = y(S_n)$ we have the Martingale property

$$E[Y_{n+1} \mid Y_0, \ldots, Y_n] = Y_n. \tag{1.35}$$

We compute the natural scale through its increments, $z_k := y(S_k) - y(S_{k-1}))$ where we define $z_0 := 1$. From (1.35) , we require that the conditional increments are zero,

$$\begin{aligned} E[Y_{n+1} - Y_n \mid S_n = k] &= x_k(y(k+1) - y(k)) + (1-x_k)(y(k-1) - y(k)) && (1.36)\\ &= x_k(y(k+1) - y(k)) - (1-x_k)(y(k) - y(k-1)) && (1.37)\\ &= z_{k+1}x_k - z_k(1-x_k) && (1.38)\\ &= 0. && (1.39) \end{aligned}$$

which gives the recurrence relation $z_{k+1} = z_k(1-x_k)/x_k$. Enforcing the boundary condition $z_0 := 1$, we obtain

$$z_k = \begin{cases} \prod_{i=0}^{k-1}(1-x_i)/x_i, & k > 0, \\ \prod_{i=k}^{-1} x_i/(1-x_i), & k < 0. \end{cases}$$

Explicitly, we can compute

$$\frac{x_k}{1-x_k} = \alpha^k \frac{p_0}{1-p_0}.$$

Let $r = \frac{p_0}{1-p_0}$ and we can write

$$z_k = \begin{cases} \alpha^{-\binom{k}{2}} r^{-k}, & k > 0, \\ \alpha^{-\binom{|k|+1}{2}} r^{|k|}, & k < 0. \end{cases}$$

Note that z_k decreases to 0 super exponentially for $\alpha > 1$ meaning that $\sum_{k\in\mathbb{Z}} |z_k| < \infty$, i.e. the natural scale is bounded $\exists\, C > 0$ such that $|Y_k| < C$ for all k. Hence, we can apply the optional stopping theorem to compute the probability of converging to either side of the boundary.

The boundary of the natural scale is defined as $y(\pm\infty) = \lim_{k\to\pm\infty} y(k)$. Let $Y_\infty := \lim_{n\to\infty} Y_n$. From Theorem 1.3 we know that Y_∞ converges almost surely to one of the boundary points i.e. $\Pr(Y_\infty := \lim_{n\to\infty} Y_n \in \{y(\infty), y(-\infty)\}) = 1$. Let $q = \Pr(Y_\infty = y(\infty))$ be the probability that the coin lands as a success at all large times. Now, boundedness and convergence of Y_n implies convergence in expectation by the bounded convergence theorem. The optional stopping time theorem implies

$$E[Y_0] = E[Y_\infty] = q\, y(\infty) + (1-q)\, y(-\infty).$$

If we further assume that Y_0 is deterministic, then we get that the probability of hitting the right boundary (i.e eventually all coins are successes) before hitting the left boundary (i.e. eventually all coin flips are failures) is given by

$$q = \frac{Y_0 - y(-\infty)}{y(\infty) - y(-\infty)}.$$

If we let $y(0) = 0$ then $Y_0 = 0$ and $y(\infty) = \sum_{k\geq 1} z_k$, $y(-\infty) = -\sum_{k\geq 0} z_{-k}$ and

$$q = \frac{|y(-\infty)|}{y(\infty) + |y(-\infty)|} = \frac{1}{\frac{y(\infty)}{y(-\infty)} + 1}. \tag{1.40}$$

We approximate $\frac{y(\infty)}{y(-\infty)}$ for $\alpha \approx 1+\epsilon$ with $\epsilon > 0$. Notice that for a fixed k $z_k(\alpha)$ is monotonically increasing as $\alpha \downarrow 1$. For each of these sequences, the pointwise limit is

$$\lim_{\alpha\downarrow 1} z_k(\alpha) = r^{-k}.$$

The monotone convergence theorem for non-negative sums permits us too interchange this limit and the infinite sum so that

$$\lim_{\alpha\downarrow 1} \sum_{k\in\mathbb{N}} z_k = \sum_{k\in\mathbb{N}} \lim_{\alpha\downarrow 1} z_k.$$

We compute

$$\lim_{\alpha\downarrow 1} \frac{y(\infty)}{y(-\infty)} = \frac{\sum_{k\in\mathbb{N}} \lim_{\alpha\downarrow 1} z_k}{\sum_{k\in\mathbb{N}} \lim_{\alpha\downarrow 1} z_{-k}} = \frac{\sum_{k=0}^{\infty} r^{-k}}{\sum_{k=0}^{\infty} r^{k}} = \begin{cases} 0 & , r > 1, \\ 1 & , r = 1, \\ \infty & , r < 1. \end{cases} \tag{1.41}$$

Substituting back into equation (1.40) we obtain the desired result. □

Here, we have shown that under exponentially growing weights, the reinforced coin yields the same outcome at all large times. When reinforcement is very weak, the coin will fixate on the initially heavier, i.e. more likely side. When reinforcement is very strong, all coin flips yield the same outcome with high probability. Since the reinforcement dynamics are similar to the trapline models from Reynolds et al. (2013) and Dubois et al (2021), the results shown here give us reason to believe similar that these models also exhibit similar asymptotic behaviour. As we will see

in the next section, we are able to extend some of these results to a slightly simplified version of these models.

1.3 Asymptotic analysis of non-backtracking bee foraging bouts

In previous sections, we explored the mathematical properties of the reinforced coin as an oversimplified version of the trapline heuristic model developed by Dubois et al. A significantly less simplified version of this model will be treated in this section. The only difference is that we assume no flower revisits within the same bout, episodically non-backtracking. Here, we extend the convergence result to episodically non-backtracking reinforced random walks, which brings us a step closer to understanding the dynamics that govern the trapline heuristic model. An episodically non-backtracking reinforced random walk is a random walk through a complete directed graph $G = (V, E)$, with $E = V \times V$ where the random walker avoids revisiting the same vertex within the same episode, also called a bout. An episode or bout starts and ends when the random walker visits the distinct vertex v^* representing the nest. Since we assume $|V| < \infty$, every episode will be of finite length. The probability that the random walker moves from location i to j changes with the number of times the edge (i, j) is traversed. As in the reinforced coin, we define these probabilities using weights that grow exponentially with the number of directed traversals.

We follow similar arguments as we did for the reinforced coin in section 1.2. We first show that for any route where every transition is positively reinforced, the probability of exclusively following this route from some time is positive, by finding an explicit lower bound for this probability — we call this the heavy route. Considering the route that follows the heaviest edge at every step, we can find a

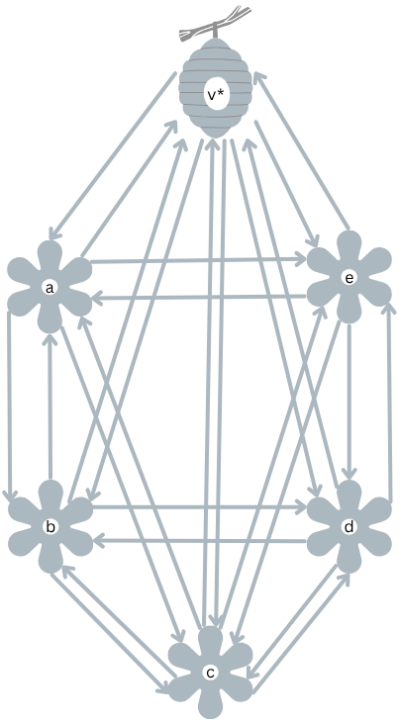

Figure 1.1: Schematic of the discretized environment with a distinct nest v^* and stationary food sources a, b, c, d, e. Every arrow is a directed edge and represents a possible movement decision the bee can make. Associated to every edge is a positive weight $W_{i,j}^{(n)}$ at time n that is used to compute the probability of the bee moving from i to j

universal lower bound on the convergence probability of the heavy route. This then allows us to argue that the number of times the process fails to follow the heavy route is finite and thus the process must converge to it with probability 1. Similar to the reinforced coin example, we then continue to investigate the effect of the learning factor on the distribution of convergent routes. We find that for very large learning factors, the convergence is decided by the first route chosen. We conjecture that for a learning factor close to but greater than 1, the convergent route is the nearest neighbour route with high probability.

1.3.1 Non-backtracking bee foraging bout model definition

Let V be the set of locations the bee may visit, and let v^* be the distinct location that we call the nest. At every time step, the bee must move to a new location. It is not allowed to move to a location it has been to since the last nest visit. At any point in time, the bee may return to the nest. Once the bee is back at the nest, it may move to any location again. The probability of moving from one location to another is proportional to the weight of the specific transition relative to the weight of all possible transitions - these exclude backtracking transitions to locations that the bee has already visited since its last bout. Each time the bee moves from one location to another, the associated weight of this transition is multiplied by a learning factor.

In summary, we assume:

1. With every nest visit, all flowers refill to the same nectar level.
2. When the bee visits a flower, it takes all the nectar present at the flower
3. The bee can remember every location it has visited since its last visit to the nest.
4. The bee only chooses to move to flowers that contain nectar, i.e. it has not visited since its last nest visit and does not revisit flowers it has visited since its last nest visit.
5. The bee chooses flowers with probability proportional to the weight associated with the action of moving from its current location to the new location.
6. The weight of a directed edge increases exponentially by a factor α with every traversal of said edge.

This parallels the traplining heuristic models by Reynolds et al and Dubois et al., with some important differences. In contrast to our simplified model in both of

their models, the bee may revisit the same flower multiple times within the same bout. In Dubois et al., movement is more restricted as bees may not perform immediate backtracks. Another key difference among these models is the timing of reinforcement. In Reynolds et al., reinforcement is route-based, meaning that edge weights only grow once a bout is completed that was of higher or equal quality to the highest quality bout experienced thus far. In Dubois et al.'s model, reinforcement is online, meaning that edge weights change during the bout, based on whether or not nectar is found upon the first visit to a flower during a bout. Since flowers are assumed to refill between bouts, this reinforcement rule, in conjunction with our non-backtracking assumption, amounts to positively reinforcing every flower visit. Note, in the online reinforcement framework, returning to the nest is not positively reinforced, since the nest is usually not considered a food source, and thus edge weights of edges leading back to the nest remain constant in Dubois et al.'s model. For simplicity, we treat nest returns essentially the same as Flower visits in our model and assume they are also always positively reinforced. Meaning that the edge weight of the traversed edge leading back to the nest grows by a factor α with every traversal. We believe this is a reasonable assumption to make, as otherwise returning to the nest and completing the foraging task is not associated with any reward.

While the episodically non-backtracking assumption might seem to restrict the bee's movement significantly, it is not uncommon. Immediate backtracks are often excluded from the analysis of flower visitation data with the argument that these backtracks are generally rare and happen accidentally as a result of reorientation flights [53]. So, we believe it is reasonable to extend this assumption to apply to all backtracks within a bout, as it will also greatly facilitate mathematical analysis and aid in developing intuition and understanding of the mechanics of such episodically edge-reinforced random walks with exponentially growing weights. In summary,

our model is best understood as a strictly non-backtracking version of Dubois et al's online-reinforcement model with added reinforced nest returns. More formally, we define the Epsiodic-Non-Backtracking Exponentially Edge-Reinforced Random Walk as follows. We abstract the environment to be a set of distinct locations of interest. These locations of interest represent flowers where the bee can collect nectar and the bee's nest that it needs to return to every so often to process the collected resource.

Definition 1.9. Let V denote the set of all locations of interest, also called the **environment**. It contains a unique and distinct element $v^* \in V$ referred to as the nest. All other elements are flowers.

We model the bee moving through the environment by randomly choosing where to go next based on where it has already been.

Definition 1.10. The **bee's position** at time t is modelled by the random variable X_t taking values in the environment V.

At the bee moves from one location to another, we call this a transition or following an edge. The terminology "following an edge" is borrowed from Graph theory, where we can imagine the environment as a graph with directed edges going both ways between every pair of locations. It is important that the edges are directed since moving from u to v is not the same as moving from v to u.

Definition 1.11. A transition, also referred to as an edge or action, is an ordered pair $(u, v) \in V \times V$ of locations.

The probability with which a bee makes certain transitions is dictated by weights associated to each directed edge. We collect the edge weight information in a single matrix called the weight matrix.

Definition 1.12. A weight matrix $W \in \mathrm{Mat}_{V\times V}(\mathbb{R}^+)$ is a matrix that assigns a positive weight $W_{u,v}$ to each edge $(u, v) \in V \times V$ for $u \neq v$. The weight of self-loops (v, v) is always 0.

Learning is a way of incorporating past experiences to guide future decisions. We let weights depend on the past experiences of the bee to inform its future decisions, we call this a reinforcement scheme.

Definition 1.13. We say the process is *exponentially-reinforced* when the entries $W_e^{(t)}$ of the weight matrix $W^{(t)}$ are given by

$$W_e^{(t)} = W_e^{(0)} \alpha^{k_t(e)}$$

where

$$k_t(e) = \sum_{1\leq i\leq t} \mathbf{1}[(X_{i-1}, X_i) = e]$$

counts the number of traversals of the directed edge e up to time t. A directed edge can only be traversed in one direction. We assume $W^{(0)}$ to be deterministic. Here, α is called the **reinforcement** or **learning** or **encounter factor**. When $\alpha > 1$, we say a transition is *positively reinforced*. A route is positively reinforced when every transition in it is positively reinforced.

Here, the superscripts of the weights in parentheses denote discrete time steps. Subscripts are directed edges of the underlying graph. We require that the initial weight of any edge is positive $W_e^{(t)} > 0$ and also the reinforcement factor must be positive $\alpha > 0$.

This parallels the reinforcement scheme of the trapline formation models given in [30], [14].

Before explaining how exactly the transition weights inform the bee's movement, we define the crucial concept of a bout. As already mentioned, the bee needs to

return to the nest regularly to process the collected resources. These returns to the nest define episodes referred to as foraging excursions or bouts.

Definition 1.14. Let (T_n) denote the ordered set of times when (X_t) is at the nest, i.e., $T_0 = 0$ and $T_{n+1} = \inf\{t > T_n \colon X_t = a\}$, let $L_n = T_n - T_{n-1}$ and for $i \in \{0, \dots, L_n - 1\}$ let $Y_{n,i} = X_{T_{n-1}+i}$. The sequence $(Y_{n,i})_{i=0}^{L_n-1}$ is called the n^{th} **bout**.

During a bout, transitions are chosen with probability proportional to their associated edge weights. Probabilities are obtained by linearly normalizing weights of allowed transitions (to locations not yet visited during this bout) to sum to unity.

Definition 1.15. The law of $Y_{n,i+1}$ is determined by the weight matrix $W^{(t)}$ and the set of locations the bee may choose from $C = V \setminus \{Y_{n,j}\}_{1 \le j < i}$. Let $a = Y_{n,i}$ be the bee's current position at time t. Then the probability that the bee moves from a to $b \in C$ is defined as

$$\Pr[Y_{n,i+1} = b \mid W^{(t)}, C] = \frac{W^{(t)}_{a,b}}{\sum_{c \in C} W^{(t)}_{a,c}}.$$

Now, the model for the bee's movement through the environment is well-defined as an episodically non-backtracking reinforced random walk where edge weights grow exponentially with the number of directed traversals. Transition probabilities are obtained by linearly rescaling weights associated with directed edges to sum to unity.

We make two useful observations. We realize that the absolute scale of weights is meaningless; all that matters when it comes to computing probabilities is the weight ratios. These change by a constant factor.

Lemma 1.16. *Let $p := \Pr(Y_{n,i} = b \mid \{Y_{n,j}\}_{0 \le j < i})$ denote the conditional probability that the bee chooses to move from location a to b after having already visited locations*

$C = V \setminus \{Y_{n,j}\}_{0 \le j < i}$ during the n-th bout. Assume the bee chooses b on the n-th bout, then when the bee returns to the location a in the $n+1$-th bout, the probability that it will choose b again is given by

$$\Pr(Y_{n+1,i} = b \mid p, \{Y_{n+1,j}\}_{0 \le j < i} = C) = \frac{\alpha p}{1 - p + \alpha p}.$$

Proof. By definition 1.15 , the probability of moving from a to b is given by

$$\begin{aligned} p' :&= \Pr(Y_{n+1,i} = b \mid p, \{Y_{n+1,j}\}_{0 \le j < i} = C) && (1.42)\\ &= \frac{W_{a,b}^{(t')}}{\sum_{c \in C} W_{a,c}^{(t')}} && (1.43)\\ &= \frac{W_{a,b}^{(0)} \alpha^{k_t[(a,b)]}}{\sum_{c \in C} W_{a,c}^{(0)} \alpha^{k_t[(a,c)]}}. && (1.44) \end{aligned}$$

The only edge leaving a for which the weight changed since time t' when the bees last visited a during the n-th bout is the edge (a, b) —or else the walk would have had to backtrack, which is not permitted by definition 1.15. Thus we write

$$\begin{aligned} p' &= \frac{\alpha W_{a,b}^{(0)} \alpha^{k_{t'}[(a,b)]}}{\alpha W_{a,b}^{(0)} \alpha^{k_{t'}[(a,b)]} - W_{a,b}^{(0)} \alpha^{k_{t'}[(a,b)]} + \sum_{c \in C} W_{a,c}^{(0)} \alpha^{k_{t'}[(a,c)]}} && (1.45)\\ &= \frac{\alpha \dfrac{W_{a,b}^{(0)} \alpha^{k_{t'}[(a,b)]}}{\sum_{c \in C} W_{a,c}^{(0)} \alpha^{k_{t'}[(a,c)]}}}{\alpha \dfrac{W_{a,b}^{(0)} \alpha^{k_{t'}[(a,b)]}}{\sum_{c \in C} W_{a,c}^{(0)} \alpha^{k_{t'}[(a,c)]}} - \dfrac{W_{a,b}^{(0)} \alpha^{k_{t'}[(a,b)]}}{\sum_{c \in C} W_{a,c}^{(0)} \alpha^{k_{t'}[(a,c)]}} + 1} && (1.46)\\ &= \frac{\alpha p}{1 - p + \alpha p}. && (1.47) \end{aligned}$$

□

Comparing this to the definition of the law of a reinforced coin, we can already start to see strong parallels, and we could expect similar behaviours from the

two processes. Recall that understanding that the conditional odds ratio changes exponentially in favour of past coin flips was a crucial step in understanding why there is a positive chance that the coin lands with the same side facing up at all large times. In fact, the conditional odds ratio of the bee following a particular transition behaves very similarly.

Lemma 1.17. *Let p and p' be as in the proof of Lemma 1.16, and also assume the bee follows the same route in two consecutive bouts. Let $O = p/(1-p)$ be the conditional odds ratio and $O' = p'/(1-p')$ the odds ratio of following the same transition in the $n+1$-th bout. We have*

$$O' = \alpha O.$$

Proof. From Lemma 1.16 we know $p' = \frac{\alpha p}{1-p+\alpha p}$. Thus, we repeat the calculation from Equation (1.2)

$$\begin{aligned} O' &= \frac{\alpha p}{1-p+\alpha p}\left(1 - \frac{\alpha p}{1-p+\alpha p}\right)^{-1} && (1.48)\\ &= \frac{\alpha p}{1-p+\alpha p}\left(\frac{1-p}{1-p+\alpha p}\right)^{-1} && (1.49)\\ &= \alpha \frac{p}{1-p} && (1.50)\\ &= \alpha O. && (1.51) \end{aligned}$$

□

1.3.2 A positive Lower Bound on the convergence probability

In the following, we assume that all initial transition probabilities are within the open interval $(0,1)$ to avoid having to state some trivial results.

Lemma 1.18. *For any positively reinforced route R of length L, the probability that*

all foraging bouts follow route R has the following lower bound:

$$\Pr(B_n = R \forall n > N) \geq \exp\left(-\frac{\alpha}{\alpha - 1} \sum_{i=1}^{L} \frac{1 - p_{0,i}}{p_{0,i}}\right)$$

where $p_{0,i}$ is the probability of the process following the route R on the ith step during the 0-th bout, conditional on the event that all previous steps have been along R

$$p_{0,i} := \Pr(Y_{0,i} = r_i \mid Y_{0,k} = r_k \forall k < i)$$

and $B_m := (Y_{m,k})_{k=0}^{L}$ is the m-th bout.

Proof. Let $(Y_{n,i})_{i=1}^{L} = B_n$ be the flower visit sequence of bout number n. Let the conditional probability of following R for another step, after having followed it for all previous steps already

$$p_{n,i} := \Pr(Y_{n,i} = r_i \mid B_m = R \forall m < n, (Y_{n,k}) = r_k \forall k < i) = \frac{\alpha^n p_{0,i}}{1 - p_{0,i} + \alpha^n p_{0,i}}.$$

The probability of following R at all times is given by the product of conditional probabilities

$$\Pr(B_n = R \forall n \in \mathbb{N}) = \prod_{n=0}^{\infty} \prod_{i=1}^{L} p_{n,i} = \exp\left(\sum_{n=1}^{\infty} \sum_{i=1}^{L} \log(p_{n,i})\right)$$

Since $p_{n,i} < 1$ implies that $\log(p_{n,i}) < 0$ we have $\sum_{n=0}^{\infty} \sum_{i=1}^{L} \log(p_{n,i}) < \infty$ if and only if $\sum_{n=1}^{\infty} \sum_{i=1}^{L} \log(p_{n,i})$ converges absolutely. This means we may switch the order of summation to get

$$\sum_{n=0}^{\infty} \sum_{i=1}^{L} \log(p_{n,i}) = \sum_{n=1}^{L} \sum_{n=0}^{\infty} \log(p_{n,i}) = \sum_{n=1}^{L} \sum_{n=0}^{\infty} \log(1 - (1 - p_{n,i})) \tag{1.52}$$

Again we apply the approximation $\log(1-x) \geq -\frac{x}{1-x}$ for $x \in [0,1]$ and obtain

$$\sum_{n=1}^{L}\sum_{n=0}^{\infty} \log(1-(1-p_{n,i})) \geq -\sum_{n=1}^{L}\sum_{n=0}^{\infty} \frac{1-p_{n,i}}{p_{n,i}} \tag{1.53}$$

$$= -\sum_{n=1}^{L}\sum_{n=0}^{\infty} \frac{1-p_{0,i}}{p_{0,i}}\alpha^{-n} \tag{1.54}$$

$$= -\sum_{n=1}^{L} \frac{1-p_{0,i}}{p_{0,i}}\frac{\alpha}{\alpha-1}. \tag{1.55}$$

Here, the first equality stems from recognizing that the multiplicative reinforcement scheme changes the conditional odds ratio by a factor of α in favor of the action taken. Substituting back into equation 1.3.2, we have the desired lower bound for the convergence probability □

Again, we consider the route that follows the most likely positively reinforced transitions at every step and show that the probability of following this route in all future bouts has a positive universal lower bound. This will allow us to argue that the process fails to follow this route only a finite number of times almost surely.

Definition 1.19. The **heavy route** denoted by $\hat{R}$ is the sequence of vertices following the most likely positively reinforced transition at every step.

Note that this route is well defined at any point in time, and following it once makes it even more likely to be followed again in the following bout. Thus, the heavy route remains the same route when it is followed repeatedly. We also observe that the probability of following a transition along this route at any point in time is at least as large as the average transition probability $1/(|V|-1)$, where $|V|-1$ is the number of locations the process could transition to. Together with Lemma 1.18, these observations imply the following theorem

Theorem 1.20. *The probability of only following the heavy route for all future bouts*

is positive and bounded below by

$$\Pr(B_n = \hat{R} \forall m > n | \mathcal{F}_n) > \exp\left(-\hat{L}\frac{\alpha|V| - 2\alpha}{\alpha - 1}\right)$$

where $\mathcal{F}_n$ is any history up to the n-th bout, and $\hat{L}$ is the length of $\hat{R}$.

The fact that at any point in time the process will follow the heavy route for all future bouts has the consequence that the process is guaranteed to follow the heavy route at all large times.

Theorem 1.21. *The process follows the heavy route almost surely at all large times:*

$$\Pr(\exists N \in \mathbb{N} \ s. \ t. \ B_n = \hat{R} \forall n > N) = 1.$$

Proof. Let $S_n := \sum_{k \leq n}(1 - \mathbf{1}_{\hat{R}}(B_k))$ count the number of times the process fails to follow the heavy route $\hat{R}$ up to the n-th bout and let $S = \lim_{n\to\infty} S_n$. Since S_n is non-negative, it suffices to show that $E(S) < \infty$. Theorem 1.20 implies that there exists $p > 0$ such that $\Pr(S_{n+1} = s + 1) < (1 - p)\Pr(S_{n+1} = s + 1 \mid S_n = s)$ and

$$E(S) = \sum_{s=0}^{\infty} s \Pr(S_n = s) < \sum_{s=0}^{\infty} s(1 - p)^s p < \infty$$

□

1.3.3 The shape of convergent routes

As with the reinforced coin, we prove that for a large learning factor $\alpha \gg 1$, the convergent route is decided in the first bout.

Theorem 1.22. *As $\alpha \to \infty$ all bouts are equal with probability one*

$$\lim_{\alpha\to\infty} \Pr(B_n = B_0 \forall n \in \mathbb{N}) = 1$$

Intuitively speaking, the first bout changes the transition probabilities so significantly that future transitions are close to deterministic.

Proof. Let $R = (v_i)_{i=0}^{L}$ be a possible route, and let the transition probabilities along the route that was taken in the first bout be denoted as

$$p_i := \Pr(Y_{0,i} = v_i \mid Y_{0,j} = v_j \forall j < i)$$

We condition on the first bout and then apply the approximation from Lemma 1.18 to obtain

$$\begin{aligned}
\Pr(B_n = B_0 \forall n > 0) &= \sum_{R \in \mathcal{R}} \Pr(B_n = R \forall n \geq 0) && (1.56)\\
&= \sum_{R \in \mathcal{R}} \Pr(B_n = R \forall n > 0) \Pr(B_0 = R) && (1.57)\\
&\geq \sum_{R \in \mathcal{R}} \exp\left(-\frac{\alpha}{\alpha - 1} \sum_{i=0}^{L(R)} \alpha^{-1} \frac{1 - p_i}{p_i} \right) \Pr(B_0 = R). && (1.58)
\end{aligned}$$

Where $L(R)$ is the length of the route R and $\mathcal{R}$ is the set of all possible routes the process could take in the first bout B_0.

Note that $\exp\left(-\frac{\alpha}{\alpha-1} \sum_{i=0}^{L(R)} \alpha^{-1} \frac{1-p_i}{p_i}\right) \to 1$ as $\alpha \to \infty$, which gives that in the limit

$$\Pr(B_n = B_0 \forall n > 0) \geq \sum_{R \in \mathcal{R}} \Pr(B_0 = R) = 1. \tag{1.59}$$

□

When reinforcement is very weak, we conjecture that the convergent route is almost surely the initial heavy route. This conjecture is based on the proof for the reinforced coin and simulations of non-backtracking random walks with weak reinforcement.

1.4 Analysis of traplines under more general reinforcement schemes

1.4.1 Stable traplines require Super-Linearly growing Edge Weights

So far, our analysis has been focused on exponentially growing weights. Here, we wish to generalize these findings to account for a wider array of functional forms for the growth of weights. This is important to future modellers who might want to choose different functional forms for how they decide weights should change in response to different experiences. As we will see, the growth paradigm of the weights can influence whether or not there is a positive probability that the simulated animal follows a trapline indefinitely. In fact, we conjecture that when all weights are allowed to grow super linearly fast, i.e. the sum of inverse weights is finite, then the model must converge to a stable trapline. We show rigorously that the probability that a simulation converges to a stable trapline is positive if and only if there exists a route for which all weights grow superlinearly fast, see Theorem 1.23. Furthermore, we show that for very fast-growing weights, the process traverses the same route during all bouts with high probability. We continue to support our conjectures with numerical simulations. We demonstrate that simulations where weights grow exponentially fast, even when backtracking is permitted, converge to a stable trapline, whereas simulations where weights grow linearly with the number of traversals do not converge to stable traplines. Finally, we also show that the set of convergent traplines shrinks as reinforcement becomes weaker but remains super linear, i.e. the reinforcement factor approaches unity from above.

In the following, assume (X_n) is a bout-non-backtracking edge reinforced random walk with transition probabilities given by edge weights. Assume edge weights grow

super-linearly with the number of times the edge was traversed. Let $\omega_e : \mathbb{N} \to \mathbb{R}^+$ be the reinforcement function such that $\omega_e(k)$ gives the weight of the edge e after it has been traversed k times. Since $\omega_e(k)$ grows super-linearly, the inverse of the reinforcement function is summable

$$\sum_{k\in\mathbb{N}} \frac{1}{\omega_e(k)} < \infty$$

For a generalized Polya urn, we know that the urn yields the same colour with probability 1 if and only if the inverse weight function is summable. This is known as the Feller-Lundberg Criterion [22]. Notice that selecting the sequence of first flower visits $\{Y_{n,1}\}_{n\in\mathbb{N}}$ after leaving the nest is essentially a sequence of draws from a generalized Polya urn, where a colour represents the edge that is chosen with weight $\omega_e(k)$. We conclude that it is necessary for a reinforcement function to grow super-linearly for the model to have a positive probability that a trapline emerges. The question remains is the summability of the inverse weight function is a sufficient criterion to conclude that traplines emerge with positive probability.

Theorem 1.23. *Let $(B_n)_{n\in\mathbb{N}}$ be a sequence of consecutive bouts with $B_n = (Y_{n,i})_{i=0}^{L_n}$ starting and ending at the nest. Let $\omega : \mathbb{N} \to \mathbb{R}^+$ be the reinforcement function such that $\omega(k_e^{(t)})$ gives the weight associated with edge e, where $k_e^{(t)})$ counts the number of times edge e has been traversed by the process by time t. The probability that all bouts are the same is positive if and only if the inverse of the reinforcement function is summable.*

$$\Pr(B_n = B_0 \,\forall n \in N) > 0 \iff \sum_{k\in\mathbb{N}} \frac{1}{\omega(k)} < \infty.$$

Proof. The Feller-Lundberg criterion tells us that there exists a random finite time $n_0 \in \mathbb{N}$ such that the first flower visit of each subsequent bout is constant $Y_{n,1} = Y_{n_0,1}$ if and only if $\sum_{k\in\mathbb{N}} \frac{1}{\omega(k)} < \infty$, showing necessity.

To see why summability of the inverse reinforcement function is a sufficient

condition, consider the following lower bound: Let $R = (v_0, \dots v_L)$ be a route and let

$$p_{n,i} := \Pr(Y_{n,i} = v_i \mid B_m = R \forall m < n, (Y_{n,j})_{0 \le i} = (r_j)_{0 \le i}) = \frac{\omega(k_{i-1,i}^{(t(n))})}{\sum_{j \in \Omega} \omega(k_{i-1,i}^{(t(n))})}$$

be the conditional probability of following the route R for another step after having followed it for all previous bouts. Let $r_{n,i} := \frac{p_{n,i}}{1-p_{n,i}}$ be the corresponding odds ratio. Note that

$$r_{n,i} = \frac{p_{n,i}}{1 - p_{n,i}} = \frac{\omega(n)}{\sum_{j \in \Omega \setminus \{i\}} \omega(0)} = c_i \omega(n)$$

where c_i is the inverse of the sum over all weights associated with edges that leave $i-1$-th stop along the route, which have never been traversed as for our assumption, we have left the $i-1$-th stop along the route n times via the edge $i-1, i$. By Lemma 1.5, the probability that all bouts follow the same route $R = (v_0, \dots v_L)$ is bounded below by

$$\Pr(B_n = R \forall n \in \mathbb{N}) \ge \exp\left(-\sum_{n \in N} \sum_{i=0}^{L} 1/r_{n,i}\right) = \exp\left(-\sum_{n \in N} \sum_{i=0}^{L} \frac{1}{c_i \omega(n)}\right)$$

By assumption $\sum_{i=0}^{L} \frac{1}{\omega(n)}$ must converge absolutely and thus $\sum_{n \in N} \sum_{i=0}^{L} \frac{1}{c_i \omega(n)} < \infty$. And thus $\Pr(B_n = R \forall n \in \mathbb{N}) > 0$ □

Remark 1.24. The proof of Theorem 1.23 does not require a non-backtracking assumption. It merely requires that there exists a route where all edges are positively reinforced. Thus, it can also be applied to episodically edge-reinforced random walks that are allowed to backtrack as long as every edge weight along the route grows superlinearly fast.

In Theorem 1.23, we establish a general criterion for both backtracking and non-backtracking- episodically-edge-reinforced random-walks to have a positive probability

to exhibit traplining behaviour. The decisive criterion here is that edge weights need to grow super linearly. When weights grow any slower than that, i.e. the inverse weights are not summable, then the probability that a bee modelled in this way follows the same route indefinitely is 0. On the other hand, if weights grow super-linearly fast, i.e. inverse edge weights are summable, there is a positive probability that the bee follows the same route at all large times. Extrapolating from the Feller-Lundberg criterion, which says that an urn permitting weights of a certain colour to grow super-linearly, the urn will yield the same colour ball almost surely, we conjecture that also for episodically reinforced random walks, the same route is chosen eventually, always with probability 1 if and only if edge weights may grow super-linearly fast. However, we show a weaker result, that the probability of following the same route eventually always is positive if and only if weights may grow super-linearly.

In biological terms, this means that if we want a model to have the potential to model the formation of stable traplines, then weights need to be allowed to grow super-linearly fast. It also means that if it turns out a model where weights grow linearly fast fits real movement data better than models where weights grow superlinearly fast, the sound biological and scientific conclusion would have to be that bees do not converge to stable traplining behaviour and even under extremely stable environmental conditions. On the other hand, if a model with superlinearly growing weights is the best fit to some data, it would be reasonable to conclude that bees do converge to following the same trap line at all large times, given that food sources are reliable enough.

1.4.2 Simulation of non-backtracking bouts

We simulate the non-backtracking process on an elongated hexagon similar to the feeder array used in [53]. Unless mentioned otherwise the initial transition probabilites

are inversely proportional to the squared distance between locations not yet visited. This means that the weights are initialized to

$$W_{i,j}^{(0)} = \frac{1}{\text{dist}(i,j)^d} \tag{1.60}$$

with the distance exponent $d = 2$ by default. We can think of the distance exponent as the inherent preference of a bee to move to a closer location. when d is large, the preference to move to closer locations is very strong.

To illustrate some of the results given above, we simulated the non-backtracking episodically edge-reinforced random walks. Meaningful investigation also requires visualization techniques that can reliably reveal the patterns hidden in the flower visitation sequences generated by the models To visualize location visitation sequences, we borrow tools such as chaos game representations (CGRSs) from DNA sequence visualizations and recurrence plots from dynamical systems.

The simplest visualization technique was to simply plot the tracks traced out by the process. This highlights the progression of how consecutive bouts inform one another.

While these plots are a very direct and easy to understand way to visualize the progression of consecutive bouts, it is difficult to visualize longer sequences of feeder visits. To over come this problem we turn to recurrence plots.

Recurrence plots are a visualization technique used in modern non-linear data analysis to detect deterministic behaviour in time series [34]. The idea is, given a time series $x_1, \ldots, x_N$ we plot a binary similarity matrix $M \in \text{Mat}_{N\times N}([0,1])$ where

$$m_{i,j} = \begin{cases} 1 & \text{if } x_i = x_j, \\ 0 & \text{if } x_i \neq x_j. \end{cases}$$

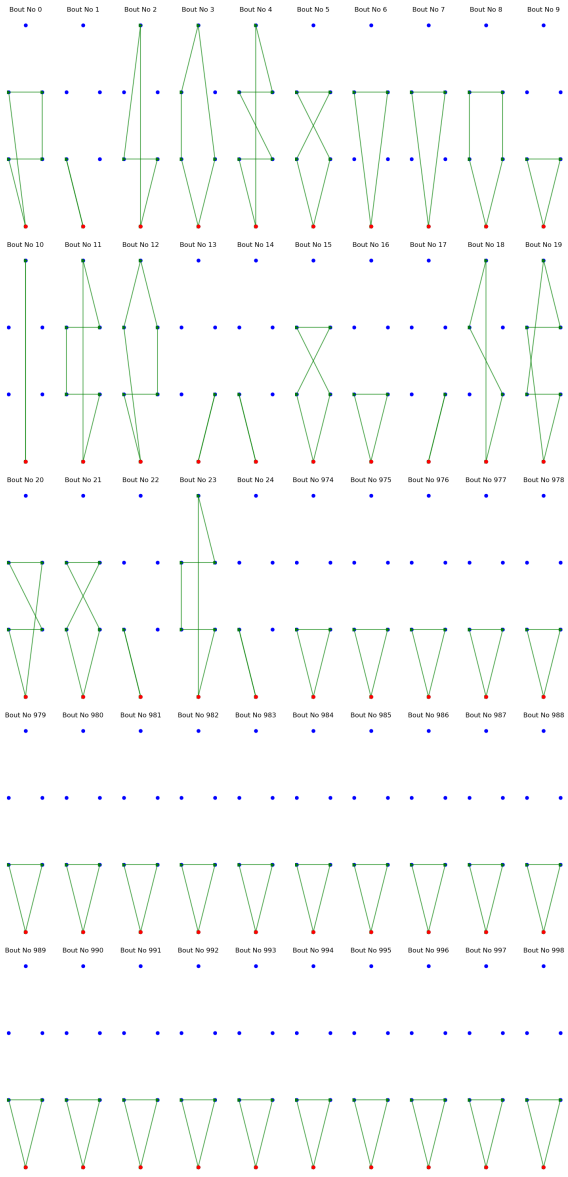

Figure 1.2: First and final 25 of a total of 100 simulated bouts of the non-backtracking episodically edge-reinforced random walk where initial edge weights grow by a factor of 1.02 after every transition.

Figure 1.3: First and final 25 of a total of 100 simulated bouts of the non-backtracking episodically edge reinforced random walk where initial edge weights grow by a factor of 10 after every transition.

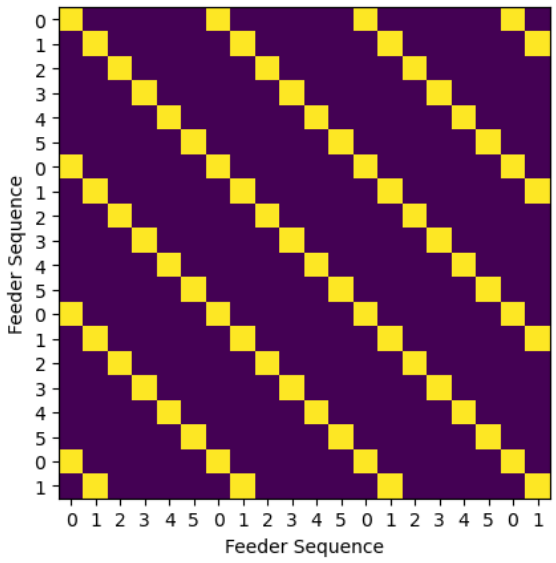

Figure 1.4: A perfectly periodic sequence produces a recurrence plot with diagonal lines that are separated by a number of cells equal to the period of the sequence.

Zero entries are plotted as white cells, and non-zero entries are black. Figure 1.4 shows an example of perfect periodic deterministic behaviour, i.e. traplining. The long diagonal lines indicate that once the system reached a certain state, it followed the same trajectory. If the system reaches the same state after, say, six time steps, another diagonal line will appear in the recurrence plot, shifted vertically by six cells.

On the other hand, when a system behaves completely randomly, the recurrence plot displays a lack of diagonal lines. An example of this is shown in Figure 1.5.

In Figure 1.6 we can clearly see that even in the beginning stages of the process simulated bees behave more deterministically than a randomly moving bee as diagonal line features appear frequently, representing repeated sequences. The more strongly reinforced process illustrated in Figure 1.7, fixates to a trapline after the first two bouts resulting in a large square of diagonal lines.

To be able to visualize even longer sequences of feeder visits we use chaos game representations (CGRs). Chaos game representations are a visualization technique commonly used to visualize structures in DNA sequences and are used to reveal

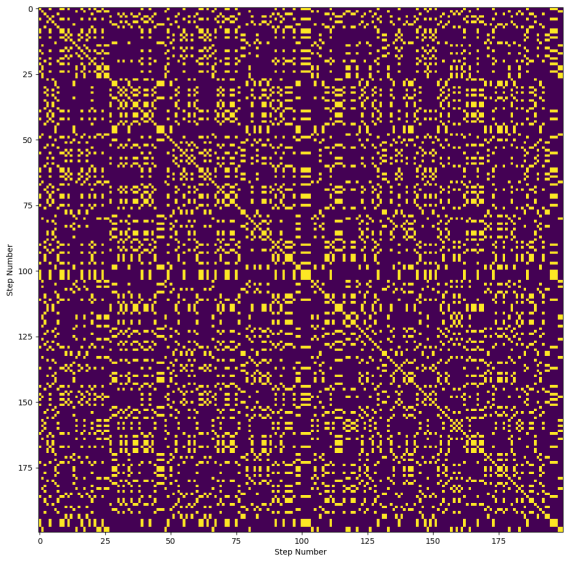

Figure 1.5: A recurrence plot of a sequence of 200 uniform draws out of a set of size 6 demonstrates a lack of structure.

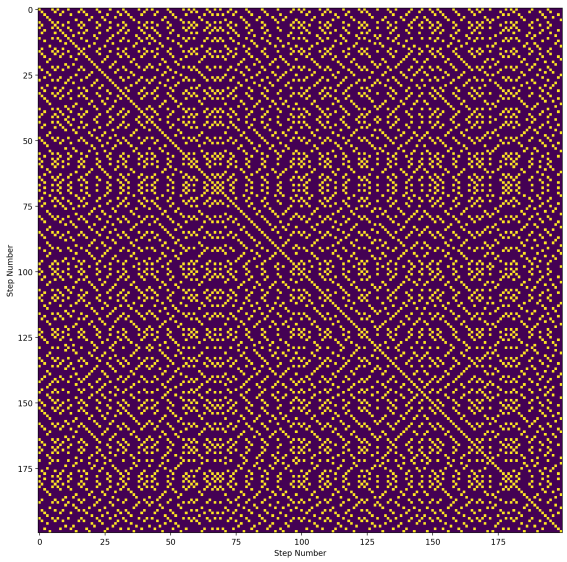

Figure 1.6: Recurrence plot of the first 200 steps of a single simulation of the non-backtracking episodically edge reinforced random walk, where initial edge weights grow by a factor of 1.02 after every transition.

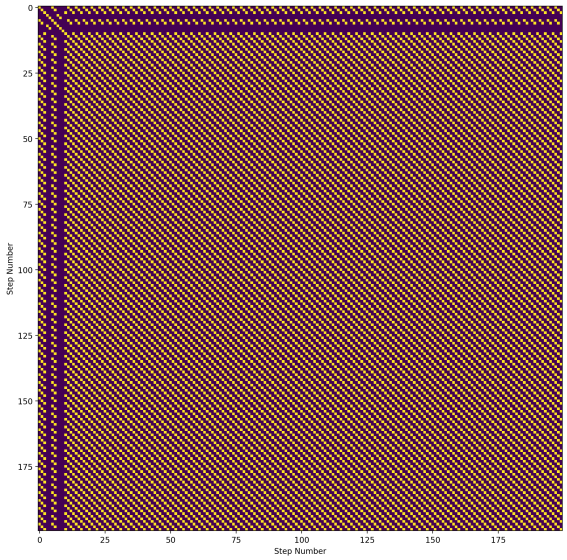

Figure 1.7: Recurrence plot of the first 200 steps of a single simulation of the non-backtracking episodically edge reinforced random walk, where initial edge weights grow by a factor of 10 after every transition.

structure and periodicity in sequences. A sequence is visualized in a regular polygon where each vertex represents an element of the alphabet in which the sequence is written. When dealing with DNA sequences, these are the amino acids adenine, cytosine, guanine, and thymine, A, C, G, T for short. For our purposes, each vertex will represent a feeder location or the nest. We number the locations $0, \ldots, N$ starting with the nest. A sequence of points is drawn in the interior of the regular polygon. Starting at the center, the next point is the halfway point between the current point and the vertex of the regular polygon corresponding to the next letter in the sequence. For a DNA sequence, this process is visualized in Figure 1.8 [32]. We use colour to visualize the progression of time; lighter colours are entries at the beginning of the sequence, and later points are darker colored. Perfectly periodic sequences converge to a set of t points, where t is the period of the sequence. In contrast, random sequences lacking periodicity will tend to scatter. There is a limitation to using CGRs with alphabets containing more than 4 letters. Namely, CGR-points from a uniformly random sequence will not be uniformly distributed in

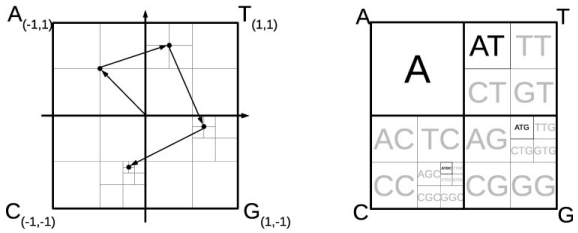

Figure 1.8: A visual explanation for how to generate a chaos game representation for a DNA sequence. The left panel shows a visualization of the sequence A, T, G, C starting in the center. A point is placed halfway towards the corner corresponding to the next letter in the sequence. The right panel shows how every region in the Square corresponds to a specific subsequence of letters. So that the number of times the letter A appears is exactly the number of points in the top left quadrant labelled A, the number of times the word AT appears in a sequence corresponds to the number of points in the region labelled AT. This figure was taken from [32].

a regular polygon that is not a square. As a baseline, we provide two visualizations of sequences generated by a uniform measure on $0, \ldots 6$ in Figure and the null model for the environmental configuration where the next location is chosen proportionally to the inverse-squared distance.

We notice stronger reinforcement leads to faster convergence, as shown in Theorem 1.22. This observation is further supported in Figures 1.12, 1.13 when considering the distribution of convergence times under different encounter factors. We declare the process has converged to a stable trapline when the route has been repeated 50 times in a row.

Considering the distribution of convergent traplines, we see that under weaker reinforcement, the probability seems to concentrate on a smaller subset of traplines compared to stronger reinforcement, as illustrated in Figures 1.14, 1.15. Interestingly enough, when we reinforce every transition, a short trapline emerges under weak reinforcement, where the final transition is not the nearest neighbour transition. This changes as we increase the initial preference to move to closer locations. Increasing the Distance exponent from two to five shows a clear preference for the nearest neighbour route, similar to the reinforced coin for which we know that it

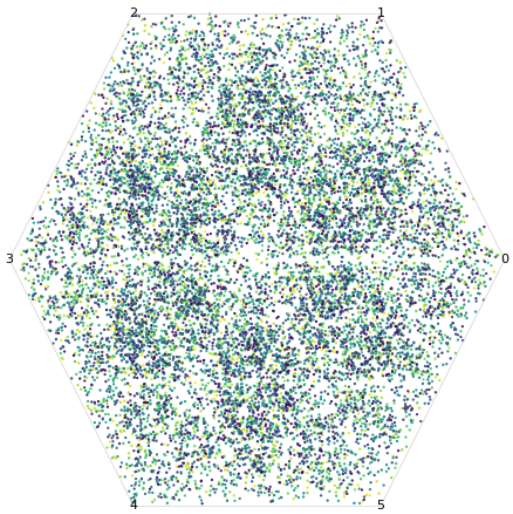

Figure 1.9: A chaos game representation of a sequence of 15000 uniform draws from $0, \ldots, 5$.

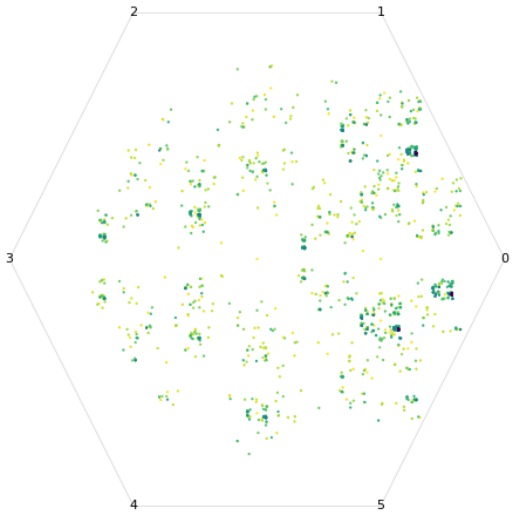

Figure 1.10: Chaos game representation of a single simulation of the non-backtracking episodically edge-reinforced random walk, where initial edge weights grow by a factor of 1.02 after every transition. The simulation was run for 1000 bouts. Darker dots represent steps taken later on in the simulation.

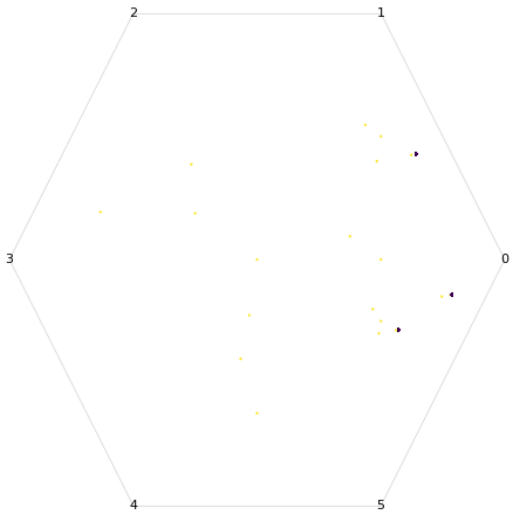

Figure 1.11: Chaos game representation of a single simulation of the non-backtracking episodically edge-reinforced random walk, where initial edge weights grow by a factor of 10 after every transition. The simulation was run for 1000 bouts. Darker dots represent steps taken later on in the simulation.

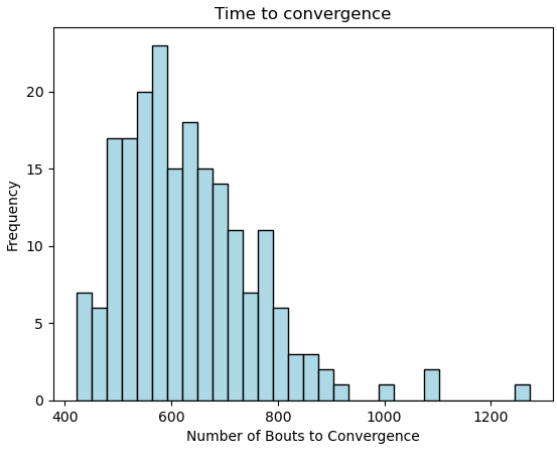

Figure 1.12: Number of bouts until a single route has been repeated 50 times in a row for an episodically edge reinforced random walk where the edge weight of every traversed edge is multiplied by a factor of 1.02.

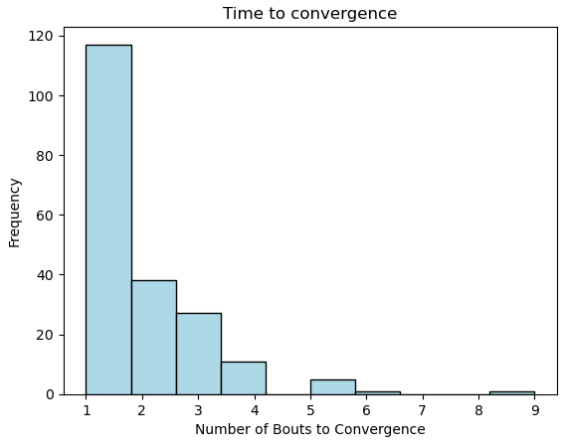

Figure 1.13: Number of bouts until a single route has been repeated 50 times in a row for an episodically edge-reinforced random walk where the edge weight of every traversed edge is multiplied by a factor of 10

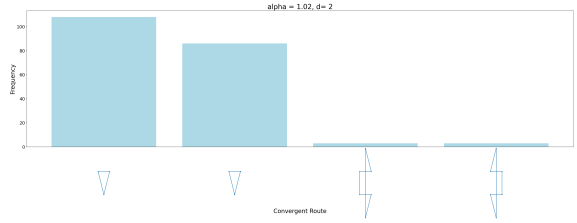

Figure 1.14: Distribution of traplines under weak reinforcement $\alpha = 1.02$. Simulated bees were not allowed to visit the same feeder twice within the same bout, and every transition was reinforced. A total of 200 realizations of this process were simulated.

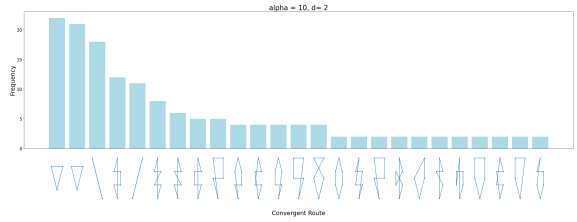

Figure 1.15: Distribution of traplines under strong reinforcement $\alpha = 10$. Simulated bees were not allowed to visit the same feeder twice within the same bout, and every transition was reinforced. A total of 200 realizations of this process were simulated.

will always converge to the initially preferred side when reinforcement is extremely weak, as we can see in Figures 1.16 and 1.17. This is a surprising outcome. Based on Theorem 1.7 , we would expect that under weak reinforcement, the process should always converge to the initially more likely outcome. The experiment illustrated in Figure 1.14, however, suggests that this heuristic explanation as to why we might always expect the closer feeder to be chosen under weak reinforcement is not exactly correct. We can see that the process clearly prefers to choose a route that is not the nearest neighbour route, even though reinforcement is fairly weak. A potential explanation for this is that two characteristics increase a route's probability of being the convergent route in a weak reinforcement paradigm. Firstly, we note that a route with fewer transitions has a higher convergence probability. To see this more clearly, consider the statement of Lemma 1.18. The lower bound on the convergence probability decreases with the number of transitions L in the route. On the other hand, we know that in a weak reinforcement paradigm, nearest neighbour transitions have a higher probability to become part of the convergent trapline, since the initial transition probabilities are inversely proportional to the distance taken to the power of the distance exponent. The strength of these two drivers of trapline selection varies with the distance exponent. In the example at hand this means the following. The fact that the final transition of the triangle-shaped route is very close to the length of the shortest transition makes it so that there is no clear preference of the route based on the initially preferred transitions when the distance exponent is small. The number of transitions has a larger effect on the convergence probability. However, when we increase the distance exponent and with that the preference for nearest neighbour transitions following the nearest neighbour route indefinitely becomes more likely than following a route with fewer transitions.

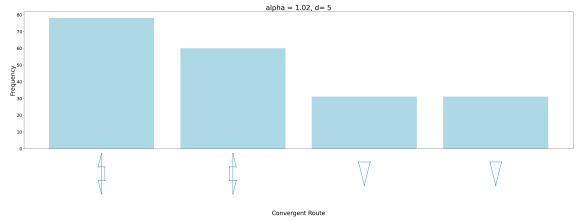

Figure 1.16: Distribution of traplines under weak reinforcement $\alpha = 1.02$ and strong preference for close locations $d = 5$. Simulated bees were not allowed to visit the same feeder twice within the same bout, and every transition was reinforced. A total of 200 realizations of this process were simulated.

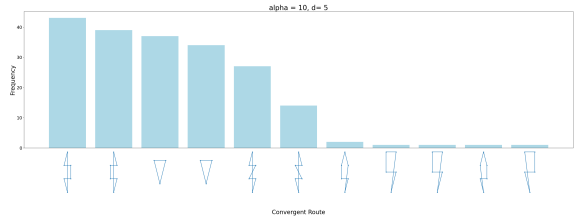

Figure 1.17: Distribution of traplines under strong reinforcement $\alpha = 10$ and strong preference for close locations $d = 5$. Simulated bees were not allowed to visit the same feeder twice within the same bout, and every transition was reinforced. A total of 200 realizations of this process were simulated.

1.4.3 Backtracking bee foraging bouts

While bees tend to avoid revisiting the same flower twice within a single bout, it is by no means uncommon [53]. In fact, learning in bees seems to manifest as a decrease over time in the number of flower revisits [30]. Thus, if we want to understand conditions under which traplines may emerge, we need to understand how these models behave when bees are allowed to revisit the same flower multiple times within the same bout.

In our simulations, we handle flower revisits similarly to Dubois et al. When a flower is visited for the second time within the same bout, we assume that all the nectar has already been taken. When the bee does not encounter nectar at its destination, we assume that, as a response to such a non-encounter, the bee will tend to avoid retracing this step that did not lead to any physical reward. Thus, instead of multiplying the weight by a number greater than 1, weights decay by a non-

encounter response factor < 1. Note that real bees might not react this way when not encountering nectar at a flower, as simply finding a flower and the promise of future rewards may already be enough to warrant a return in the future. Furthermore, we adopt reinforced nest returns by using the same criterion as Reynolds et al, of requiring a high route quality to warrant the edge weight of the edge leading back to the nest to grow and decay otherwise

We formalize this idea by defining the encounter and non-encounter.

Definition 1.25. The quality of a route is defined as the ratio of the squared number of resources collected during a bout and the total distance travelled during this bout

$$q_b = \frac{(|\{y_{b,s}\}_{s=0}^{L_b}|)^2}{\sum_{1\leq s\leq L} \operatorname{dist}(y_{b,s-1}, y_{b,s})}$$

Note, this is the same definition for route quality as used in [14]

Definition 1.26. We say a traversal of an edge (i, j) is an encounter if it is the first time within this bout that the process visits j, or $j = v^*$ is the nest, and the route quality is at least as high as any route quality previously achieved; otherwise, it is a non-encounter.

We distinguish between two responses, the encounter and non-encounter responses, which change the weight matrix:

Definition 1.27. Weights change exponentially as a response to encounters and non-encounters

$$W^{(n+1)}_{y_{b,s-1},y_{b,s}} = \begin{cases} (1+\ell)W^{(n)}_{y_{b,s-1},y_{b,s}} & \text{if } y_{b,s-1}, y_{b,s} \text{ is an encounter,} \\ (1+a)W^{(n)}_{y_{b,s-1},y_{b,s}} & \text{if } y_{b,s-1}, y_{b,s} \text{ is a non-encounter.} \end{cases}$$

The parameter ℓ is called the *encounter increment* and a is called the *non-encounter* or *abandon increment*. $(1 + \ell)$ is the reinforcement, learning or encounter factor,

$1 + a$ is the *abandon* or *non-encounter factor*

For our simulations, we assume $\ell > 0$ and $a = \frac{1}{1+\ell} - 1 < 0$ so that the effects of an encounter and a non-encounter cancel each other out. So $1 + \ell$ corresponds to the learning factor α from the previous section. If the response increment is positive, this implies a tendency of the process to return. If the response increment is negative, this implies avoidance.

We know that whenever edge weights can grow superlinearly with the number of traversals, there is a positive probability for a stable trapline to emerge by Theorem 1.23. For bout-non-backtracking models, we were able to show in Theorem 1.21 that when weights grow exponentially fast with the number of traversals, the simulated bee is guaranteed to follow the same route at all large times. The argument hinges on the fact that there is a universal lower bound, such that from some time onward, the simulated bee follows the same route. When we allow backtracking, this universal lower bound is no longer valid since traversing the same edge twice in the same bout causes the edge weight to shrink, potentially negating the positive reinforcement of the first traversal. Still, we believe that, similarly to non-backtracking EERRWs, stable traplines emerge under similar conditions in backtracking EERRWs. More precisely, we conjecture for backtracking EERRWs

1. If edge weights grow exponentially fast with the number of positive traversals, then a stable trapline emerges almost surely.

2. When reinforcement is very weak, the reinforcement factor is close to 1, $\alpha = 1 + \epsilon$ with $0 < \epsilon \ll 1$, then the emergent stable trapline is almost surely the nearest neighbour route.

1.4.4 Simulation of backtracking bouts

As a representative case for weak reinforcement, we chose an encounter factor of $1 + \ell = 2$ and correspondingly a non-encounter factor of $1 + a = 0.5$. Accordingly, the encounter and non-encounter increments are 1 and $-.5$. As a representative case for strong reinforcement, we consider an encounter factor of 10, a non-encounter factor of $1/10$. Accordingly, the encounter and non-encounter increments are $\ell = 9$ and $a = 0.9$.

Environment Plots

Figure 1.18 and 1.19 show example trajectories for the null and online learning process with exponentially growing weights, respectively. Comparing these two realizations, we can see how 1.19 gradually begins to resemble preceding bouts, slowly converging to a trapline, whereas the null model remains random.

To further investigate the shape of convergent traplines, we simulate 200 bees under with an encounter increment of $\ell = 1$, a non-encounter increment of $a = -0.2$ and a distance exponent $d = 2$ to simulate the weak exponential reinforcement case of parameter configurations for 1000 bouts and plot traplines that the process converged to.

Quantitative analysis of simulated Tracks

We simulated 200 bees foraging on 5 flowers for 1000 bouts. The flowers are arranged in an elongated hexagon. As previously, a nest return is positively reinforced when a route is better than or equally as good as previously experienced routes. We say a bee converged on a trapline if the bee followed the same route for the final 50. The simulated bee followed the same route. We found that almost all simulation runs converged to a trap-line within the first 1000 bouts. Convergent trap-lines were

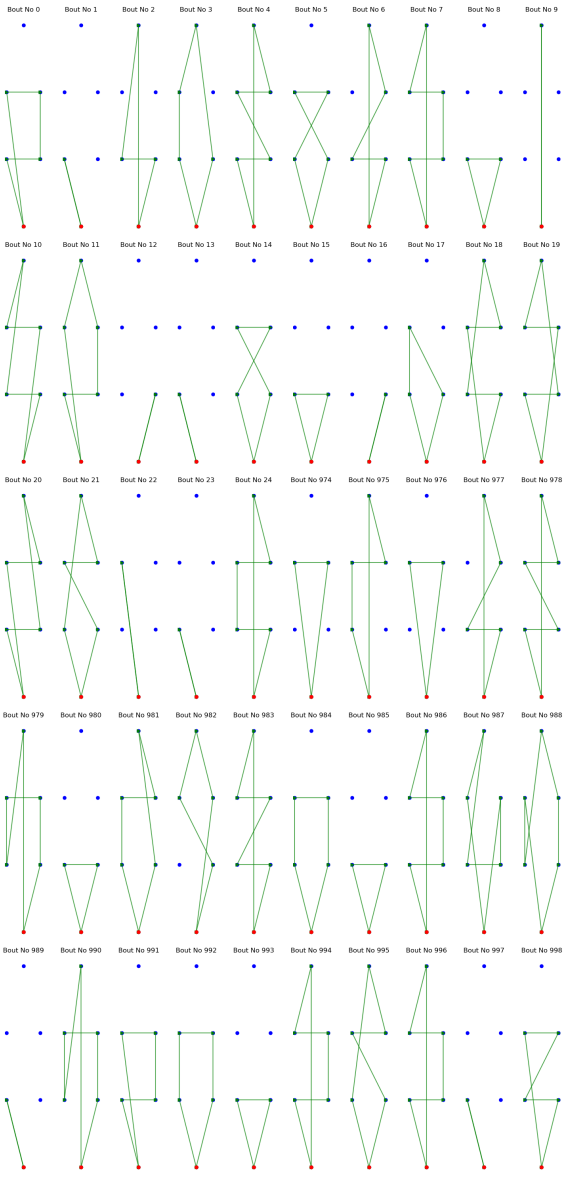

Figure 1.18: Example realization of the null model: a bee moving from feeder to feeder with probability proportional to the inverse square distance. Blue dots are feeder locations, the red dot is the nest, start and end point of every bout.

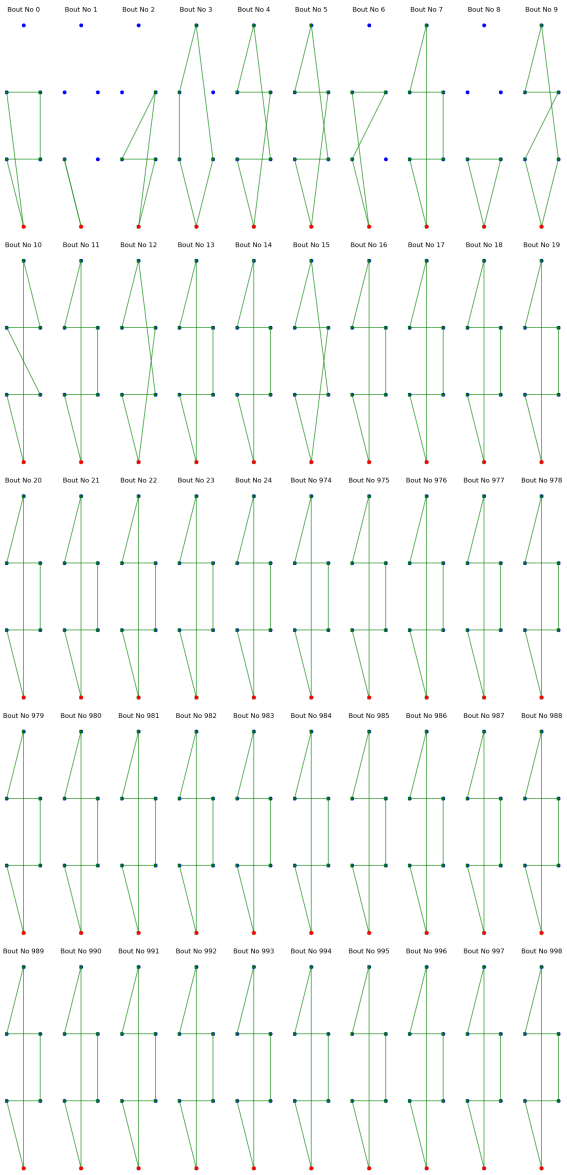

Figure 1.19: Example realization of a trapline model allowing backtracking within a bout. The encounter increment is set to $\ell = 1$, and the non-encounter response increment is set to $a = -1/2$. Weights grow and decay exponentially fast with rates $\alpha = 1 + \ell$ and $\beta = 1 + a$ with the number of times a food is encountered or not after traversing an edge, respectively. Edge-weights are initialized to be proportional to the inverse square distance. Returning to the nest after completing a bout with route-quality at least as high as the best route-quality achieved thus far triggers the same response as encountering food at the end of the edge leading back to the nest. When a lower route quality is achieved, the non-encounter response is triggered.

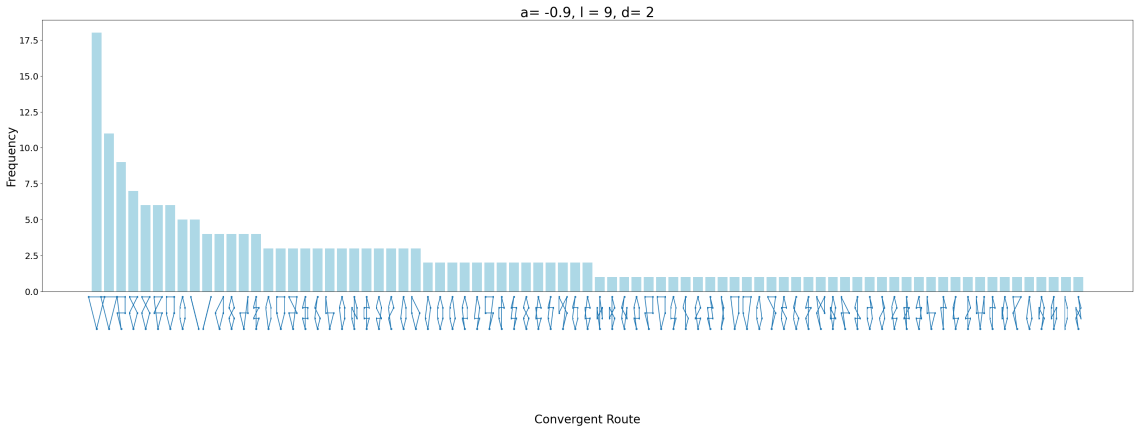

Figure 1.20: Frequencies of convergent routes when backtracks are allowed under a strong reinforcement paradigm. A total of 200 simulations were run on the elongated hexagon, with an encounter increment of 9 and a non-encounter increment of $-.9$ and a distance exponent of $d = 2$. Convergence was declared if the final 50 of 10000 all followed the same route, as shown on the x-axis. To reach 200 convergent simulations, 201 simulations had to be run. There are 81 distinct emerging traplines.

always non-backtracking and visited at least 2 and most frequently all 5 flowers. The exact shape of the convergent trapline depends on the initial transition probabilities and the learning factor.

Smaller learning factors and a larger distance exponent lead to a narrower array of convergent routes. In simulation runs with a learning factor close to one and a large distance exponent, the convergent route was the nearest-neighbour route. Increasing the encounter increment from 1 to 9 yielded a more diverse set of convergent routes, as seen in Figures

From Figure 1.23, we can see the emergent traplining behaviour as darker points, representing later visits, are more concentrated on a small region, indicating that the process converged to a strict periodic visitation sequence. Comparing Figures **??** and 1.23 , we can see that the simulated bee has greatly reduced its variance in route choice as the point cloud in Figure **??** is much more scattered than Figure 1.23.

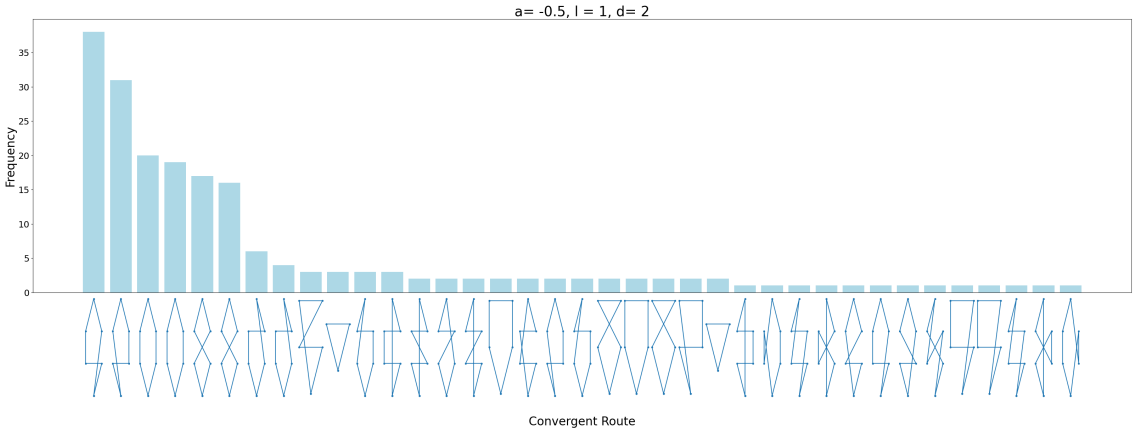

Figure 1.21: Frequencies of convergent routes when backtracks are allowed under a weak reinforcement paradigm. A total of 200 simulations were run on the elongated hexagon, with an encounter increment of 2 and a non-encounter increment of $-.5$ and a distance exponent of $d = 2$. Convergence was declared if he final 50 out of 10000 all followed the same route, shown on the x-axis. To reach 200 convergent simulations, a total of 218 simulations had to be run.

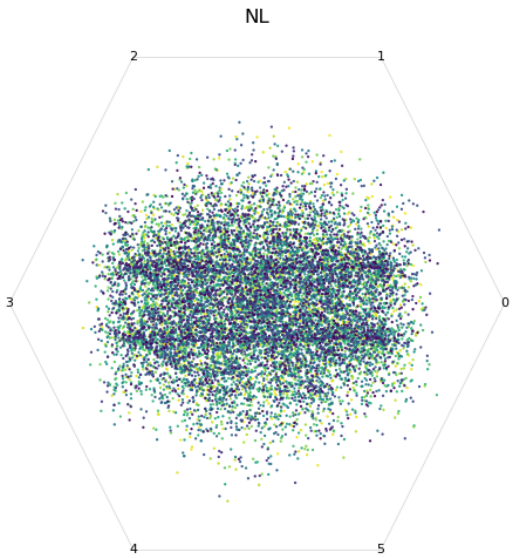

Figure 1.22: A chaos game representation sequence of approximately 15000 points generated by the null model NL where feeder locations are chosen with probability inversely proportional to the squared distance to the current location. Locations are arranged in an elongated Hexagon.

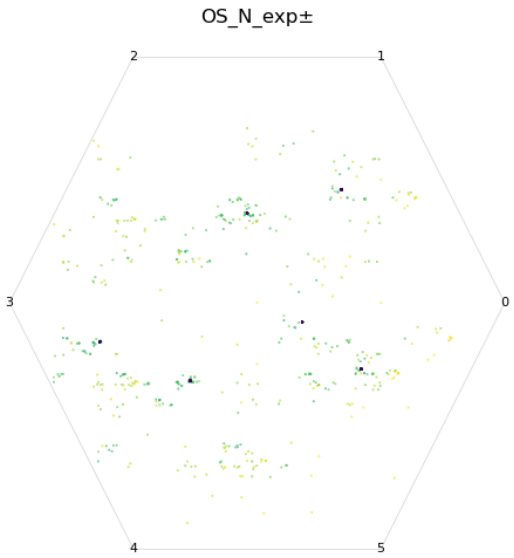

Figure 1.23: A chaos game representation of a location visitation sequence generated by the simplified online learning model where weights grow and decay exponentially at a rate of 3 and 1/2 with the number of encounters and non-encounters. Each vertex of the polygon represents a location of interest in the environment, where 0 represents the nest all other locations are flowers. Lighter-colored points represent location visits early on in the sequence darker-colored points are later in the sequence. The dividing rate is 1/2.

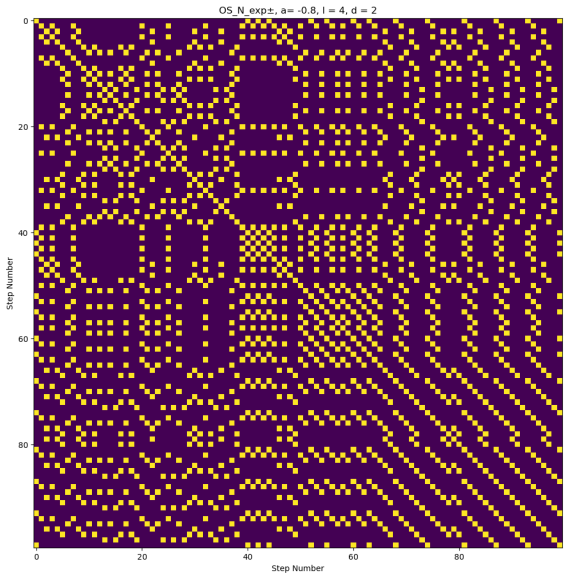

Figure 1.24: Recurrence plot of an episodically edge reinforced random walk with balanced encounter and non-encounter response strengths of 5 and 0.2, respectively.

Recurrence plots

We can see traplines emerge around the 60-th step, in the model when the encounter factor is set equal to 5 and the non-encounter increment is set to 1/5 , as diagonal lines show from step 60 onward in Figure 1.24

1.5 Dicussion

1.5.1 Goals and summary

In this chapter, we set out to understand the conditions under which reinforcement-based models of foraging can produce stable traplines - in the sense of repeating the same route at all large times. In particular, we investigate models present in the literature as given by Lihoreau et al, Reynolds et al and Dubois et al [30, 44, 14]. In essence, these models are episodically edge-reinforced random walks (EERRW). By investigating simplified and generalized versions of these models, we derive conditions that determine when traplines can emerge from models of this

type. Answering the question of when traplines can emerge from simple stochastic movement rules is a fundamental step in understanding memory-informed movement, a thriving area of research. Going beyond animal movement research, the theorems presented here apply to any system that can be framed as an episodically reinforced random walk on a complete graph. This includes a wide array of reinforcement learning tasks, so these results could also be of interest to researchers in artificial intelligence or reinforced random walks [48, 33].

Despite this broad field of applications, we will keep the discussion focused on the original question that we set out to answer: when can stable traplines emerge in bumblebee foraging behaviour? We consider a trapline stable if, from some point in time onward, the bee follows the same route without deviating from it. Furthermore, different hypotheses about the shape and the decision rules that could give rise to traplines have been put forward, the *order of encounter hypothesis*, proposed in [25] and the seemingly competing *nearest-neighbour-hypothesis* hypothesis [2]. However, both of these models fail to account for the inherent stochasticity of animal movement. Here we argue that the EERRW-models developed by [44, 30, 14] are a compromise between these two hypotheses as initial movement favours nearest neighbour transitions, and later on in the process the bee retraces its steps. Most interestingly, we provide evidence that the shape of the convergent route is greatly determined by reinforcement strength. When reinforcement is very strong the bee will fixate on a single route very quickly, mainly determined by the order in which feeders were encountered, closely resembling the first bouts with high probability as seen in Theorem 1.22. However, when the encounter increment is small but positive, reinforcement is weak and the process is more likely to converge to the nearest neighbour route. These predictions are based in the asymptotic analysis of the reinforced coin and where we showed that in the edge case of an encounter factor approaching unity from above the coin converges to the initially heavier side as seen

in Theorem 1.7. This trend continues in the simulations made in Section 1.4.4.

There has been discussion about the importance of negative reinforcement, i.e. avoidance of non-beneficial feeding sites [14]. We find that in the models negative reinforcement does not play a critical role in whether or not the process can converge to a stable trapline, as non of the convergence theorems depend on the non-encounter factor. We show that edge weights must grow superlinearly for stable traplines to emerge. Traplines cannot emerge when weights may only decay. However, the non-encounter factor might still play a role in the shape of convergent traplines. A negative non-encounter increment has the consequence that stable traplines must not visit the same vertex twice. Moreover, we believe that the relaitive strengh of encounter and non-encounter facorrs might further influence trapline shape and time to convergence. In our simulations, to balance the effects of encounter and non-encounter reinforcement events, we chose the non-encounter factor to be exactly the inverse of the encounter factor in all of our simulations. As part of future work, it would be interesting to see the effect of unbalanced encounter to non-encounter factors.

1.5.2 Conditions for the emergence of stable traplines

Most fundamentally, we point out conditions under which traplines can emerge at all. We show that the emergence of traplines requires super-linear reinforcement of edge weights (1.23). Reframing the reinforcement scheme of the reinforced random walk models in the literature, we show that traplines can emerge in these models with positive probability. Making the simplifying assumption that the model bees do not visit any flower twice within the same bout, we prove that traplines must emerge when the bees can collect enough nectar so that the nest return is considered a success. This finding supports the possibility of stable traplines consistent with the nearest-neighbour-hypothesis and the order of encounter-hypothesis, but rules

out the idea that reinforcement with only linear growth could stabilize routes. With multiplicative or exponential reinforcement, convergence to a single trapline is inevitable provided that backtracking is not allowed, as shown in Theorem 1.21. This means that a bee, according to this model, will inevitably be stuck in a trapline, moving from flower to flower in a seemingly deterministic fashion provided flowers refill after every bout. Removing the non-backtracking assumption, allowing the bee to revisit the same flower multiple times within the same bout, does not qualitatively change the behaviour as long as we limit the number of times an edge can be positively reinforced within a single bout to be finite. If we were to allow an edge weight to grow to infinity within a single bout, there would be a positive probability that the bout becomes infinitely long, and the bee could be stuck in an infinite loop. Simulations, in which an edge is positively reinforced when food is found at the destination, i.e. the first time the location is visited during this bout or returns to the nest after a 'successful' bout in which it has collected enough nectar by visiting a minimum number of flowers, suggest traplines must emerge even when backtracking is permitted. In support of this conjecture, we can prove that the probability for a stable trapline to emerge is at least positive when edge weights can grow superlinearly, i.e. the inverse edge weights after consecutive positive traversal are summable.

Traplines can not emerge solely as a result of avoidance. Avoidance is modelled as negative reinforcement. We show that in non-trivial cases, negative reinforcement cannot lead to the emergence of traplines.

This work provides a baseline explanation of how memory is used in animals that follow stable traplines, repeating the same foraging route at all times after an initial exploration phase. Furthermore, this suggests that if we do see strict convergence to a single foraging route that is followed from some point onwards, with food sources refilling in the episodic fashion assumed in the theorems, any reinforcement

model should use superlinearly growing weights. Also, notice that the non-encounter factor is effectively irrelevant to the qualitative asymptotic behaviour. The decisive parameter in these models is the learning factor, which governs the growth of edge weights. To be more concrete, we show that traplines cannot emerge when the agent only attempts to avoid certain locations.

1.5.3 Determinants of trapline shape

We argue that the EERRW type model with exponential reinforcement is best understood as a stochastic trade-off between the order of encounter hypothesis and the nearest neighbour hypothesis, since traplines associated with either of these hypotheses can emerge. When reinforcement is weak, the learning factor is close to unity, the trapline is more likely to be the nearest neighbour route. On the other hand, when reinforcement is very strong and $\alpha \gg 1$ the model predicts that the stable trapline is the order-of-encounter-route with high probability. Simulations and some theoretical work presented in section 1.4 suggest that these results generalize even to models that allow a location to be visited multiple times within a single bout.

1.5.4 The mathematical context

Today, the literature on reinforced random processes is vast, and yet there are still many unexplored models. For example, at least to our knowledge, the literature on edge-reinforced random walks (ERW) is largely focused on simple graphs with few exceptions treating directed traversals of simple edges, while results for edge-reinforced random walks on directed graphs seems to be lacking entirely. Similar to the model treated here and yet quite different is the Ant RW by Erhard et al. [16]. In their model, traversing an edge from x to y increases the weight by a factor $\exp(\beta)$ and decreases the probability of traversing the edge in the reverse direction from y to x by a factor of $1/\exp(\beta)$. This has the consequence that if a simple edge

is crossed as many times from x to y as it has been crossed from y to x , the weight of traversing the edge in either direction is the starting weight. In contrast, in the model presented here, traversing the edge (x, y) does not change the weight of (y, x). Similarly to our result, Erhard et al. find that the Ant RW localizes on a cycle in finite graphs that are not trees. We show that, instead of the walk localizing to a single edge as is the case for strongly reinforced ERRWs on simple graphs [10], the strongly reinforced ERRW on a digraph may localize on a cycle. Also new to the literature is the idea of episodic reinforcement. Foraging bouts naturally motivate the idea of episodic reinforcement, where reinforcement and action choice is very much non-Markovian, making this analysis rather difficult.

1.5.5 Limitations and future work

The models discussed here, even though insightful, also have their biological and mathematical limitations. Probably the most significant limitation from a mathematical and biological perspective is that most of this work is concerned with the episodically non-backtracking case. We found that this was a necessary first assumption to make a rigorous proof about convergence to a stable trapline possible. From a biological perspective, the episodic-non-backtracking assumption means that the bee knows which flowers it has visited and which ones it has not since its last nest visit and actively avoids them. While it seems possible that bees can remember or sense in some way which flowers have recently been visited, it certainly is not true that bees actively avoid recently visited flowers to the point that they never visit the same flower twice within a single bout, as this commonly happens in bee tracking data. We address this limitation in the simulation study by examining convergence behaviour under different parameter values of the model, where we permit the bee to backtrack. Furthermore, we show in Theorem 1.23 that when we remove this assumption, the model may still converge to a stable trapline. However, a rigorous

proof that this is always the case eluded our efforts. Here, the sticky point of the argument is that edge weights along the heavy route, the route following the most likely positively reinforced transition, may shrink. Thus, it is not possible to establish a universal upper bound on the probability that the process may fail to follow the heavy route from some time n onwards. While we were unable to prove that the backtracking process must converge with probability one, we were unable to disprove it. And as such, we believe that this might still hold true.

From a biological perspective, convergence of the process is only relevant if convergence takes place quickly enough. In the weak reinforcement regime, convergence takes too long to have significant biological relevance, as flowers would not bloom long enough for the bee to complete enough bouts to converge to a stable trapline.

Another limitation in this work is that treated models do not distinguish between recent and distant experiences. Similarly to the episodic non-backtrack assumption, we found that this was necessary to make rigorous mathematical analysis tractable. Biologically, it seems reasonable to expect that a bee would weigh more recent experiences than long-past ones or even forget experiences from the distant past entirely. Future investigations of these models could include an explicit memory model that could include a set of experiences that the bee can base its movement decisions on. A simpler way to address the relevance gradient of recent versus long-past experiences could be addressed by introducing a weight decay towards null weights.

Finally, we would like to point out that choosing the exact order in which to visit potential food sources is only one way in which learning in bees could manifest. When considering radar tracks of bees, the straightness of flight paths also seems to decrease significantly over time [53]. In the discrete space models presented here, we are unable to model changes in flight paths. Future work could use the conceptual framework of reinforced random processes to develop continuous space

analogs, which may be able to explain the straightening of flight paths.

1.5.6 Significance

Taken together, these results clarify the mathematical conditions under which reinforcement models produce stable traplines and show how the two hypotheses, nearest-neighbour and order-of-encounter, arise as special cases of the same model. In order for stable traplines to emerge, there needs to be a strong incentive to retrace previously successful steps. Avoidance of unsuccessful actions alone cannot give rise to repetitive foraging behaviour under non-degenerate assumptions.

Chapter 2

Inference from Feeder Visitation Sequence Data

2.1 Introduction

Animals use memory to navigate environments more efficiently and complete foraging tasks. Yet most animal movement models make the simplifying assumption to disregard memory and learning when it comes to modelling animal movement and instead include environmental covariates such as habitat quality. In some cases of animal behaviour, the influence that memory and learned behaviour have on an animal's movement is undeniable. One such example is the phenomenon called "traplining behaviour". It refers to a kind of repetitive foraging behaviour where animals move between stationary resources in a near-deterministic fashion. This behaviour has been observed across a wide range of taxa from primates to bees.

Darwin is often credited with first describing this type of behaviour in bees. Despite the fact that traplining has been a subject of study in behavioural ecology for over 100 years, our understanding of how bees navigate and decide on which routes to take is still lacking. But recent research on cognition in insects is promising that we shall gain deeper insights into how bees navigate, choose food sources and solve

multi-destination routing problems. It has long been known that many insects follow hardwired behaviours. For instance, wood ants seem to be innately attracted to tree-like objects, or honey bees that leave their angle at an angle x to the sun will always return to the nest at the supplementary angle $180x$ [1, 4]. Furthermore, researchers have been able to formulate navigational algorithms that can be implemented in neural networks that resemble the structure of a bees central complex [27, 47]. Moël et (2019) show that the path integration algorithm developed by Stone et al (2017) can integrate multiple vector memories. In plain English, this means a bee may remember several locations and successfully navigate to it. This gives a rigorous scientific basis for the biological plausibility of trapline formation models such as the ones developed by Reynolds et al (2013) and Dubois et al (2021), which make use of the assumption that a bee may actively choose to move to one location over another. Le Moël et al. also mention that the exact mechanism for how the bee decides on which feeder to visit next is not yet understood [27], *a research gap we will address in this paper*. For example, we still do not understand if bees make routing decisions solely based on immediate rewards or if they also take into consideration factors such as the overall route quality [14]. We also do not know if bees only tend to revisit flowers when they find food or if they tend to return to flowers just to check for food availability [29]. Also, we ask if such route learning behaviour is hardwired in bees or does every bee have its own individual foraging strategy? These are the questions we hope to answer in what follows.

We believe part of the reason why these questions have not yet been answered is due to a lack of a precise conceptual framework for how to model learning and memory-informed movement in animals and methods to confront these models with data.

Previously, scientists have struggled to reliably define traplining behaviour using statistical methods [51]. Thomson et al. suggest three indices to detect and define

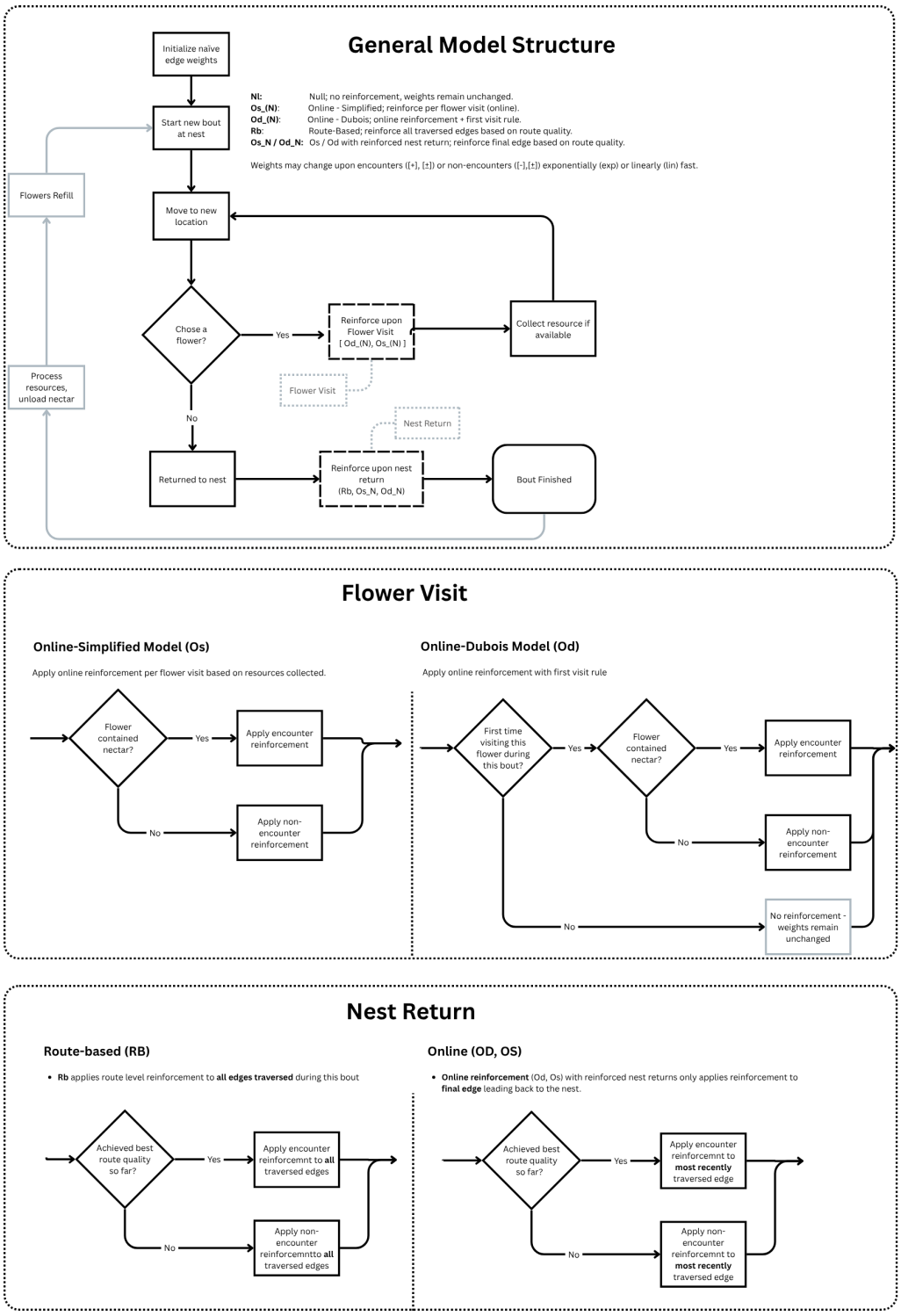

Figure 2.1: An overview of different models for trapline formation. Models are distinguished by which location the random walker is allowed to move to, when reinforcement is applied to the edge weights and criteria for applying positive or negative reinforcement. Edge weights can grow exponentially or linearly.

traplining behaviour: the asymmetry index, the variance of return lengths and the similarity index. The asymmetry index detects bias in the directionality of movement. The null assumption is that under no traplining, bees would have no preference for going from a to b or from b to a. A shortcoming of this index is that the directionality of movement could also be caused by external factors such as prevailing wind directions or nest position.

The variance-return-length-index is a normalized version of the variance in the number of steps taken to return to the same flower. It captures the idea that subsequent bouts should have low variance. To test for statistically significant traplining behaviour with respect to the variance return length index, it is necessary to produce null location visit sequences. If the variance in return length of observed data is within the lowest 5% of the return length variances produced by the null model, we consider this to be indicative of significant traplining.

Finally, Thomson et al. invented the internal alignment index is based on a symmetric similarity measure between two sequences called maximum alignment. It is based on the Smith-Waterman algorithm for aligning DNA sequences. This sequence alignment algorithm counts how many deletions or insertions are needed to render two sequences the same. The authors describe computing the similarity measure of two sequences of length m and n and lining them up as row and column headers of an $m \times n$ matrix. matrix entries are 1 if row and column headers agree and 0 otherwise. In a second step, insert dummy rows to maximize the number of 1 along the main diagonal. The similarity index is computed by dividing the total number of matches along the diagonal by the number of cells in the diagonal. They further develop methods to test for traplining based on these indices. For the asymmetry index, Thomson et al. (1997) suggest using the binomial probability test to test for directionality preference in each location-location pair. These hypothesis tests can then be combined using Fisher's method of combining probabilities. The

test derived from the similarity index is based on the Mantel test and can be used to test for significant differences in sequence similarity among individuals, as well as traplining when comparing

In [51] Thomson et al conclude that overall, no single one of their methods provides a complete solution to the challenge of creating a comprehensive operational definition of traplining. And suggest that as many metrics as possible should be used when the goal is to show an animal retraces its steps with statistical significance. Here, we add another such method. But more importantly, we take an additional step towards answering a more relevant and useful question that an investigator might ask: What sort of repetitive foraging patterns might one expect from bees using a particular tactic or facing a particular spatial problem and how, if at all, can we reliably detect these patterns and underlying mechanisms? Thomson emphasizes the importance of flexibility in the approach, given the data available and the question at hand. We believe that the approach presented here is flexible and precise enough to account for different experimental setups.

Later on, researchers like Ohashi, Lihoreau, Chittka, Woodgate and Reynolds adopt the hypothesis tests laid out by Thomson in 1997 [38, 28, 53, 44]. All of these studies were able to detect some form of statistically significant retracing of previous foraging routes. In 2012, Lihoreau et al. developed the trapline heuristic model, a model that marks an important first step towards understanding the underlying mechanics of traplining in bees. The idea is simple: whenever a bee moves from one location to another, the probability of repeating this action either increases or decreases, depending on whether or not the outcome of performing this action is desirable. Essentially, though not explicitly formulated as such, Lihoreau et al. frame traplines as an emergent solution to a reinforcement learning task. Now the question becomes which reinforcement learning algorithm could produce the behaviour observed in the real world?

In the seminal paper [44] Reynolds and Lihoreu propose a route-based learning algorithm that gives rise to traplines which seem to agree with observed flower visitation sequences. At its heart, this algorithm is best understood as an episodically edge-reinforced random walk in a complete directed graph. Let us unpack what this really means. The environment is discretized into locations of interest such as potential food sources, flowers, and the animal's central foraging place, the nest. For simplicity's sake, we are only interested in modelling the order of visits to each of these locations. At every location, the bee is faced with a "choice" of which location to visit next. We visualize this as arrows leading to all other locations, each arrow representing a certain action or choice. These arrows will also be called directed edges or transitions. The tendency of the bee to move from one location to another is given by a weight associated with each directed edge. These weights define transition probabilities and can change based on the bees' experience. In the very first proposed trapline heuristic model in [44] these weights change upon the bee returning to its nest. If the overall route quality was higher than the route qualities achieved during previous bouts, all weights associated with traversed directed edges increase and decrease otherwise. This simple learning can give rise to emergent stable traplines that seem to agree with patterns observed in the real world, when tested with methods such as the similarity index developed by Thomson [51, 44].

Most recently, Dubois et al. suggested a novel adaptation of this model, thought to be more parsimonious. Their critique of the route-based learning model given in [44] and supported in [31] is that in order to perform route-based learning, bees would have to remember entire routes until they return to the nest to determine the overall route quality. Instead, Dubois et al believe it is simpler if bees only relied on immediate rewards to update their tendency to retrace certain steps. In their model, the weight of the most recently traversed edge is updated every time the bee visits a location it has not yet visited during the current bout. Ironically, this first visit rule

again requires the bee to somehow know which flowers it has visited within a given bout. So part of this paper includes proposing a simplified online learning algorithm where the edge weight associated with the most recently traversed directed edge is always updated, regardless of whether the destination had been visited before or not.

Now that researchers have developed these sophisticated biologically relevant models for trapline formation, the question arises: which model best captures the true underlying mechanism for trapline formation? To answer this question, we fit these models to flower visitation sequence data with maximum likelihood estimation and perform model selection using the Bayesian information criterion (BIC). The benefits of using maximum likelihood are that we also obtain parameter estimates. These can paint a precise picture of trapline formation.

In Chapter 1, we show that the no-immediate-backtracking assumption can be avoided and the model structure made more parsimonious when we assume that after every step, positive or negative reinforcement is applied based on the outcome of the transitions without changing the qualitative behaviour of having infrequent immediate backtracks and self-avoiding stable convergent traplines. Thus, we suggest a third parsimonious simplified model based on the online reinforcement model proposed by Dubois et al. We call the model the simplified online reinforcement model and define this in detail in section 2.2.

In summary, we have three structurally distinct edge-reinforced random walk models, each representing a different hypothesis about bee memory and cognition. The route-based reinforcement model structure assumes relatively large memory capacity and is based on the hypothesis that bees aim to optimize foraging bouts as a whole. Online reinforcement, on the other hand, hypothesizes that bees optimize their bouts on a local scale based on immediate feedback of transitions. While the Online reinforcement model by Dubois et al. retains more of the global optimization

characteristic from the route-based reinforcement structure, the simplified online reinforcement structure is guided entirely by local immediate feedback and requires the least amount of memory capacity to execute out of all models that include some amount of memory.

2.2 Model formulation

We adopt the framework of an edge-reinforced random walk [11, 40] to model learning and memory reinforced movement. In general, we assume that the bees' movement is governed by a transition probability matrix that changes over time

$$\mathbf{P}^{(n)} = \left(p_{i,j}^{(n)}\right)_{i,j\in\Omega}$$

where $p_{i,j}^{(n)}$ denotes the probability of moving from location of interest i to location of interest j at time n. This transition probability matrix is derived from a dynamic transition weight matrix $\mathbf{W}^{(n)}$ with non-negative entries and at least one positive entry by rescaling every row to sum to 1,

$$p_{i,j}^{(n)} = \frac{W_{i,j}^{(n)}}{\sum_{k\in\Omega} W_{i,k}^{(n)}}. \tag{2.1}$$

Learning is modelled by the way in which transition weights change over time. Transition weights can change for multiple edges all at the same time in between bouts when the bee has returned to its nest (route-based learning) or one edge at a time after every transition (online learning). Weights can change differently depending on what the bee experienced. We distinguish between encounter, i.e. finding food or returning to the nest after a successful foraging bout, and non-encounter experiences, not finding food or returning to the nest empty-handed. We expect that, as a reaction to encounter experiences, weights will grow, increasing

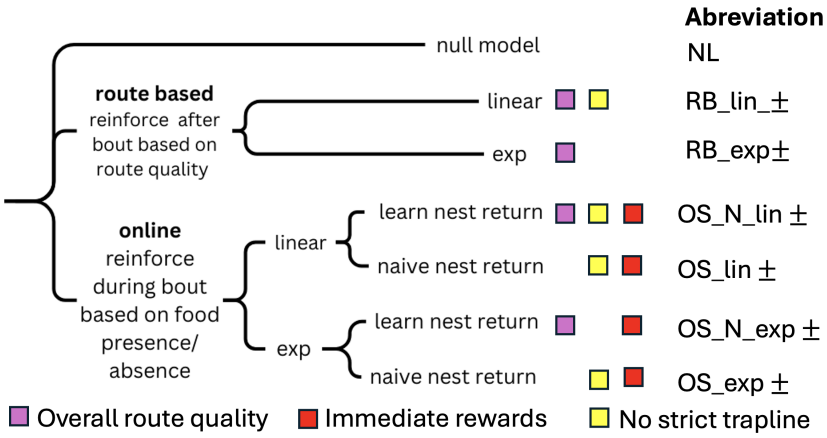

Figure 2.2: A dendogramm of the seven models fitted to data and the hypotheses each of them encodes. models with a purple tag take into consideration the over all route quality in reinforcement decisions. models with a red tag take into consideration immediate rewards encountered at flowers. Models with a yellow tag predict that in the long run a bee will not converge to a strict trapline in the sense of following the same route at all large times. In the righ most column the Model's name or abbreviation for that matter is given.

the probability that the action that brought about the encounter outcome will be repeated. On the other hand, when the experience was non-encounter, we expect the weight associated with this non-encounter experience will shrink. In Chapter 1, we showed that the exact functional form in which weights change is essential to the model's predicted long-term behaviour. In a nutshell, when weights are allowed to increase exponentially with the number of encounter experiences associated with an action and this action reliably leads to encounter outcomes, the action will fix over time, meaning that the action is carried out at all large times. This leads to what we might call *strong traplining behaviour*, where a bee would follow the same foraging route from some time onward. On the other hand, we have a much weaker form of traplining behaviour that emerges when weights increase and decrease linearly with the number of encounter and non-encounter experiences, respectively. Figure 2.2 gives an overview of all models fitted to data and their hypotheses they encode as well as the names for these models used in the following.

Before precisely formulating the models, we establish some notation. The set of

locations is Ω. Denote the location of the bee at time n by $v_n \in \Omega$. A *bout* starts and ends at the nest. Let $b(n) = \sum_{0 \leq m \leq n} \mathbf{1}(v_m = v^*)$ count the number of times the bee has been at the nest up to time n , and let $s(n) = \sum_{0 < m < n} \mathbf{1}(b(m) = b(n))$ be the amount of time that has passed since the last nest visit. We relabel the visitation sequence $u_{b(n),s(n)} = v_n$ to be the location during the $b(n)^{\text{th}}$ bout at the time step $s(n)$ and the b^{th} bout. More precisely, the b^{th} bout is defined as the sequence of visits starting and ending between consecutive nest visits, with the first nest visit inclusive and the last nest visit belonging to the next bout $B_b = (u_{b(n)=b,s(n)})$. Similarly, we can relabel the Weight matrices $\mathbf{U}^{(b(n),s(n))} = \mathbf{W}^{(n)}$.

2.2.1 The null model

The null model predicts that bees do not learn, i.e. adjust weights based on experiences; they move from flower to flower in a naive fashion. Since even for completely random movement, closer flowers are more likely to be the first flowers to be hit by chance, we follow [44] (2013) and define the transition weight matrix $\mathbf{W}^{(0)} = (W^{(0)}_{i,j})_{i,j \in \Omega}$ with

$$W^{(0)}_{i,j} := \frac{\text{dist}^{-d}(i,j)}{\sum_{k \in \Omega} \text{dist}^{-d}(i,k)}. \tag{2.2}$$

So the initial weight associated with the action of going from location i to j is inversely proportional to the distance taken to the power d. We call d the *distance exponent*. A large distance exponent implies a large preference for closer flowers over farther ones. Further, we notice that $\mathbf{W}^{(0)}$ is a stochastic matrix where all rows sum to unity. Since the null models assume that no learning takes place, we have that $\mathbf{W}^{(n)} = \mathbf{W}^{(0)}$, and movement is essentially a Markov chain with the transition matrix $\mathbf{P}^{(0)}$.

2.2.2 Learning opportunities - route-based and online

The model developed by [44] encapsulates the hypothesis that bees make decisions based on overall route quality and adapt their behaviour from bout to bout. Transition weights change upon nest return. Let $T_b = \{(u_{b(n),s(n)}, u_{b(n),s(n)+1}) \in \Omega \times \Omega : b(n) = b\}$ be the set of edges traversed during the b-th bout. Then

$$W_{i,j}^{(n+1)} = \begin{cases} r(W_{i,j}^{(n)}; \xi_n) & \text{if } b(n+1) = b(n) + 1 \text{ and } (i,j) \in T_{b(n)} \\ W_{i,j}^{(n)} & \text{otherwise.} \end{cases} \tag{2.3}$$

Here $r : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is the reinforcement function responsible for modelling learning based on encounter, non-encounter or neutral experiences. This type of reinforcement assumes that the bee is able to remember all transitions taken during the bout so that the experience of returning to the nest can be associated to each of these transitions.

In *online reinforcement*, only the weight of the most recent transition changes at every time step, modelling immediate adaptations in behaviour based on the immediate experience. Explicitly, the Weight transition matrix at each time step is defined as

$$W_{i,j}^{(n+1)} = \begin{cases} r(W_{i,j}^{(n)}; \xi_n) & \text{if } (i,j) = (v_{n-1}, v_n) \\ W_{i,j}^{(n)} & \text{otherwise} \end{cases} \tag{2.4}$$

where r is the reinforcement function. The experience ξ_n is directly associated with the most recent traversal. Note that when the experience associated with a transition is neutral, weights will not change $r(w, \xi = 0) = w$.

Encounter, non-encounter and neutral experiences

In any model, the bee distinguishes in two different kinds of transitions, ones leading to flowers (flower-transition) and those leading back to the nest (nest-transitions). The bee might respond to the experience ξ of using these experiences may be classified in three different ways, which we will define and explain here: *encounters*, non-*encounters* and *neutral* experiences. Flower transitions are classified as encounters $\xi^{(\text{flower})} = 1$ when the bee finds food at the destination, and a non-encounter $\xi^{(\text{flower})} = -1$ when no food is found at the flower

$$\xi_n^{(\text{flower})} = \begin{cases} 1 & \text{if food is found at flower } v_n \neq v^* \\ 0 & \text{if returned to the nest } v_n = v^* \\ -1 & \text{no food is found at flower } v_n \neq v^* \end{cases} \tag{2.5}$$

In the online reinforcement model presented by Dubois et al. [14] the authors additionally include a first visit rule for flower-transitions, where a flower-transition is only a learning experience when the flower is visited for the first time during a bout

$$\nu_n = \mathbf{1}(u_{b(n),s(n)} \notin \{u_{b(n),s(m)}\}_{m<n}). \tag{2.6}$$

When a bee returns to the nest, the experience is classified based on route quality. If the bee has just completed a route of the highest quality experienced thus far, we classify it as an encounter. If the bee has experienced better routes before, this is a non-encounter. We formalize this by defining the nest return experience at time n

$$\xi_n^{(\text{nest})} = \begin{cases} 1 & \text{if } v_n = v^* \text{ and } q_{b(n)} = \max_{m \le n} q_{b(n)} \\ 0 & \text{if } v_n \neq v^* \\ -1 & \text{if } v_n = v^* \text{ and } q_{b(n)} < \max_{m \le n} q_{b(n)}. \end{cases} \tag{2.7}$$

Here $q_{b(n)}$ is the route quality of the $b(n)$-th bout and is calculated as

$$q_b = \frac{(\sum_{s \in \mathbb{N}} f_{b,s})^2}{\sum_{s \in \mathbb{N}} \text{dist}(u_{b,s-1}, u_{b,s})}. \tag{2.8}$$

the route quality is the ratio of the squared total amount of food gathered and the distance travelled during the b-th bout. We say an experience is neutral or does not apply to the given transition and write $\xi = 0$.

In summary, for different models, the experience ξ_n is computed as

$$\xi_n = \begin{cases} \nu_n \xi_n^{(\text{flower})} & \text{for model OD} \\ \nu_n \xi_n^{(\text{flower})} + \xi_n^{(\text{nest})} & \text{for modelOD\_N} \\ \xi_n^{(\text{nest})} & \text{for model RB} \\ \xi_n^{(\text{flower})} & \text{for model OS} \\ \xi_n^{(\text{flower})} + \xi_n^{(\text{nest})} & \text{for model OS\_N} \end{cases} \tag{2.9}$$

The set of edges the experience is associated with varies based on the model structure. In *online learning* , the experience ξ_n is associated with the n-th transition only. In *route-based learning* the final transition of the bout leading back to the nest is associated with the entire route.

2.2.3 Linear and exponential reinforcement

The reinforcement function $r(w;\xi)$ indicates how a single edge weight w changes depending on the experience ξ associated with it. For *linear learning* or *reinforcement*

$$r_{\text{lin}}(w;\xi) = \begin{cases} \max\{w+\ell, 0\} & \text{if } \xi > 0, \\ w & \text{if } \xi = 0, \\ \max\{w+a, 0\} & \text{if } \xi < 0. \end{cases} \tag{2.10}$$

In this formulation ℓ , and a are called the *encounter* and *non-encounter increment*, respectively. For *exponential learning* or *reinforcement*

$$r_{\text{exp}}(w;\xi) = \begin{cases} w \cdot (1+\ell) & \text{if } \xi > 0, \\ w & \text{if } \xi = 0, \\ w \cdot (1+a) & \text{if } \xi < 0. \end{cases} \tag{2.11}$$

Here $(1+\ell)$ and $(1+a)$ are called the encounter and non-encounter factors, respectively. In summary, weights can grow in two qualitatively distinct ways,

$$r(w;\xi) = \begin{cases} r_{\text{lin}}(w;\xi) & \text{if model linear learning,} \\ r_{\text{exp}}(w;\xi) & \text{if model exponential learning.} \end{cases} \tag{2.12}$$

When a bee does not learn from encounter or non-encounter experiences, then $\ell = 0$ or $a = 0$ respectively. A negative encounter or non-encounter increment implies that the bee is less likely to repeat the same action given an encounter or non-encounter experience. We expected that the encounter increment is positive and the non-encounter increment is negative, but found that both are positive.

2.2.4 Exponential learning as resource selection functions

Alternatively, the exponential reinforcement rule could also be formulated as a resource selection function combined with a naive movement kernel, following the conceptual framework of [45]. In this formulation, the number of encounter and non-encounter experiences associated with an edge becomes the resource weight. More explicitly, we formulate the exponentially growing weights model in terms of a resource selection function that is multiplied by a naive movement kernel. Let $S_n^+(i,j)$ and $S_n^-(i,j)$ count the number of times the edge (i,j) was associated with an encounter and non-encounter experience by time n , respectively. Then the exponential model is

$$P_{i,j}^{(n)} \propto P_{i,j}^{(0)} \cdot (1+\ell)^{S^+(i,j)} \cdot (1+a)^{S^-(i,j)} \tag{2.13}$$

$$= P_{i,j}^{(0)} \cdot \exp\left(\log(1+\ell)\, S^+(i,j) + \log(1+a) S^-(i,j)\right) \tag{2.14}$$

$$\propto P_{i,j}^{(0)} \cdot \frac{\exp\left(\log(1+\ell)\, S^+(i,j) + \log(1+a) S^-(i,j)\right)}{\sum_{k\in\Omega} \exp\left(\log(1+\ell)\, S^+(i,k) + \log(1+a) S^-(i,k)\right)}. \tag{2.15}$$

Here, the final line is the kernel of the probability mass function that dictates the bees' movement. The final term is a resource selection function. The action or resource-preference-score is proportional to the effective number of encounter and non-encounter experiences associated with the edge.

We chose linear and exponential growth paradigms for edge weights to model both stable trapline behaviour, using exponential growth in edge weights, and behaviour where the bee learns but continues to explore after some time. Recall that the model may exhibit traplining behaviour if and only if inverse edge weights are summable. Therefore, we selected exponential and linearly growing edge weights as representative cases for either of these model classes: models that support stable traplines and those that do not. We chose the exponential model as a representative

of the class of models that support stable traplines due to its previous use in literature [14] and its alternative formulation as a mixture of a naive movement kernel and a resource selection function. We chose linearly growing weights to represent the class of models that will never fully converge to a stable trapline since it seemed parsimonious to us and resembles the well-studied and widely appreciated Polya urn scheme[22].

2.2.5 Model fitting

In order to fit these models above to data, we require complete sequences of feeder visits $\mathbf{x} = (x_0, \ldots, x_N)$, with $x_n \in \Omega$ the distances between all locations $D = \text{dist}(v_i, v_j)$ for all $v_i, v_j \in \Omega$ and the information on whether or not food was found at the location or not $\mathbf{f} = (f_0, \ldots, f_N)$ and finally, we also need to be able to identify the nest among the locations. Based on this information, we can compute the likelihood of the data under given parameter values $\theta = (a, \ell, d)$ for the non-encounter increment a, encounter increment ℓ and distance exponent d precisely as

$$p(\mathbf{x}|\theta, \mathbf{f}, D) = \prod_{n=1}^{N} p(x_{n+1} \mid \theta, (x_m)_{0=m}^{n}, (f_m)_{m=0}^{n}, D). \tag{2.16}$$

Using Nelder-Mead optimization, we can use maximum likelihood estimation and parametric bootstrapping to compute parameter estimates and confidence intervals. To verify the validity of this method, we run a simulation study. In this simulation study, we further investigate how well BIC performs as a model selection criterion.

2.3 Flower visitation sequence data

The flower visitation sequence data used for this analysis was collected by Woodgate et al. and published in visual form in [53]. These data were collected using

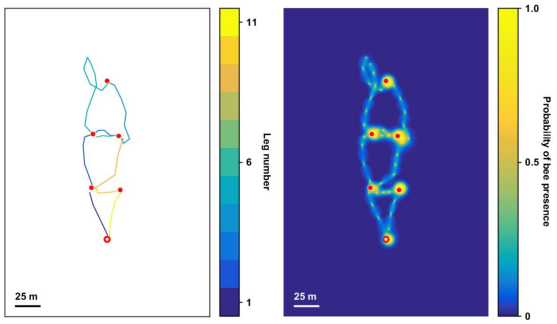

Figure 2.3: An example track of a bee foraging on the feeder array from Woodgate et al, Supplementary material S1 Figure S47 [53].

harmonic radar tracking. Bumble bees were *Bombus terrestris audax*, were pre-trained to forage feeders, which consisted of a blue-painted wooden platform (20cm × 20cm)mounted on top of a white pole (92cm). At the center, there was an acrylic plate with a hole to hold a precisely measured amount of 40% sucrose solution. Training took place in two stages, in a flight room and outdoors. Only a small percentage of bees continued to visit feeders at each stage. Bees were tracked continuously over the course of 1 to 2 days. Figure 2.3 shows the final bout of Bee #1 from [53]. During the experiment, the feeders and the nest were arranged in an elongated hexagon 150m, long and 25m wide, with the nest at one of the far ends. The amount of nectar in the 5 feeders was calibrated for each bee individually so that it would constitute one-fifth of its crop capacity.

2.4 Simulation study

To verify accurate model selection with the amount of data available to us from [53]. We selected models based on the Bayesian Information Criterion (BIC) as BIC is better for selecting the true underlying model, whereas AIC is more geared towards selecting the most predictive model [5].

We simulated 150 transitions from each of the 31 models 100 times. To all $31 \times 150 \times 100 = 465000$ resulting in-silico data sets, we fit all 31 models using

maximum likelihood estimation and selected the best model based on BIC. We simulated the data with encounter increments of $a = -0.25, \ell = 0.1$ for exponential reinforcement and $a = -0.05, \ell = 0.1$ for linear reinforcement, and the distance exponent was simulated at a value of $d = 2$. Based on previous research, these were values deemed to be realistic. In our simulations, we assumed that flowers refill in between bouts as per the experimental design of Woodgate et al [53].

When the true generating model is of the OD-type, this method tends to select the null model instead of the true generating model most of the time. Furthermore, distinguishing between models that allow for only a single reinforcement, where only encounter or only non-encounter experiences lead weights to change, is inaccurate.

We pre-select a subset of models that will allow us to answer biological research questions. To achieve this, we eliminate the online-dubois type models and models that restrict learning opportunities to only encounter or non-encounter experiences. Thus, we reduce the number of candidate models to 7 (NL, OS_exp[±], OS_lin[±], OS_N_exp[±, OS_N_lin[±], RB_lin[±], RB_exp[±]]). In this subset of models, the true generating model is selected at least 70% of the time. We deemed this sufficient to conduct a statistical analysis. Parameter estimates are centred around the true parameters and 95% parametric bootstrap confidence intervals using a bootstrap sample size of 200 model fits cover the true parameter values approximately 95% of the times, as expected.

2.5 Model fits to bee data

We fitted a total of seven distinct models to the flower visitation sequence data collected by Woodgate et al. (2017). These seven models are the null model (NL), the route-based model with linearly changing weights with encounter and non-encounter experiences (RB_lin±), route-based model with exponentially changing

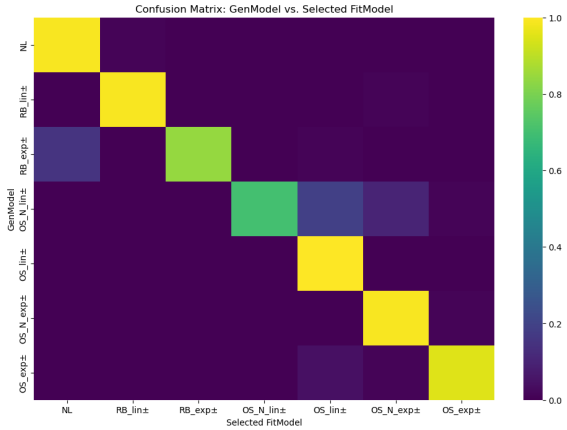

Figure 2.4: Empirical Confusion matrix of selecting the true generating Model based on 200 simulations for each data-generating model. The true generating model is on the y-axis and the selected model is on the x-axis.

weights with encounter and non-encounter experiences (RB_exp±), the simplified online learning model with linearly changing weights with encounter and non-encounter experiences (OS_lin±), the simplified online learning model with expoentially changing weights with encounter and non-encounter experiences (OS_exp±), the simplified online learning model with linearly changing weights with encounter and non-encounter experiences and reinforced nest returns (OS_N_lin±) and the simplified online learning model with exponentially changing weights with encounter and non-encounter experiences and reinforced nest returns (OS_N_exp±).

2.5.1 Missing data

Since weights are products of the entire movement history on the array, we only considered complete flower visitation sequences for each bee and discard data after a flower visit has been missed. Some sequences did not include the nest as the final flower visit, in which case we first consulted the radar tracks. For Bee B74 (Bee 3), Bout 12 (Supplementary Figure S110) in [53] , we could see that the radar track ended right at the nest, so that we assumed Bout 12 finished at the nest. On the other hand, B61 (Bee 2 in the paper), Bout 33 had Feeder 2 as the last entry in the

visit sequence. Looking at the flight path (Supplementary Figure S80), the radar track ended abruptly with the bee in between Feeder 2 and Feeder 1. We could be certain that the bee did return to the nest, but we could not guarantee that it did not stop at any other feeders along the way, so we classified this as an incomplete sequence. B61, Bout 41 and B56 (Bee 1), Bout 22 were likewise incomplete. After discarding data from later bouts, we had at least 170 transitions for each of the 4 bees. Bee Y01 only had a total of 16 recorded transitions, a number too small to permit reliable statistical analysis. Hence, we excluded Bee Y01 from further analysis. In total, we analyzed 1250 transitions, which made up approximately 60% of the total data collected.

2.5.2 Results

For 3 out of the 4 bees, the Route-Based linear learning (RB_lin[±]) is the best model fit, and the online simplified model was the second best fit to the data, with ΔBIC between 2 and 5, see Figure 2.5. The non-encounter increment in the best fit model was estimated to be $0.118, 0.660, 0.179, 0.204$ with 95% confidence interval, lower bound $0.028, 0.178, 0.068, 0.042$, and upper bound of $0.244, 1.49, 0.331, 0.472$ for Bee B56, B61, B74, O72 respectively. The encounter increment in the best fit model was estimated to be $0.116, 0.444, 0.383, 0.391$ with a 95% confidence interval, lower bound $0.014, 0.157, 0.177, 0.118$, and upper bound of $0.282, 1.08, 0.708, 0.868$ for Bee B56, B61, B74, O72 respectively. The distance exponent in the best fit model was estimated to be $1.539, 1.326, 1.574, 1.332$ with a 95% confidence interval, lower bound $0.941, 0.535, 0.821, 0.683$, and upper bound of $2.068, 2.184, 2.489, 2.216$ for Bee B56, B61, B74, O72 respectively.

For bee B56, the second most supported model was the online-simplified-linear reinforcement model with reinforced nest returns (OS_N_lin±). The ΔBIC for this model is 2.66. All other models have a ΔBIC greater than 19. The non-

encounter and encounter increment for the OS_lin_N± model are 0.096 and 0.130 with confidence intervals of $(0.030, 0.190)$ and $(0.033, 0.257)$ respectively.

Similarly, for bee B61 OS_N_lin± is the second best fit model with a ΔBIC of 3.285 all other models had a ΔBIC greater than 86. The model OS_N_lin± estimates the non-encounter and encounter increment are as 0.513 and 0.500 with confidence intervals ranging from 0.158 and 0.129 to 1.133 and 1.060.

Also, for bee O72 OS_N_lin± is the second best fit model with a ΔBIC of 2.372. Here, the non-encounter and encounter increment are estimated to be 0.205 and 0.257 with confidence intervals ranging from 0.048 and 0.069 to 0.381 and 0.557 respectively. The third most supported model is RB_exp± with a ΔBIC of 3.070. For bee O72 RB_exp± estimates the non-encounter and encounter increment as 0.167 and 0.519 with confidence intervals ranging from -0.001 and 0.167 to 0.294 and 1.028. Additionally, for this bee, the fourth moderately supported model is OS_N_exp±. With a ΔBIC of 9.107 we estimate encounter and non-encounter increments as 0.190 and 0.201 with confidence intervals ranging from 0.089 and -0.066 up to 0.331 and 0.403

Finally, for bee O72, the OS_N_exp± is by far the best supported model, with all other models having a ΔBIC greater than 10.

2.6 Discussion

Each model that we fitted to the data represents a distinguishable biological hypothesis. To evaluate the evidence based on ΔBIC we distinguish in four evidence classes, weak, positive or moderate, strong and very strong, with ΔBIC ranging from 0 to 2, $2-6$, $6-10$ and > 10 respectively [43]. Thus, if a modek is assigned ΔBIC of three, we would consider this moderate evidence in favour of the best model. In general a lower ΔBIC means that the model is more likely to be the true generating model.

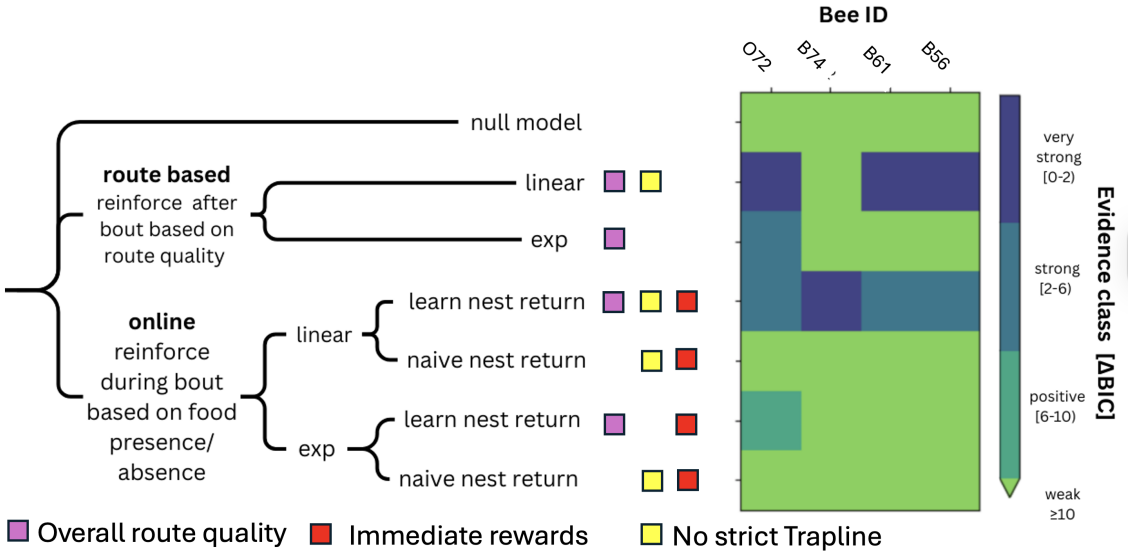

Figure 2.5: ΔBIC for individual bees and Model fits. Low ΔBIC scores are a better fit. Three of the four bees are best modelled by routebased linear reinforcement.

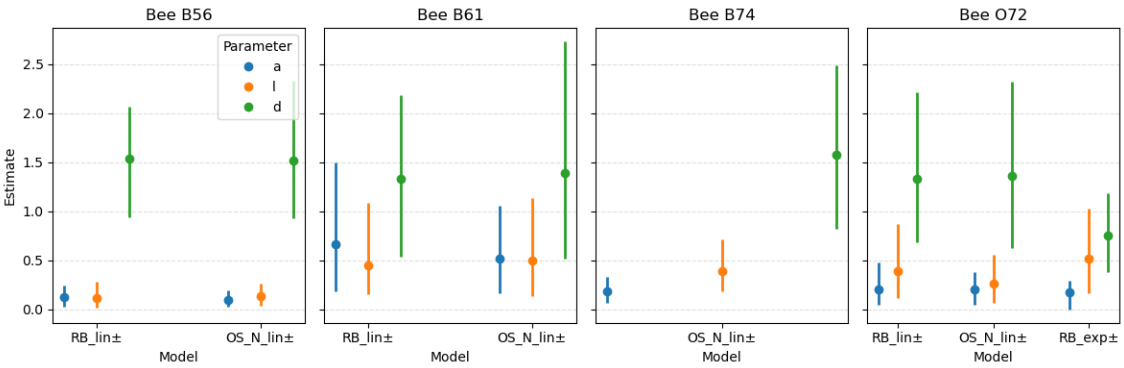

Figure 2.6: Parameter estimates for each bee per panel. The models on the x-axis are sorted by best fit of ΔBIC left to right, lowest to highest. Models with a ΔBIC greater than 6 are excluded from this plot. a and ℓ are non-encounter and encounter increments, d is the distance exponent.

2.6.1 Do bees learn multi-destination routes when foraging on an array of flowers?

We consider the results of this study as very strong evidence that bees exhibit learning behaviour in complex multi-destination routing and foraging tasks by adjusting the order in which they visit different feeder sites. The null model assumes that transition probabilities of moving from flower to flower remain constant throughout the entire foraging process. For all bees, the null-model predicting no-learning had the highest ΔBIC, far exceeding the ΔBIC = 10 threshold to be considered very strong evidence for learning. Recall that a higher ΔBIC implies a worse model fit.

Furthermore, in the hypothesis testing framework, we find that confidence intervals for the learning parameters do not include 0 in the best models for all bees. Often with a considerable effect size, with non-encounter and encounter increment, the maximum likelihood estimates exceed 0.2.

2.6.2 Do bees converge to stable traplines?

Based on the analysis conducted here, we find that overall linearly growing weights seem to best model bee foraging behaviour. Technically, when weights grow linearly, a random process does not converge to following the same transitions at all large times as would be the case when a stable trapline is established as we have seen in Chapter 1. The models that would predict development of a stable trapline, i.e. models that assume weights grow/ decay exponentially with the number of encounter/non-encounter experiences associated with an edge as shown in Chapter 1, consistently performed worse than models predicting no formation of stable traplines under stable environmental conditions. Having said that, linearly growing weights can still lead to significant retracing of routes that could be considered traplining, but to call it stable is likely an exaggeration.

2.6.3 Do bees make global or local decisions

For three out of the 4 bees analyzed the route-based learning model performed moderately better than the online learning model. Interestingly, adding reinforced nest returns to the model always improved the ΔBIC score of the model significantly. Thus indicating that the bees tend to repeat routes that have improved route quality.

2.6.4 Are foraging strategies hardwired in bees?

Since one bee displayed a different foraging strategy from the others, we believe that individual bees may pursue different route formation strategies. However, to come to a consistent conclusion, we need to observe more bees. It is also worth mentioning that the OS_N_lin$\pm$ model is the first or second-best model for all bees, being only moderately worse at explaining the data than the route-based learning model with linear reinforcement RB_lin$\pm$. Thus, it seems plausible that maybe there is a true hardwired process of how bees find efficient foraging routes that uses aspects of RB_lin$\pm$ and OS_N_lin$\pm$.

2.6.5 Sample and shift strategies vs. complete traplining

We expected that the non-encounter increment would be negative as a result of a sample and shift strategy where bees would use previous reward experiences as information, and respond to them in a 'win-stay, lose-shift' manner. Instead, we found that in none of the best models ΔBIC < 6, the non-encounter increment is estimated to be negative. In fact, it is estimated to be strictly positive, with confidence intervals not including 0. What this means is that even when no nectar was found at a location, the bumblebee tended to return to this same feeder location. In fact, the tendency to return to the same feeder location after no food is found is not significantly different from the tendency to return to a location after food

was found. Therefore, we believe that all foragers tested in this study follow a complete traplining approach rather than a sample and shift strategy. Ohashi and Thomson have investigated differences in foraging efficiencies of these two strategies based on computer simulations. They found that the complete traplining strategy tends to be more robust under heavy competition and always performs at least as well as the sample-shift strategy [37]. A heuristic explanation for the benefits of traplining in general is that searching for new food resources is costly and does not necessarily lead to higher rewards when the amount of nectar located at each flower is bounded above. Ohashi et al. explain this as follows: when resources are allowed to grow indefinitely, searchers and trapliners perform equally well. Trapliners have a steady flow of resources in both scenarios, whether resources are bounded or unbounded. Searchers collect very little nectar from most flowers. When flowers linearly accumulate nectar over time, searchers get lucky every once in a while and "win the jackpot" collecting a lot of nectar all at once. When the resource amount at each flower is limited, these jackpots can no longer reach sizes large enough to compensate for the generally bad harvest. A similar logic might apply for sample-shift trapliners and complete trapliners, since a sample-shift trapliner invests more time and energy searching, a complete trapliner generally has steady benefits. These findings are further backed up by theoretical results stating that systematic foraging, such as traplining, can increase the amount of resources collected compared to an independent search strategy[41].

Our findings support the idea that bees follow foraging strategies that minimize the variance of encountered resources by consistently returning to the feeder location regardless of whether or not these locations contained food or not, at least on the time scale of our data. Having said this, empirical studies of trapline formation in the presence of competition, i.e. other foragers, show that bumblebees reduce overlap of the flowers they forage on over time [36]. This behaviour could be interpreted as

evidence in favour of a sample shift strategy, but does not necessarily imply that the shift is a result of not encountering food. We conjecture based on the discrepancy of evidence for bumblebees in vs. out of competition, and that shifting may be prompted not by the absence of food, but by some mechanism that tells the second bee that another bee had depleted this resource beforehand. One such mechanism could be electroreception. It is well known that bees and other pollinators carry positive electric charges, whereas flowers are negatively electrically charged [8, 15]. Furthermore, bumble bees and honeybees are not only capable of perceiving the presence or absence of weak electromagnetic fields but can also distinguish between different shapes resulting from different charge distributions on flowers. When a bee visits a flower, it reduces the charge of the flower, weakening the magnetic field. Thus, it seems plausible that bees can detect whether or not a flower has been visited based on the surrounding electromagnetic field.

2.6.6 Limitations

A major limitation of our analysis is that the exactly true model is likely not among the models tested for in our analysis. This means we only select the best of our candidate models, neither of which is likely to be an accurate depiction of an arbitrarily complex reality. However, injecting more realism into the models in question does not always increase accuracy in predictions. For example, it may be that the model could be more realistic if we included a precise decision mechanism based on energetics. However, given the similar spatial scale of distances between feeder locations and the fact that we would need to rely on general estimates for a bee's vital rates, we doubt that, in the end, establishing a direct link between energetics and the bees' movement would yield more precise results. This is especially true when considering that bees in this experiment had to carry harmonic radar transceivers, which pose a significant weight addition to a small bee. While the idea

of testing many models with high degrees of realism may be intriguing, we caution from testing too many models that in the end cannot be properly distinguished using the data available.

Additionally, we want to remind the reader that part of the goal was to demonstrate that it is possible to fit these models to flower visitation sequence data and obtain precise parameter estimates. Thus, we hope that future work will take advantage of the techniques developed here and build on our results. In particular, we hope that sequences of at least 170 consecutive transitions will be collected from more bees on differently shaped feeder arrays. Being able to compare tracks from a large number of bees would allow us to make progress in answering whether or not the navigational decision mechanism is hard-wired.

Furthermore, the method developed here only uses the order of feeder visits. However, the order in which feeders are visited is only one way in which bees optimize their foraging behaviour. Woodgate et al. (2017) point out that by increasing the straightness of flight paths, bees are able to increase foraging efficiency significantly. Unfortunately, our methods are blind to such improvements, but could potentially be adapted to include such information.

We also do not expect that these models will generalize well to scenarios where bees need to forage on a large array of flowers, as memory and energetic limitations will likely play a more important role than they are given in our models. In our models, we assume that at every location, the bee can access information about the accumulated experiences it had at a given location, the number of encounters and non-encounters it had when deciding to visit a certain feeder while currently being located at another. For a large number of flowers, this becomes a large number of experiences the bee needs to remember. Future models could use limited memory

Finally, we would like to point out that our models make the strong assumption that a bee decides where to go next based on the accumulated experiences and does

not distinguish between more or less relevant experiences. In future models, bees might weigh recent experiences more heavily than ones from the distant past.

2.6.7 Conclusion

We used maximum likelihood estimation to fit variants of the traplining heuristic model to flower visitation sequence data collected by Woodgate et al. [53]. We found that linear route-based reinforcement best models the formation of weak traplining behaviour observed. Unexpectedly, bees tended to retrace steps even when they did not result in the bee increasing its foraging efficiency or finding more food. This foraging behaviour suggests that bees identified the artificial flowers as potential feeding locations that they checked for food availability. For a bee, the mere presence of a flower might be sufficient to warrant returning and checking for food, as this location has already been identified as a potential food source. More generally, this work demonstrates the possibilities that fitting precise mechanistic models integrating data on food availability and movement opens up for understanding animal cognition and resource selection.

Conclusion

What is the nature of traplining behaviour in bees? Traplining is a taxonomically widespread behaviour, where animals foraging on patchily distributed, stationary, renewing food sources will repetitively visit feeding sites. Even though traplining already intrigued early naturalists like Darwin, who observed traplining in bees, the ontogeny and the underlying mechanisms for how traplines emerge remain poorly understood to this day, over 100 years later. In 2013, Reynolds et al proposed an iterative improvement heuristic to model trapline formation in bees. In their model, bees move from flower to flower with probabilities given by a probability matrix. Transition probabilities change after every bout. When the bout was successful, the probability of repeating the successful route increases and decreases otherwise. Later on, this model was adapted by Dubois et al. to base reinforcement on local rewards such as finding food at the visited flower. Simulations suggest that these models do converge to stable traplines in a similar fashion to how real bees do. In Chapter 1, we give rigorous proof that versions of this model must converge to stable traplines, revisiting the same sequence of feeders at all large times. We further argue that the precise parameter configuration also influences the shape of the emergent traplines. Weaker and slower reinforcement is likely to give rise to nearest neighbour routes that emerge very slowly, whereas strong reinforcement is more likely to give rise to traplines that emerge very fast and are similar to the first route that was taken. Furthermore, we establish a necessary and sufficient condition for when an

episodically reinforced random walk can exhibit stable traplining behaviour; edge weights must be allowed to grow superlinearly in order for traplines to emerge with positive probability. Future research should involve finding more general conditions under which traplines are guaranteed to emerge.

In chapter 2 we use maximum likelihood estimation to fit variations of the trapline heuristic model to flower visitation sequence data from [53]. We select best best-fit models based on the Bayesian information criterion (BIC). The validity of these methods is verified in a simulation study. As a result of fitting these models directly, we are able to provide evidence that bees do not converge to perfectly stable traplines as models with linearly growing weights performed much better than models with exponentially growing weights. Furthermore, bees seem to generally take into consideration the overall route quality. Thus, it seems plausible that bees may be able to remember entire foraging routes to base their decisions on. All four tested bees showed a significant change in behaviour towards retracing previous steps it came as a surprise that bees did not seem to distinguish between retracing highly efficient routes over routes that were less efficient.

A deeper theoretical understanding of trapline heuristic models and their convergence behaviour prompted us to investigate alternative formulations of these models. Confronting these models with data, we are able to test assumptions baked into variants of distinct trapline heuristic models. This work demonstrates the importance to developing a deep understanding of biological models as it reveals implicit assumptions. In [44], the simple assumption of a multiplicative reinforcement mechanism was made. This is equivalent to making the assumption that bees do converge to perfect traplining behaviour, following the same route at all large times. Understanding why exponentially growing weights might be equivalent to assuming the emergence of perfect traplines allowed us to reformulate the model with linearly growing weights, encapsulating the hypothesis that bees do not follow one and the same trapline

at all large times. Further realizing that it is possible to use maximum likelihood estimation and a multiple working hypothesis approach, we were able to provide empirical evidence that bees might not follow perfect traplines, but still have a tendency to retrace their steps.

Appendix A

Appendix

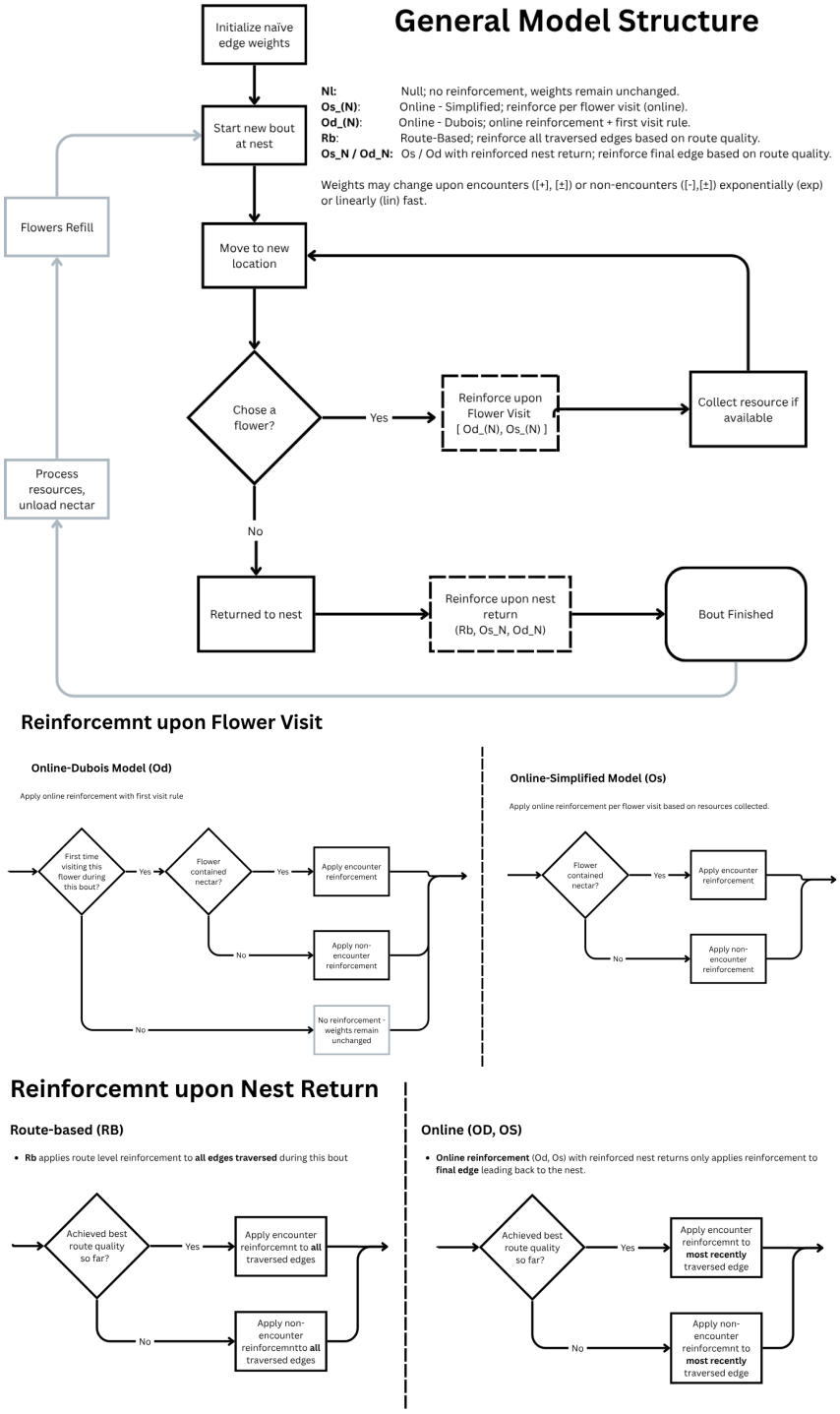

Figure A.1: An overview of different trapline model structures, route-based and online reinforcement with and without reinforced Nest returns.

Model abbreviation	Learning Opportunities	edges to be reinforced	weight growth
NL	-	-	-
RB_exp[±]	positive and negative nest returns	all traversed during bout	exponential
RB_exp[+]	positive nest returns	all traversed during bout	exponential
RB_exp[−]	negative nest returns	all traversed during bout	exponential
RB_lin[±]	positive and negative nest returns	all traversed during bout	linear
RB_lin[+]	positive nest returns	all traversed during bout	linear
RB_lin[−]	negative nest returns	all traversed during bout	linear
OS_exp[±]	positive and negative flower transitions and nest returns	last transition	exponential
OS_exp[+]	positive flower transitions and nest returns	last transition	exponential
OS_exp[−]	negative flower transitions and nest returns	last transition	exponential
OS_lin[±]	positive and negative flower transitions and nest returns	last transition	linear
OS_lin[+]	positive flower transitions and nest returns	last transition	linear
OS_lin[−]	negative flower transitions and nest returns	last transition	linear
OD_exp[±]	positive and negative flower transitions to new flowers and nest returns	last transition	exponential
OD_exp[+]	positive flower transitions to new flowers and nest returns	last transition	exponential
OD_exp[−]	negative flower transitions to new flowers and nest returns	last transition	exponential
OD_lin[±]	positive and negative flower transitions to new flowers and nest returns	last transition	exponential
OD_lin[+]	positive flower transitions to new flowers and nest returns	last transition	exponential
OD_lin[−]	negative flower transitions to new flowers and nest returns	last transition	exponential
OS_N_exp[±]	positive and negative flower transitions and nest returns	last transition	exponential
OS_N_exp[+]	positive flower transitions and nest returns	last transition	exponential
OS_N_exp[−]	negative flower transitions and nest returns	last transition	exponential
OS_N_lin[±]	positive and negative flower transitions and nest returns	last transition	linear
OS_N_lin[+]	positive flower transitions and nest returns	last transition	linear
OS_N_lin[−]	negative flower transitions and nest returns	last transition	linear
OD_N_exp[±]	positive and negative flower transitions to new flowers and nest returns	last transition	exponential
OD_N_exp[+]	positive flower transitions to new flowers and nest returns	last transition	exponential
OD_N_exp[−]	negative flower transitions to new flowers and nest returns	last transition	exponential
OD_N_lin[±]	positive and negative flower transitions to new flowers and nest returns	last transition	exponential
OD_N_lin[+]	positive flower transitions to new flowers and nest returns	last transition	exponential
OD_N_lin[−]	negative flower transitions to new flowers and nest returns	last transition	exponential

Table A.1: Overview of structural model configurations and names. The first two letters represent the general model family, RB stands for route-based, OD stands for online-dubois, and OS stands for online simplified. In route-based models, reinforcement is applied to all traversed edges upon nest return. In OD models, reinforcement is applied to the most recently traversed edge if a new flower is visited. In OS models, reinforcement is applied on a rolling basis to transitions leading to flowers. In online reinforcement models, the $\_N$ suffix indicates that nest returns may also be reinforced. $+, -, \pm$ indicate if edge weights may change upon encounters, also referred to as positive experiences $(+)$, or non-encounter also referred to as negative experiences $(-)$ or both encounters and non-encounter $(\pm)$

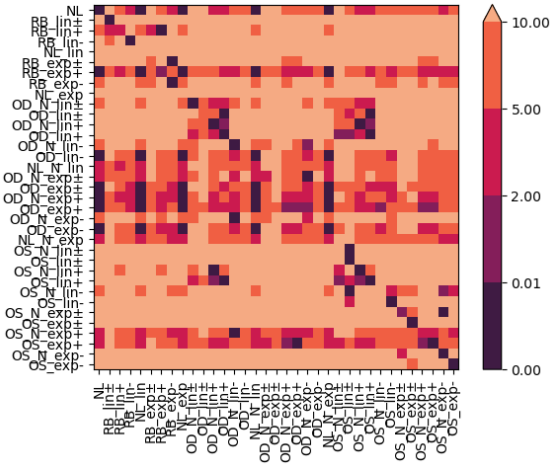

Figure A.2: Confusion Matrix showing mean ΔBIC over 200 model runs where data-generating models are row headers, and models fit to generated data are column headers.

Bibliography

[1] M. Agarwal and V. Sunil. Basic behavioural patterns in insects and applications of behavioural manipulation in insect pest management. *Journal of Entomology and Zoology Studies*, 8(2):991–996, 2020.

[2] D. J. Anderson. Optimal foraging and the traveling salesman. *Theoretical Population Biology*, 24(2):145–159, 1983.

[3] A. Buatois, J. Mailly, T. Dubois, and M. Lihoreau. A comparative analysis of foraging route development by bumblebees and honey bees. *Behavioral Ecology and Sociobiology*, 78(1):8, 2024.

[4] C. Buehlmann and P. Graham. Innate visual attraction in wood ants is a hardwired behavior seen across different motivational and ecological contexts. *Insectes Sociaux*, 69(2):271–277, 2022.

[5] A. Chakrabarti and J. K. Ghosh. AIC, BIC and recent advances in model selection. In *Philosophy of Statistics*, pages 583–605. North-Holland, 2011.

[6] N. Chatterjee, D. Wolfson, D. Kim, J. Velez, S. Freeman, N. M. Bacheler, K. Shertzer, J. C. Taylor, and J. Fieberg. Modelling individual variability in habitat selection and movement using integrated step-selection analysis. *Methods in Ecology and Evolution*, 15(6):1034–1047, 2024. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.14321.

[7] L. Chittka, J. Kunze, C. Shipman, and S. L. Buchmann. The significance of landmarks for path integration in homing honeybee foragers. *The Science of Nature*, 82(7):341–343, 1995.

[8] D. Clarke, E. Morley, and D. Robert. The bee, the flower, and the electric field: electric ecology and aerial electroreception. *Journal of Comparative Physiology A*, 203(9):737–748, 2017.

[9] M. Collett, D. Harland, and T. S. Collett. The use of landmarks and panoramic context in the performance of local vectors by navigating honeybees. *J Exp Biol*, 205(6):807–814, 2002.

[10] C. Cotar and D. Thacker. Edge- and vertex-reinforced random walks with super-linear reinforcement on infinite graphs. *Annals of Probability*, 45(4), 2017.

[11] B. Davis. Reinforced random walk. *Probability Theory and Related Fields*, 84(2):203–229, 1990.

[12] M. Dorigo and L. M. Gambardella. Ant colonies for the travelling salesman problem. *Bio Systems*, 43(2):73–81, 1997.

[13] S. S. Dragomir. Some inequalities for applications to weighted means. *Palestine Journal of Mathematics*, 9(1):537–548, 2020.

[14] T. Dubois, C. Pasquaretta, A. B. Barron, J. Gautrais, and M. Lihoreau. A model of resource partitioning between foraging bees based on learning. *PLOS Computational Biology*, 17(7):1–19, 2021. Publisher: Public Library of Science.

[15] S. J. England and D. Robert. Electrostatic pollination by butterflies and moths. *Journal of the Royal Society Interface*, 21(216):20240156, 2024. Publisher: Royal Society.

[16] D. Erhard, T. Franco, and G. Reis. The directed edge reinforced random walk: The ant mill phenomenon. *Journal of Statistical Physics*, 190(1):18, 2022.

[17] S. N. Ethier and T. G. Kurtz. *Markov processes : characterization and convergence*. Wiley series in probability and mathematical statistics. Wiley, New York, 1986.

[18] A. Feuerbacher. Pollinator declines, international trade and global food security: Reassessing the global economic and nutritional impacts. *Ecological Economics*, 232:108565, 2025.

[19] R. Freeman. Charles darwin on the routes of male humble bees. *Bull Br Mus Nat Hist*, 3, 1968.

[20] B. Gibson, M. Wilkinson, and D. Kelly. Let the pigeon drive the bus: pigeons can plan future routes in a room. *Animal Cognition*, 15(3):379–391, 2012.

[21] A. Goldshtein, M. Handel, O. Eitan, A. Bonstein, T. Shaler, S. Collet, S. Greif, R. A. Medellín, Y. Emek, A. Korman, and Y. Yovel. Reinforcement learning enables resource partitioning in foraging bats. *Current Biology*, 30(20):4096–4102.e6, 2020.

[22] T. M. Gottfried. Theory and application of a pólya urn with non-linear feedback, 2024.

[23] C. Gómez-Martínez, M. A. González-Estévez, J. Cursach, and A. Lázaro. Pollinator richness, pollination networks, and diet adjustment along local and landscape gradients of resource diversity. *Ecological Applications*, 32(6):e2634, 2022.

[24] B. Heinrich. "majoring" and "minoring" by foraging bumblebees, bombus vagans: An experimental analysis. *Ecology*, 60(2):246–255, 1979. Publisher: Ecological Society of America.

[25] D. Janzen. Euglossine bees as long-distance pollinators of tropical plants. *Science*, 171(3967):203–205, 1971. Num Pages: 3 Place: Washington Publisher: Amer Assoc Advancement Science Web of Science ID: WOS:A1971I224800030.

[26] A. J. King and H. H. Marshall. Optimal foraging. *Current Biology*, 32(12):R680–R683, 2022. Publisher: Elsevier.

[27] F. Le Moël, T. Stone, M. Lihoreau, A. Wystrach, and B. Webb. The central complex as a potential substrate for vector based navigation. *Frontiers in Psychology*, 10:380097, 2019. Publisher: Frontiers.

[28] M. Lihoreau, L. Chittka, and N. E. Raine. Travel optimization by foraging bumblebees through readjustments of traplines after discovery of new feeding locations. *The American naturalist*, 176(6):744–757, 2010. Place: United States.

[29] M. Lihoreau, L. Chittka, and N. E. Raine. Trade-off between travel distance and prioritization of high-reward sites in traplining bumblebees. *Functional Ecology*, 25(6):1284–1292, 2011. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2435.2011.01881.x.

[30] M. Lihoreau, N. E. Raine, A. M. Reynolds, R. J. Stelzer, K. S. Lim, A. D. Smith, J. L. Osborne, and L. Chittka. Radar tracking and motion-sensitive cameras on flowers reveal the development of pollinator multi-destination routes over large spatial scales. *PLOS Biology*, 10(9):e1001392, 2012. Publisher: Public Library of Science.

[31] M. Lihoreau, N. E. Raine, A. M. Reynolds, R. J. Stelzer, K. S. Lim, A. D. Smith, J. L. Osborne, and L. Chittka. Unravelling the mechanisms of trapline foraging in bees. *Communicative & integrative biology*, 6(1):e22701, 2013. Place: United States.

[32] H. F. Löchel and D. Heider. Chaos game representation and its applications in bioinformatics. *Computational and Structural Biotechnology Journal*, 19:6263–6271, 2021.

[33] Q. Ma, A. Johansson, A. Tero, T. Nakagaki, and D. J. T. Sumpter. Current-reinforced random walks for constructing transport networks. *Journal of The Royal Society Interface*, 10(80):20120864, 2013.

[34] N. Marwan, M. Carmen Romano, M. Thiel, and J. Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5):237–329, 2007.

[35] R. Noser and R. W. Byrne. How do wild baboons (papio ursinus) plan their routes? travel among multiple high-quality food sources with inter-group competition. *Animal Cognition*, 13(1):145–155, 2010.

[36] K. Ohashi, A. Leslie, and J. D. Thomson. Trapline foraging by bumble bees: V. effects of experience and priority on competitive performance. *Behavioral Ecology*, 19(5):936–948, 2008.

[37] K. Ohashi and J. D. Thomson. Efficient harvesting of renewing resources. *Behavioral Ecology*, 16(3):592–605, 2005. Place: United Kingdom Publisher: Oxford University Press.

[38] K. Ohashi, J. D. Thomson, and D. D'Souza. Trapline foraging by bumble bees: IV. optimization of route geometry in the absence of competition. *Behavioral Ecology*, 18(1):1–11, 2007.

[39] C. H. Papadimitriou. The euclidean travelling salesman problem is NP-complete. *Theoretical Computer Science*, 4(3):237–244, 1977.

[40] R. Pemantle. A survey of random processes with reinforcement. *Probability Surveys [electronic only]*, 4:1–79, 2007. Publisher: Sponsored by Institute of Mathematical Statistics and by the Bernoulli Society.

[41] H. P. Possingham. The distribution and abundance of resources encountered by a forager. *The American Naturalist*, 133(1):42–60, 1989. Publisher: [The University of Chicago Press, The American Society of Naturalists].

[42] S. G. Potts, J. C. Biesmeijer, C. Kremen, P. Neumann, O. Schweiger, and W. E. Kunin. Global pollinator declines: trends, impacts and drivers. *Trends in Ecology & Evolution*, 25(6):345–353, 2010.

[43] A. E. Raftery. Bayesian model selection in social research. *Sociological Methodology*, 25:111–163, 1995. Publisher: [American Sociological Association, Wiley, Sage Publications, Inc.].

[44] A. M. Reynolds, M. Lihoreau, and L. Chittka. A simple iterative model accurately captures complex trapline formation by bumblebees across spatial scales and flower arrangements. *PLoS Computational Biology*, 9(3):e1002938, 2013.

[45] U. E. Schlägel, M. A. Lewis, and L. Börger. Detecting effects of spatial memory and dynamic information on animal movement decisions. *Methods in ecology and evolution /*, 5(11), 2014. Place: Hoboken, N.J. : Publisher: John Wiley.

[46] M. V. Srinivasan, S. Zhang, M. Altwein, and J. Tautz. Honeybee navigation: Nature and calibration of the "odometer". *Science*, 287(5454):851–853, 2000.

[47] T. Stone, B. Webb, A. Adden, N. B. Weddig, A. Honkanen, R. Templin, W. Wcislo, L. Scimeca, E. Warrant, and S. Heinze. An anatomically constrained model for path integration in the bee brain. *Current Biology*, 27(20):3069–3085.e11, 2017.

[48] B. A. G. Sutton R.S. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[49] J. Tautz, S. Zhang, J. Spaethe, A. Brockmann, A. Si, and M. Srinivasan. Honeybee odometry: Performance in varying natural terrain. *PLOS Biology*, 2(7):e211, 2004. Publisher: Public Library of Science.

[50] M. C. Tello-Ramos, T. A. Hurly, and S. D. Healy. Traplining in hummingbirds: flying short-distance sequences among several locations. *Behavioral Ecology*, 26(3):812–819, 2015. _eprint: https://academic.oup.com/beheco/article-pdf/26/3/812/13897864/arv014.pdf.

[51] J. D. Thomson, M. Slatkin, and B. A. Thomson. Trapline foraging by bumble bees: II. definition and detection from sequence data. *Behavioral Ecology*, 8(2):199–210, 1997. _eprint: https://academic.oup.com/beheco/article-pdf/8/2/199/825317/8-2-199.pdf.

[52] T. Wang, K. Choi, and H. Wang. Derivations of animal movement models with explicit memory, 2024.

[53] J. L. Woodgate, J. C. Makinson, K. S. Lim, A. M. Reynolds, and L. Chittka. Continuous radar tracking illustrates the development of multi-destination routes of bumblebees. *Scientific Reports*, 7(1):17323, 2017. Publisher: Nature Publishing Group.