Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Ahmed Roman	Date

Learning new physics from biology and data

By

Ahmed Roman
Doctor of Philosophy

Physics

Ilya Nemenman, Ph.D
Advisor

Dr. Gordon Berman, Ph.D
Committee Member

Dr. Eric Weeks, Ph.D
Committee Member

Dr. Samuel Sober, Ph.D
Committee Member

Dr. Justin Burton, Ph.D
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D
Dean of the James T. Laney School of Graduate Studies

Date

Learning new physics from biology and data

By

Ahmed Roman
B.S., Virginia Polytechnic Institute and State University, 2013
B.S., Virginia Polytechnic Institute and State University, 2015
M.S., Virginia Polytechnic Institute and State University, 2015

Advisor: Ilya Nemenman, Ph.D

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Physics
2022

Abstract

Learning new physics from biology and data

By Ahmed Roman

In an attempt to understand various phenomena in living systems and properties of data, we found new physical phenomena that were previously unstudied. In an effort to better understand interface propagation in *Dictyostelium discoideum*, we deduced a new interface growth rule that may describe a broad class of biological interfaces. Simulations and analysis of the new growth rule gave rise to a new universality class of interface growth with three dynamic exponents instead of the usual two. By studying thermal learning in *C. elegans*, we constructed a new model of associative learning that incorporates classical and operant conditioning, and generalization. Our new model gives rise to a mechanism that explains learning phenomena such as extinction and spontaneous recovery, which previous learning models could not explain. Finally, we attempt to understand the underlying process of Bayesian entropy estimation. We show that Bayesian entropy estimators depend on a few emergent data statistics that rely on states that are sampled once or more.

Learning new physics from biology and data

By

Ahmed Roman
B.S., Virginia Polytechnic Institute and State University, 2013
B.S., Virginia Polytechnic Institute and State University, 2015
M.S., Virginia Polytechnic Institute and State University, 2015

Advisor: Ilya Nemenman, Ph.D

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Physics
2022

Acknowledgement

We have finally arrived at the proximal destination: this thesis, after much pain and hardship. I am grateful for the advice of Ilya Nemenman along the way. His advice and attention to detail have improved me as a scientist and thinker. You are one of the few people who had a profound positive impact on me, Ilya. You are, indubitably, one of the most creative people I know. I am aware of many of your (not so obvious) attempts to help. Thanks, Ilya.

I am grateful to William Ryu for his advice. His mature and perceptive views that he shared with me before I joined Ilya's group served me well. His humbleness might mask his thoughtful and meticulous nature to the casual onlooker. Will, I have deep respect for you. There is much still to be learned from you.

I am grateful to have worked with Daniel Weissman. I have learned much from you, and you were kind enough to patiently teach me. I am grateful to Benjamin Good for getting me started on population genetics and perturbation theory. I am thankful for the time and conversations I had with Ajit Srivastava. I am also grateful for the advice of Gordon Berman during my times of need. I would like to thank Justin Burton for his help when I had personal troubles and otherwise.

I am fortunate that I met wonderful people in the Nemenman Lab: Eslam Abdelaleem, Michael Martini, Baohua Zhou, Damian Hernandez, David Hofmann, Itai Pinkovezky, Michael Pasek, Catalina Rivera, Caroline Holmes, Mia Morrell, Audrey Sederberg, and many others. Thanks for your friendship and encouragement, and for making N117 fun.

I am thankful for all the help that Barbara Conner has given me throughout my Ph.D. here at Emory. You saved me significant time and gave me good advice. I really appreciate everything you did for me and my colleagues. Thank you, Barb.

I would like to thank my friends Mostafa Ali, Haithem Taha, and Shariar Kabir for their support and ongoing advice on personal and academic matters despite their limited time. You are the best friends I could have hoped for. Your honesty, truthfulness, and moral character are exemplary. I am better because of your friendship. You embody what it means to be a responsible Muslim, and I am proud to be your friend.

My mom is the best woman I know, despite the many significant hardships and obstacles that stood in her way. Despite the systemic injustice she endured and the sacrifices she made to protect herself and her children, she stayed strong and persevered. Despite our very limited time together, my inclination to rarely give up is in large part due to her. You are a role model for what it means to be responsible. I love you, mom.

My father spent significant time teaching me algebra, geometry, and theology. Without the thousands of problems we solved together, I might not have started this thesis. Even though the time I spent with my father was limited, I am still grateful for what we had.

I want to thank my incredibly intelligent and gorgeous wife, Maryam, who is also my best friend. You are a wonderful, truthful, honest, and wise woman. It is in large part due to your wisdom and patience that I am a better person today than I was when we first met. Spending many nights with me on campus so that I can finish my work is partly why this thesis is written now.

Yosra, you are my favorite. Your support helped me more than you could imagine. I love you more than I let on. Despite our strange and complicated childhood, you were the constant in an ever-changing environment. Even though you are smarter than me, I have learned a few things you might not have learned yet. Listen to your older brother, and we might reach the final destination together.

I am thankful to Allah for everything I have. I am nothing without Your wisdom. I am nothing without Your provision. I am not lost because of Your knowledge. Our Lord! Let not my heart deviate from the truth after You have guided me, and grant me mercy from You. Truly, You are the Bestower. My Lord, enable me to be grateful for Your favor which You have bestowed upon me and upon my parents and to work righteousness of which You will approve and make righteous for me my offspring. Indeed, I have repented to You, and indeed, I am of the Muslims.

سورة الحج

يَا أَيُّهَا النَّاسُ ضُرِبَ مَثَلٌ فَاسْتَمِعُوا لَهُ ۚ إِنَّ الَّذِينَ تَدْعُونَ مِنْ دُونِ اللَّهِ لَنْ يَخْلُقُوا ذُبَابًا وَلَوِ اجْتَمَعُوا لَهُ ۖ وَإِنْ يَسْلُبْهُمُ الذُّبَابُ شَيْئًا لَا يَسْتَنْقِذُوهُ مِنْهُ ۚ ضَعُفَ الطَّالِبُ وَالْمَطْلُوبُ ﴿٧٣﴾

Surah al-Hajj

73. O, people! A parable is presented, so listen to it: those you call upon besides Allah will never create a fly, even if they came together for that purpose. And if the fly snatches anything from them, they cannot recover it from it. Weak is the pursuer and the pursued.

Table of Contents

List of Figures

Chapter 1 Introduction

Modeling of biological experiments typically relies on existing physical processes that are well understood, where the modeler attempts to recast the biological experiment in a mathematical language that best reveals this underlying known physical process. On rare occasions, new physics arises in the process of modeling because the biological phenomena are either a consequence of previously unstudied physical phenomena or inspire new physical phenomena. In flock dynamics, for example, two different modeling paradigms (the aligning of spins paradigm and the attractive/repulsive particle interactions through a potential paradigm) were integrated to explain how individuals in flocks maintain the coherence of the flock. This led to the discovery of new physics of propagation through new propagation laws and correlation exponents [15]. Much like this flocking example, we are interested in situations where previously unobserved physics phenomena manifest themselves in living systems. In this thesis, we construct a new rule for interface[1] growth that is inspired by the behavior of *Dictyostelium discoideum* cells and the signalling process between them. We also construct a new dynamical system of learning that manifests itself in how the roundworm *C. elegans* decides to move up and down a temperature gradient given its thermal history and feeding state. Finally, we study how features of data samples influence the entropy estimate of the distribution generating the data, and provide a reliable estimate of entropy in terms of those features. In each of these examples, we discover new physics. In the interface growth model, we found a new universality class of interface growth that has three instead of two dynamic exponents. In animal learning, we uncover a mechanism that explains salient features of learning that previous learning models failed to explain. In entropy estimation, we surprisingly found that a handful of data features determine the entropy estimates obtained from Bayesian estimators.

[1] An interface is the boundary between two distinguishable media.

1.1 Interfaces in nature

Interfaces are common in our everyday experience. Examples range from the boundary of the spilled coffee stain as it conquers a segment of the tablecloth to the fire burning a piece of paper or snow growing vertically on the side of the glass windows. It might be surprising that the morphology of these interfaces changes drastically depending on the scale at which we observe them. When we look at the moon, we see that its surface appears smooth. However, close-up observations show that the moon's surface is abounding with craters and hills. Collisions with space debris increase the craters of the moon over time, rendering the surface morphology dynamical. How can we describe the time-dependent morphology of an interface that seems to be smooth to the eye on large observation length scales but rough on short observation length scales? It turns out that the morphology of many of the interfaces that we observe in nature are *self-affine*, which means that if the interface/surface is described via a function I, then it satisfies $I(x_1, \cdots, x_n, t) \sim \left(\prod_{i=1}^{n+1} b_i^{-\alpha_i}\right) I(\alpha_1 x_1, \cdots, \alpha_n x_n, \alpha_{n+1} t)$, where $\sim$ is the asymptotic symbol and α_i is real and positive for all i [9]. This means that if we make different scale changes in directions parallel and perpendicular to the interface, then the morphology stays the same. This is contrasted with fractal interfaces, which appear the same on every observation scale. Through these different scale changes in different directions, we can show that the morphology of the interface at certain lengths and time scales is similar [2] to the morphology at other lengths and time scales. To organize these interfaces, symmetry principles codified in group theory are used to classify properties of rough interfaces [9]. Self-affine interfaces that satisfy the same symmetries will share the same long time and length scale statistical properties. The set of interfaces, whose microscopic details are different, but whose macroscopic statistical properties are the same are grouped in what is called a *universality class* [3, 9, 37, 55, 86].

[2] Similar means that the morphology of the interface at certain times and lengths is related via a similarity transformation to the morphology at other times [9].

1.1.1 Ballistic deposition

Over the years, many models of interface growth have been developed. Examples include random deposition, ballistic deposition, the Eden model, solid on solid models, and many more [9]. Each of these models stresses certain features of the growth process that others do not. Here we focus on the ballistic deposition model and attempt to clarify the salient features of interface growth through this model.

The simplest version of the growth process in the ballistic deposition model occurs on a lattice of length L, but off-lattice versions have been studied as well. A site $i \in \{1, \cdots, L\}$ is sampled randomly, then a particle is deposited at the chosen site from a height much larger than the maximum height of the interface. In the simplest version of the model, the particle falls perpendicular to the interface until it sticks to it. The deposited particles form an aggregate of a particular geometry.

1.1.2 Roughening

The interface of the aggregate is defined as the set of particles are highest in each column. After depositing N particles, the time $t = N/L$ is the average number of deposited layers. A quantitative description of the growth process is obtained by defining the mean height function of the aggregate

$$\bar{h}(t) = \frac{1}{L} \sum_{i=1}^{L} h(i,t), \tag{1.1}$$

where $h(i,t)$ is the height of the i^{th} column at time t. In ballistic deposition, a particle arriving at site i sticks to the interface according to the rule

$$h(i,t) = \max\{h(i-1,t), h(i,t)+1, h(i+1,t)\} \tag{1.2}$$

with the initial condition $h(i,0) = 0$ for all $1 \leq i \leq L$. An important quantity in studies of interface growth is the width of the interface $w(L,t)$ which characterizes the roughness of the interface. The width is defined as the root mean squared fluctuation in the height of the

interface

$$w(L,t) = \left(\frac{1}{L} \sum_{i=1}^{L} [h(i,t) - \bar{h}(t)]^2 \right)^{1/2}. \tag{1.3}$$

Since the initial condition is a flat interface, the width initially is zero. As particles are deposited, the interface roughens. A typical width vs. time plot has three different regimes: a Poisson regime, a growth regime, and a saturation regime [9]. For $t < 1$, the Poisson regime dominates and $w(t) \sim t^{1/2}$. For $1 < t \ll t_\times$, the interface is in the growth regime where the width scales as $w \sim t^\beta$. The exponent β is called the *growth exponent* and it summarizes the time dependence of the interface roughening. Due to the finite size of the lattice, the power law growth does not continue indefinitely and for times $t \gg t_\times$, the width saturates to a value $w_{sat} \sim L^\alpha$. The exponent α is called the *roughness exponent*, and it summarizes the roughness of the saturated interface. The cross-over time $t_\times$ occurs when $t_\times \sim L^z$ where z is the *dynamic exponent*.

1.1.3 Dynamic scaling

The exponents, α, β and z are not independent. The asymptotic relation $t_\times^\beta \sim L^\alpha$ is used to calculate the cross-over time $t_\times$. This yields $t_\times \sim L^{\alpha/\beta}$ and $z = \alpha/\beta$. If we had multiple simulations of the ballistic deposition model for various lattice sizes, we would obtain multiple width vs. time curves $w(t, L_1), \cdots, w(t, L_k)$, which would saturate at different times $t_{\times,1}, \cdots, t_{\times,k}$ and different saturation widths $w_{sat}(L_1), \cdots, w_{sat}(L_k)$. If we scale each width by its saturation value, $\frac{w(t,L_j)}{w_{sat}(L_j)}$, then the new width vs. time curve will saturate at $\frac{w(t,L_j)}{w_{sat}(L_j)} = 1$. Plotting each curve $\frac{w(t,L_j)}{w_{sat}(L_j)}$ against $t/t_{\times,j}$, all curves will saturate at the same characteristic time $t/t_{\times,j} = 1$. These two observations suggest that $\frac{w(t,L)}{w_{sat}(L)}$ is a function of $t/t_\times$ only, i.e.

$$\frac{w(L,t)}{w_{sat}(L)} \sim f\left(\frac{t}{t_\times}\right), \tag{1.4}$$

where $f(\cdot)$ is a scaling function. Replacing w_{sat} and $t_\times$ by their asymptotic values L^α and L^z, we obtain the *Family-Vicek scaling relation*

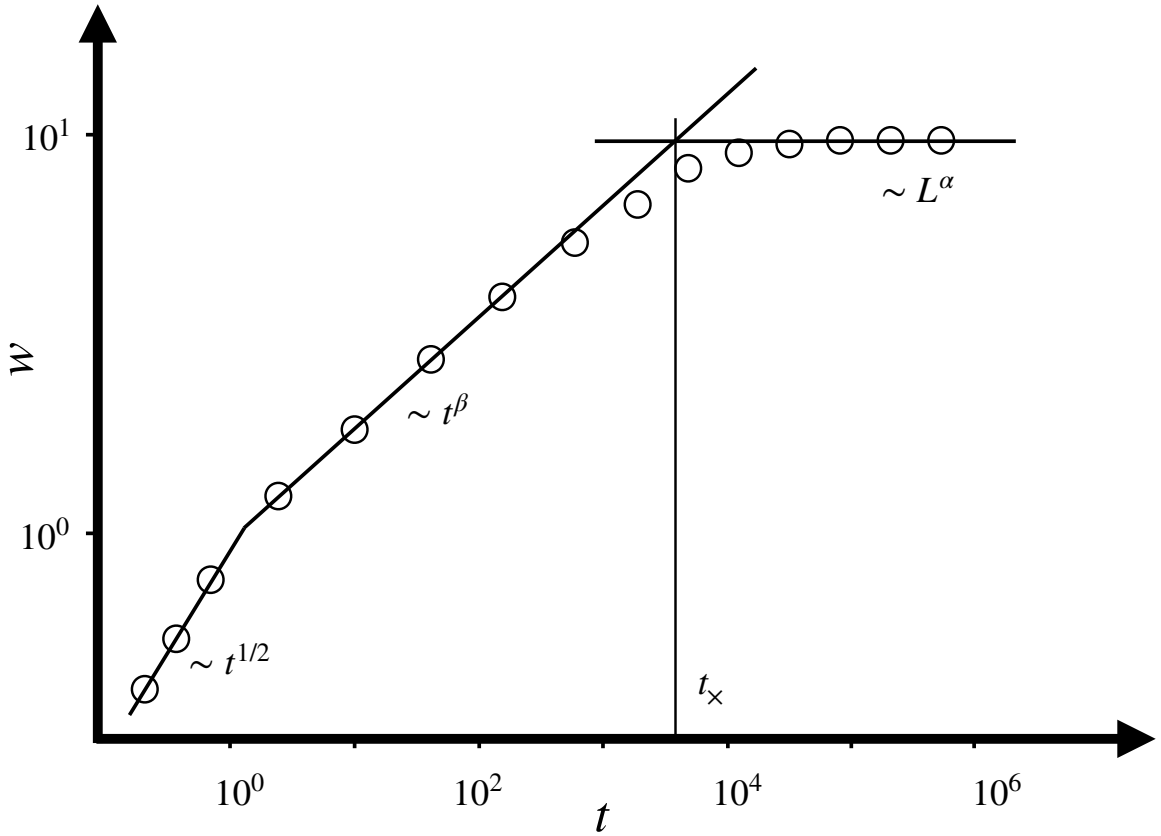

Figure 1.1: width vs. time curve shows three regimes: a Poisson growth regime, a growth regime, and a saturation regime that scale as $t^{1/2}$, t^{β}, and L^{α} respectively. The cross-over time scales as L^{z}.

$$w(L,t) \sim L^{\alpha} f\left(\frac{t}{L^{z}}\right). \tag{1.5}$$

1.1.4 Interface growth with memory

Historically, studies of interfaces were confined to systems where there was no feedback between the constituents that compose the interface and the latter's morphology [9]. In Chapter 2, we abstract out the rules for a new model of interface growth from the behavior of a classic biological model of collective signaling, collective motility, and development: *Dictyostelium discoideum*. Dictyostelium cells secrete the chemical cyclic AMP and use it as a signal. In this system, a spreading wave of cyclic AMP activates a cell, but only if the temporal derivative of the cyclic AMP concentration is positive and large [127]. This

induces feedback between the interface (the boundary separating cells that secreted AMP from those that did not), and the interface's constituents (the front of the cells that secreted AMP). By simulating these interface growth rules and studying the ensuing interface dynamics, we discovered a new universality class of interface growth. Given that there are few universality classes of interface growth for non-living systems, this new universality class represents a new physics discovery that is entirely inspired by biology.

1.2 Animal learning

A fundamental question in behavioral neuroscience is how animals make decisions and select actions in response to punishment and reward. This question has been experimentally studied in behavioral psychology via classical and operant conditioning paradigms. There is significant evidence about the associations that influence various aspects of learned behavior. Animals use associations between unconditioned stimuli US (e.g., food or tail pinch) and conditioned stimuli CS (e.g., light, tone) to predict when to expect a reward or a punishment [32, 60, 84, 90, 126]. Despite plentiful evidence of learning based on associations, models of conditioning that quantitatively account for intricacies of behavior remain rare [36, 87, 88, 139].

Early models of animal learning explain overshadowing [3], blocking[4] and conditioned inhibition[5] by assuming that learning occurs when unexpected events occur [97]. The difference between what is expected and what occurs, the error, is proportional to the rate of change of the association strength between the conditioned and unconditioned stimuli. The sign of the error is what increases the association strength if the error is positive (what

[3]Overshadowing is when two or more stimuli are present, and one stimulus produces a stronger response than the other because it is more relevant or salient.

[4]The blocking effect occurs when the attempt to pair a second conditioned stimulus CS_2 to an unconditioned stimulus US is blocked due to a preexisting association between the US and another conditioned stimulus CS_1. For example, we pair light (CS_1) and food (US). This causes the animal to salivate upon the presentation of light. We then attempt to pair light (CS_1) and tone (CS_2) to food (US). The combination induces salivation. When we present the tone alone after that (CS_2), it does not induce salivation. So tone is blocked by the pairing between light and food.

[5]In conditioned inhibition, a conditioned stimulus predicts the absence of the unconditioned stimulus.

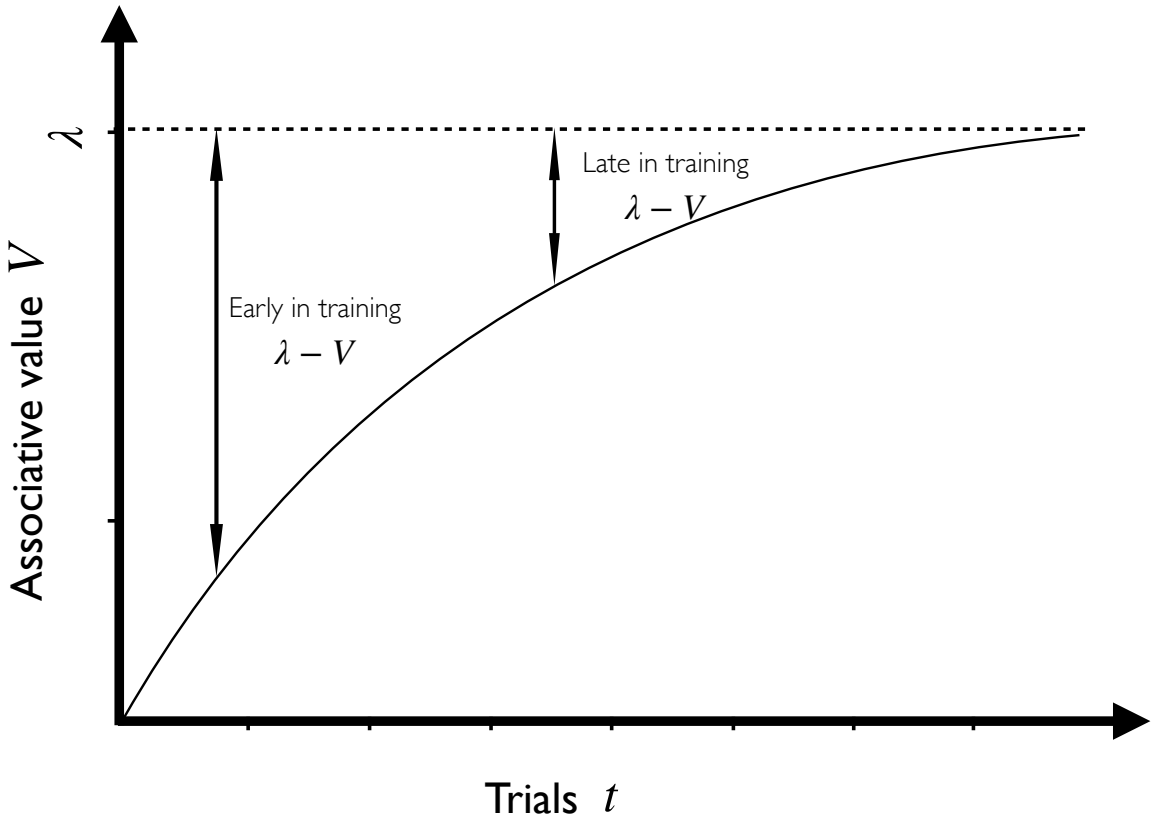

Figure 1.2: Associative strength V vs trial number t. The curve shows a diminishing error effect where the error is large early in training and smaller later in training.

occurs exceeds expectations), and diminishes it if the error is negative (what occurs fails expectations). The animal then chooses an action (or makes a decision) based on the association strength.

1.2.1 The Rescorla-Wagner model

Rescorla and Wagner [97] assumed that the magnitude of the error, or *surprise* (a proxy for the effectiveness of a US), depends on the difference between the US presented and what the learner expects. They further assumed that there was a relationship between the conditioned stimuli that precede the US and the expectation of the US. This means that a strong (weak) response to conditioned stimuli indicates a high (low) expectation of the US.

Using λ to represent the US presented on a given trial and V to represent the expected value of the stimuli preceding the presentation of the US, the magnitude of the *surprise* is the difference $\lambda - V$ between what occurs and what is expected. At the beginning of

the conditioning experiment, the surprise $\lambda - V$ will be large. As more CS-US pairs are presented, the surprise diminishes and V converges to λ. This is summarized through

$$\Delta V = \alpha(\lambda - V), \tag{1.6}$$

where α is the learning parameter which is related to the salience of the US and CS pair.

1.2.2 Rescorla-Wagner and the blocking effect

Imagine the following experimental paradigm: In phase one, we repeatedly present the (CS, US) pair (A,US) until a strong conditioned response to A occurs. In phase two, we repeatedly present the CSs $A + B$ with the US. Finally, we present the CS B alone in test trials to see if it elicits a conditioned response. Interesting, little response occurs to the presentation of B during the testing phase, despite repeated pairing with the US. This is called the *blocking effect* and it is a non-trivial prediction of the Rescorla-Wagner model.

At the end of phase one, the association strength V_A between A the US reaches its maximum value of λ. In phase two, there are two cues (A and B) whose combined association with the US is V_{A+B}. Applying the Rescola-Wagner model to the stimulus B, we obtain $\Delta V_B = \alpha(\lambda - V_{A+B})$. In the beginning of phase two, $V_B = 0$. Thus, the combined association $V_{A+B} = V_A + V_B = \lambda + 0 = \lambda$. As a result, $\Delta V_B = \alpha(\lambda - V_{A+B}) = \alpha(\lambda - \lambda) = 0$. This means that the Rescorla-Wagner model predicts that stimulus B does not acquire an association with the US. In other words, stimulus A has blocked stimulus B from forming an association with the US.

1.2.3 Rescorla-Wagner and associative strength loss despite pairings with the US

Imagine that in phase one of an experiment we separately pair stimuli A and B with the US (one food pellet) until each stimulus is perfectly predictive of the US (a single food pellet). At the end of phase one, $V_A = \lambda$ and $V_B = \lambda$. In phase two, the compound stimulus $A+B$ is presented together for the first time and is followed by the US (a single food pellet).

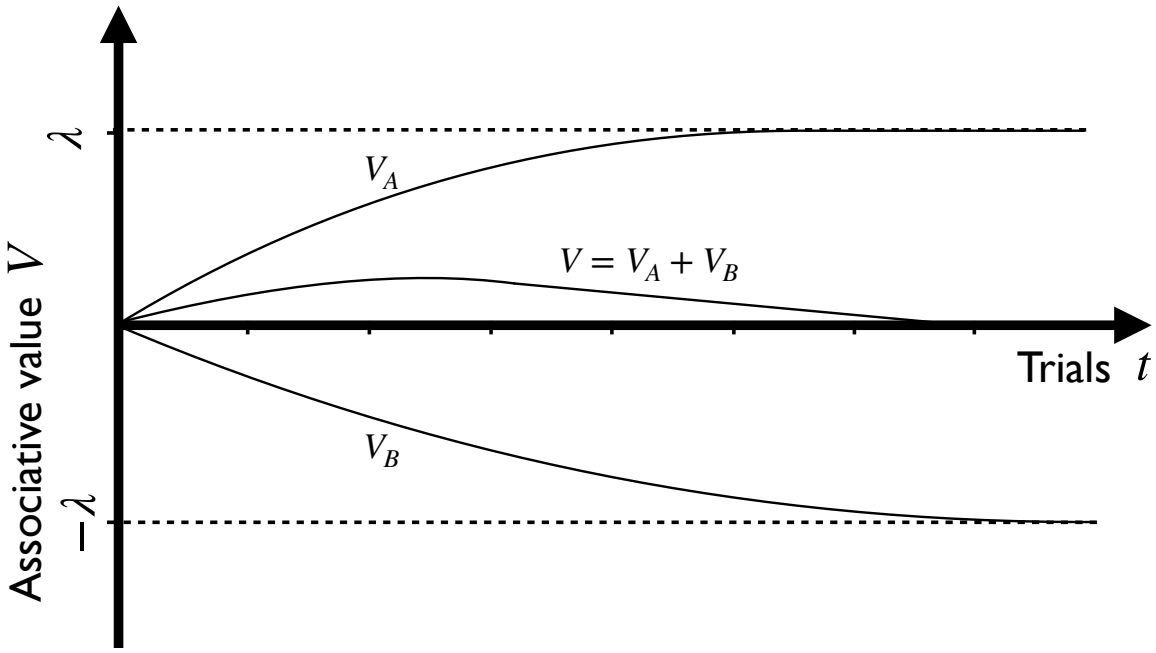

Figure 1.3: Associative strength V vs trial number t for conditioned inhibition experiments. V_A and V_B denote the associative strength for the reinforcing trials and non-reinforcing trials respectively. The combined strength $V = V_A + V_B$ shows non-monotonic behavior.

Naively, we expect that the association strength V_A and V_B to remain unchanged since the US did not change. But the Rescorla-Wagner model predicts a decrease in both V_A and V_B. In the beginning of phase two, $V = V_A + V_B = 2\lambda$ which implies

$$\Delta V = \alpha(\lambda - V) = \alpha(\lambda - 2\lambda) = -\lambda. \tag{1.7}$$

This means there is an over-expectation of the US (2 food pellets instead of the one presented). This negative error reduces the association strength V and its constituents V_A and V_B.

1.2.4 Rescorla-Wagner and conditioned inhibition

In a conditioned inhibition experiment, we randomly alternate between two different types of trials: reinforcing and non-reinforcing. In the reinforcing trials, we pair the reinforcing CS A (tone) with the US (a food pellet). In the non-reinforcing trials, we present both the reinforcing and non-reinforcing stimuli $A+B$ (tone and light) but in the absence of the US (so no food is presented). After many repetitions of both trial types, the associative

strength $V_A = \lambda$ and the combined stimulus $V = V_A + V_B = 0$. This is because stimulus A perfectly predicts the US while the combined stimulus predicts no US. The Rescorla-Wanger model predicts that $V_B = -\lambda$. This means that while positive associative strength V_A predicts presence of the US, negative associative strength V_B is the model's explanation of how stimulus B predicts the absence of the US.

Interestingly, if the learning parameter is different between the two trial types (reinforcing trials having a larger learning rate[6] than that of the non-reinforcing trials), then $V = V_A + V_B$ will initially increase and subsequently decrease to zero after many trials. This occurs because the rate at which V_A approaches λ is larger than the rate at which V_B approaches $-\lambda$. Under the difference in time scales assumption, this non-monotonic behavior of V becomes a prediction of the Rescorla-Wagner model.

1.2.5 Problems with the Rescorla-Wagner model

While models such as the Rescorla-Wagner model explain some features of associative learning, such models of conditioning are not of sufficient complexity to serve as *quantitative* models of real life animal behavior [84]. They fail to account for certain *qualitative* features of conditioning, including (i) extinction of an association that is no longer presented (not merely unlearning) and its subsequent spontaneous recovery [31, 33, 84], (ii) existence of many reinforcement cues (some potentially associated with non-reward-driven actions), whose collective output influences the behavior [29, 84], (iii) asymmetric responses to reward vs. punishment [31] and to conditioned association vs. conditioned inhibition [84], (iv) generalization among similar, but different cues [84], and so on. The weakness of models is a consequence of the rarity of informative experiments: measuring behavior with high spatial and temporal resolution for a long time is difficult; behavior itself often modifies the conditioning contingency; behavior is often measured as a discrete and noisy signal, and hence is not a reliable readout of a CS-US association; and biology of different reinforcement pathways is often unclear.

[6] Or equivalently a smaller time scale as we show in section 3.6.1

1.2.6 A new learning model

In Chapter 3, we abstract out a model of animal learning that includes classical-conditioning, operant-conditioning, generalization, and extinction from long term, high spatio-temporal resolution thermal learning experiments of the model system *C. elegans*. By analyzing the data, we show that the structure of thermal learning in the brain of the round-worm *C. elegans* must depend on at least four dynamics processes. We learn the parameters of the model by fitting the themotactic index of wild type worms. We then predict the four hour long thermotactic trajectory of a certain mutant without any fitting parameters. We also predict that the phenotypic effect of various genetic perturbations will modify the star-vation avoidance behavior. We finally predict that there is a circuit in the brain of the worm that computes a weighted average of the temperatures for which the worm is either hungry or cannot access food (it is not clear how to distinguish these two states from each other). We discovered that worms *simultaneously associate a single CS (temperature) with mul-tiple reinforcement cues (absence of food and the presence of a preferred temperature) to make a decision*. This is likely a *generic feature of learning* in real animals that is discov-ered from biology. By failing to include this important feature, previous models of learning in behavioral psychology failed to describe certain aspects of associative learning. Our new model of learning, therefore, represents a new physics discovery that is the product of detailed analysis of the biology of learning in a real animal.

1.3 Entropy Estimation

Estimating mutual information between two random variables has become a common task in various scientific disciplines. Estimating the entropy of probability distributions of various data sets is a first step towards calculating mutual information in modern data analysis. A common problem is that the number of independent samples obtained in ex-periments is limited, so that many states are under-sampled and naïve entropy estimators are inaccurate. This occurs because there are often many states whose probability is less

than the inverse sample size, which means that these states are unlikely to be sampled. This means that, in order to obtain proper estimates of entropy, we must make assumptions on the tail(s) of the underlying distribution (where many states are unlikely to be sampled), then choose the tail hypothesis that is most likely given the samples. Previous studies found that the statistics of states that occur in data sets multiple times (coincidences) provide useful corrections to entropy estimates in the extremely under-sampled regime by selecting the most likely tail hypothesis from those allowed by the prior [5, 6, 81, 82]. These previously obtained corrections are largely numerical in nature and so provide little insight to which features of the dataset cause them. Before we delve in the process of how this problem is solved, we will first attempt to understand why coincidences are useful through a simpler problem.

1.3.1 The birthday problem

Suppose we are at a party of K people. What should K be so that the chance of getting two people with the same birthday is one half? For simplicity, we will assume that a year is 365 days (so we remove February 29^{th}) and that the probability of being born on any day of the year is 1/365 [7]. The probability of no birthday match among the K party goers is[8]

$$\text{Prob}[\text{no match}|K \text{ person party}] = \prod_{k=0}^{k=K-1} \frac{365-k}{365} \approx e^{-\binom{K}{2}/365}. \tag{1.8}$$

As a sanity check, we see that if $K = 1$, then probability of no match is one. From Eq. (1.8) we determine that at $K = 23$, the probability of one or more matches is about 50.7%. In fact, the probability of at least one match is

$$\text{Prob}[\text{at least one match}|K \text{ person party}] \approx 1 - e^{-\binom{K}{2}/365} \approx \begin{cases} 50.7\% & \text{for } K = 23 \\ 97\% & \text{for } K = 50 \\ 99.97\% & \text{for } K = 75. \end{cases} \tag{1.9}$$

[7] This turns out not be quite a correct assumption. There are seasonal effects that vary from country to country. It is more likely, for example, that babies are born about nine months after a holiday.

[8] we obtain this probability by observing that the first person has 365 possible birthdays, while the second person has only 364 once the first person's birthday is fixed etc.

These probabilities seem surprising. They say that already with only $K = 23$ people, the chance of at least two people having the same birthday is 50.7%. The surprise comes from the faulty intuition that there are 365 days a year and so for the probability of a match to be near a half we need about 365/2 people. The reason this intuition is faulty is that the quantity that matters here is the number of coincidences or the number of possible pairs that could have the same birthday. The number of pairs is $\binom{K}{2}$ is roughly speaking the important quantity here and this appears in the expected number of pairs X with the same birthday

$$\langle X \rangle = \frac{K(K-1)}{2 \cdot 365}. \tag{1.10}$$

The expected number of pairs with the same birthday at $K = 28$ is $\langle X \rangle \approx 1$ pair!.

1.3.2 Inferring the number of states from coincidences

Generalizing Eq. (1.10) to a year with d days, we see that the expected number of pairs with the same birthday is

$$\langle X \rangle = \frac{K(K-1)}{2 \cdot d}. \tag{1.11}$$

From this we see that the expected number of coincidences X is related to the number of days d via the formula

$$d = \frac{K(K-1)}{2 \cdot \langle X \rangle}. \tag{1.12}$$

This formula means that if we have access to the number of coincidences in sample party of size K, then we can estimate the number of days in year. Of course nothing about this analysis is particular to the birthday problem. In fact we can apply this argument to a sample of size K of a system with d equiprobable states. A coincidence is when a state is visited more than once in the sample. The entropy of this system then becomes

$$S \approx \log\left(\frac{K(K-1)}{2 \cdot \langle X \rangle}\right). \tag{1.13}$$

1.3.3 Problems with inference

From knowledge of the variance of the coincidences X, we can estimate the error in S. This estimate, however, suffers from a serious deficit. By looking at Eq. (1.11), we note that we do not typically have access to $\langle X \rangle$, but instead we have access to a realization of the random number of pairs X in the a specific sample of size K. Therefore, for any inference to occur X must at least be one. This happens on average whenever $K(K-1) \geq 2d$ or

$$K \gtrsim \sqrt{2d}. \tag{1.14}$$

That is we need the sample size to be at least of the order of square root of the number of states in order for us to expect at least one coincidence. Of course, this becomes more difficult as the number of states approaches infinity, which is the typical case. Also our estimate now will suffer from the problems associated with using X instead of $\langle X \rangle$ in Eq. (1.13). A similar argument to the one made here was first utilized by Ma to estimate entropy for random systems [69]. Ma generalized this argument for special cases where groups of states are not equally probable. Ultimately, we have to make assumptions on the underlying probability distribution over the states. Nemenman and colleagues [81, 82] made further progress on entropy estimation by making assumptions on the probabilities of the states from which the samples are obtained. Finally, Archer and colleagues [5] further relaxed the assumptions made by Nemenman and colleagues [81, 82]. Details of these works are out of the scope of this introduction and are available in Chapter 4.

1.3.4 What is new here?

In Chapter 4, we study the common and well established Pitman-Yor Mixture entropy estimator, which attempts to estimate the entropy of undersampled multinomial observations. This estimator makes an assumption that the tail of the observed distribution is either exponential, power law, or a mixture of the two. This Bayesian estimator is re-analyzed, and its entropy estimate is rewritten in terms of the coincidence vector $\mathcal{K}$, whose i^{th} element is the number of states sampled i or more times. Surprisingly, we found that the tail

hypotheses assumed by Bayesian estimators are equivalent to making assumptions about a few data statistics: the number of coincidences among the samples, and the dispersion of the coincidences.

These analytical coincidence-based estimates provide a method for understanding which coincidences influence the entropy estimates most, by allowing us to study synthetic data samples generated from known distributions. These emergent quantities that determine the entropy estimate represent a physics discovery that is inspired by data analysis.

1.4 Conclusion

Finally, in Chapter 5 we conclude with a discussion on further developments that naturally arise from the problems addressed in this thesis. We outline possible theoretical examination of animal learning through a theoretical lens in ways that may not have been as clear before. We also discuss how coupling of learning with other dynamical systems such as ecology may prove useful in understanding animals in more realistic environmental conditions. We also comment on using the entropy estimates found in Chapter 4 to obtain estimates of the mutual information between two random variables in the under-sampled regime.

Chapter 2 Ballistic deposition with memory: a new universality class of surface growth with a new scaling law

2.1 Summary

[1] Motivated by recent experimental studies in microbiology, we suggest a modification of the classic ballistic deposition model of surface growth, where memory of a deposition at a site induces more depositions at that site or its neighbors. By studying the statistics of surfaces in this model, we obtain three independent critical exponents: the growth exponent $\beta = 5/4$, the roughening exponent $\alpha = 2$, and the new (size) exponent $\gamma = 1/2$. The model requires a modification to the Family-Vicsek scaling, resulting in the dynamical exponent $z = \frac{\alpha+\gamma}{\beta} = 2$. This modified scaling collapses the surface width vs. time curves for various lattice sizes. This is a previously unobserved universality class of surface growth that could describe surface properties of a wide range of natural systems.

2.2 Introduction

Interface growth, and its ensuing roughening, is a paradigmatic nonequilibrium statistical physics process, with applications to many domains of physics [9]. Analytical, computational, and experimental studies have shown that the statistics of interface roughness in such processes usually is characterized by one of well-known universality classes: Poisson, Edwards-Wilkinson (EW), and Kardar-Parisi-Zhang (KPZ) [37, 55], as well as their extensions to quenched disorder, correlated noise, and so on [3, 86]. In the first, interface heights at every point are uncorrelated. In the second, peaks in the interface are smoothed through diffusion. Finally, in the third, nearby sites in the interface help each other grow,

[1] This chapter presents the paper [103] with the same title as this chapter. The work was done in collaboration with Ruomin Zhu and Ilya Nemenman.

resulting in a nonlinear amplification of fluctuations. Competition between the smoothing and the nonlinearity leads to the interface roughness that increases with time and eventually saturates at a system-size dependent value.

More concretely, we denote the height of a 1-d interface at point x at time t by $h(x,t)$. Then the standard deviation of the interface height defines the interface roughness

$$w(L,t) = \langle (h(x,t) - \langle h(x,t) \rangle_L)^2 \rangle_L^{1/2}, \tag{2.1}$$

and the average here is over a domain of size L. Such growth processes are generally characterized by three critical exponents: β, the *growth* exponent, which measures how the roughness grows with time; α, the *roughness* exponent, which parameterizes the dependence of the roughness of the saturated interface on the system size, and z, the dynamical exponent, which relates the time at which the width of the interface stops growing to the system size. The three exponents are related by the celebrated Family-Vicsek dynamical scaling [38]

$$w(L,t) \sim L^{\alpha} f(t/L^z), \quad \text{with} \tag{2.2}$$

$$f(u) \propto \begin{cases} u^{\beta} & u \ll 1 \\ 1 & u \gg 1 \end{cases}, \tag{2.3}$$

which results in $z = \alpha/\beta$.

What unites all of these cases is that there is no memory or inertia in the interface growth – deposition is Markovian in time. This is a reasonable assumption when the interface is built by or from stateless agents. However, when the agents are more complex, such as when they are living cells with a multitude of internal states, such memory-less assumption should be questioned. For example, in cyclic AMP signaling in *Dictyostelium discoideum*, which is a classic biological model of collective signaling, collective motility, and development, a spreading wave of cyclic AMP activates a cell, but only if the temporal derivative of the cyclic AMP concentration is positive and large [127], cf. Fig. 2.1a. In another example, an action potential propagates in a bacterial film only if a concentration of a previously secreted extracellular potassium has not yet decayed through diffusion [72,

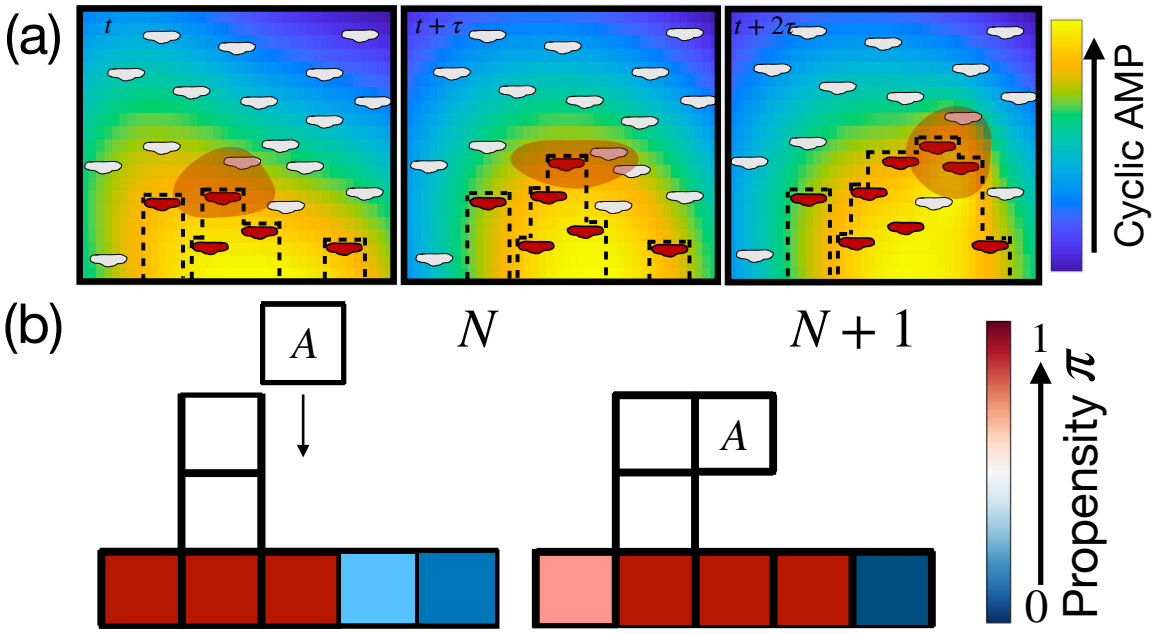

Figure 2.1: (a) Three successive snapshots of the activation front (dashed lines) of Dictyostelium cells. An inactive cell (white) activates (turns to red) in response to cAMP (concentration shown by the color, from blue to yellow). However, to activate, it must sense a large positive temporal change of cAMP, which exists transiently only in the pink oval-shaped region. As cAMP diffuses, only cells that are next to recently active cells can be activated. (b) A ballistic deposition with memory (BDM) model of the process. A particle A is deposited (activated) at time step N into the middle column. The propensity for deposition (shown by color) in the column and its neighbors becomes one at the next time step, while the propensity of all other sites decreases.

93]. All such processes possess memory: the interface at a certain point can grow, but only if it grew here recently. Theory of such interface growth processes with memory is not yet established. In particular, we do not know the relevant critical exponents, how many different universality classes there are, and whether the Family-Viscek scaling is satisfied in such settings.

Here we develop a model of Ballistic Deposition with Memory (BDM), one of likely many possible extensions of the traditional memoryless surface growth processes, which is inspired by the microbiological systems mentioned above. We derive the critical exponents, and verify them numerically. We show that the process falls into a new universality class with a new scaling law and a new scaling relation. The KPZ universality class is an unstable fixed point in the BDM dynamics. Finally, we discuss the effect of varying mem-

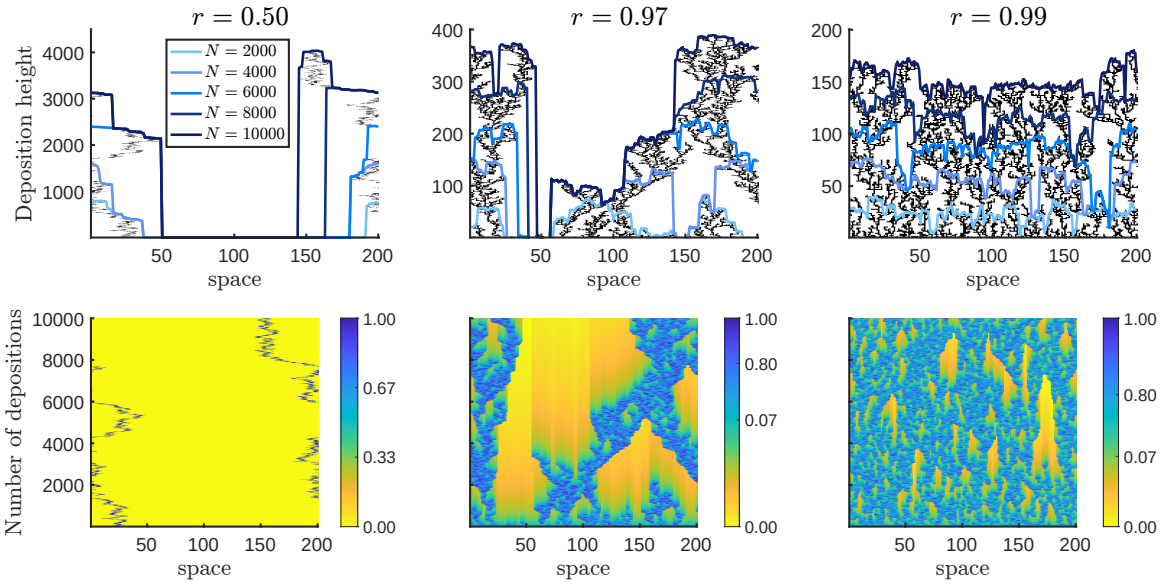

Figure 2.2: Examples of the BDM interfaces. In the top row, the deposition height vs. space coordinate is shown for each deposition as a black dot. Different columns are for different values of r. The blue curves show the surface height as a function of the spatial coordinate for different number of depositions, from $N = 2000$ (light blue), to $N = 10000$ (dark blue). The bottom row shows the propensity, encoded by color, as a function of space and time.

ory duration, and show that the standard KPZ interfaces are achieved in a particular limit of the memory parameters.

2.3 Model formulation

We consider the deposition of particles on a one-dimensional substrate of length L. Each site i has a propensity value $0 \leq \pi_i(N) \leq 1$, which determines the probability that the site will receive a particle deposition at step N. Initially, all sites are equally likely to receive a deposition; i.e., $\pi_i(0) = 1$ for all i. However, unlike in the ballistic deposition model, if a site j receives a deposition, then the propensity at that site and its nearest neighbors is set to one, while the propensity of all other sites is reduced by a factor r, thus reducing the probability of receiving a deposition if no deposition has happened for a long time:

$$\pi_i(N+1) = \begin{cases} 1 & j-1 \leq i \leq j+1 \\ r\pi_i(N) & \text{otherwise.} \end{cases} \tag{2.4}$$

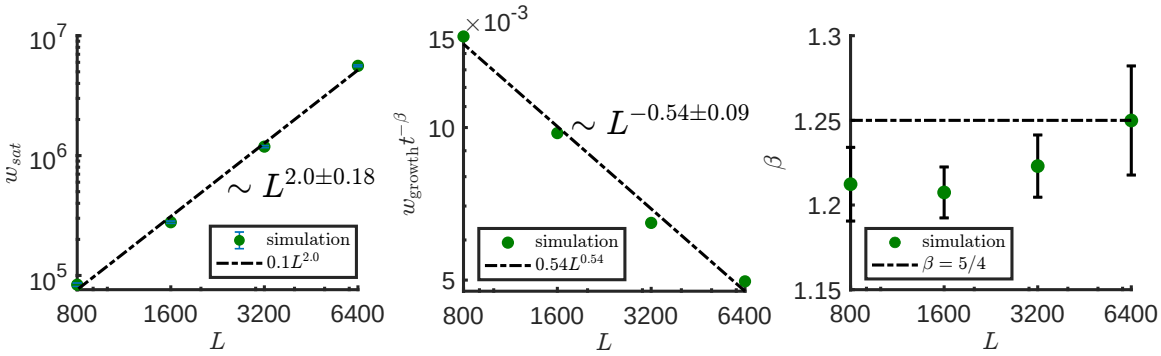

Figure 2.3: (Left) The saturation width of the BDM interface as a function of the system length L. The scaling L^{α}, where α is the roughening exponent, obtained from simulation, is shown. (Middle) The scaling of w/t^{β} in the growth regime as function of the system size is L^{γ}; $t(N)$ is defined as in Eq. (2.11). The fitted scaling with the size exponent $\gamma = 0.54 \approx 0.5$ is also plotted. (Right) The growth exponent β as a function of the system size L. The black dash-dotted line at $\beta = 5/4$ is the theoretical prediction, in agreement with simulations once finite size effects become negligible. The value $r = 0.5$ is used for all subplots.

Overall, the probability to receive the deposition at site i at the Nth deposition event is

$$\mathcal{P}[i, N] = \frac{\pi_i(N)}{\sum_{i=1}^{L} \pi_i(N)}. \tag{2.5}$$

At step N, the height of the interface at site i is $h(i, N)$, with $h(i, 0) = 0$. After a site i is randomly selected for the deposition according to Eq. (2.5), its height increases from $h(i, N)$ to

$$h(i, N + 1) = \max\{h(i - 1, N), h(i, N) + 1, h(i + 1, N)\}, \tag{2.6}$$

allowing for overhangs, as in the traditional Ballistic Deposition, cf. Fig. 2.1a. We model the process with periodic boundary conditions, $h(L + 1, N) = h(1, N)$. The dynamics of the surface and the propensity are shown in Fig. 2.2 for various values of r. For low values of r, a single propensity finger moves randomly, causing the deposition sites to follow a random walk, and overhangs form every time the random walk reverts. In the intermediate r regime, multiple propensity fingers move randomly, merge and split. These fingers deposit particles that form shapes, reminiscent of the Diffusion Limited Aggregation [131], though whether the similarity is more than qualitative is unclear. For values of $r \approx 1$, many deposition fingers merge into a deposition front, whose fluctuation is KPZ-like (as we will

discuss later), but with chasms that have a much lower height, and whose frequency gets lower as $r \to 1$.

2.4 A random walker

We start with the fast propensity decay limit, defined as $r \ll 1/L$ (see *Appendix* for details). Here the probability that any site j receives a deposition at step $N+1$ given that a non-neighboring site i received a deposition at step N is

$$\sum_{\substack{j \notin \{i-1,i,i+1\} \\ 1 \le j \le L}} \mathcal{P}[j, N+1 | i, N] = \sum_{\substack{j \notin \{i-1,i,i+1\} \\ 1 \le j \le L}} \frac{\pi_j(N)}{\sum_{k=1}^{L} \pi_k(N)} < \frac{Lr}{3} \ll \frac{1}{3}, \tag{2.7}$$

where we used the bounds

$$\sum_{k=1}^{L} \pi_i(N) > \sum_{k \in \{i-1,i,i+1\}} \pi_k(N) = 3 \tag{2.8}$$

and

$$\sum_{\substack{j \notin \{i-1,i,i+1\} \\ 1 \le j \le L}} \pi_j(N) < (L-3)r < Lr \tag{2.9}$$

to bound the denominator and numerator, respectively. Therefore, the location x_N of the deposition after N steps is well approximated by a 1-d random walker.

2.5 Determining the unit of time

Classical models of interface growth, such as ballistic deposition or KPZ, usually define time in the units of the mean number of deposited layers, $t \sim N/L$. However, when the measure of where particles can be deposited concentrates, different relations between N, L, and t emerge [35, 73]. Similarly, in our case, the usual definition would imply that, for small r, in one time unit, one deposits L particles on $\sqrt{L}$ sites. The flux per unit time at those $\sqrt{L}$ sites would then be $\sim \sqrt{L}$, which is infinite in the thermodynamic limit, $L \to \infty$. Here we show that, to avoid this pathology, $t \propto N$ is the only acceptable choice.

We begin by defining the space-averaged propensity $\pi(N) = \frac{1}{L}\sum_{i=1}^{L}\pi_i(N)$, which can be decomposed into two different contributions. At step $N-1$, $L-3$ sites do not receive a deposition and are not neighbors of the deposition site. The total propensity of those sites at step N is $(L-3)r\pi(N-1)$. The total propensity of the site that receives a deposition and its two neighbors is three. Thus we obtain the recursion relation for $\pi(N)$

$$\pi(N) = \frac{1}{L}[(L-3)r\pi(N-1) + 3]. \tag{2.10}$$

The solution of Eq. (2.10) (see *Appendix*) has a characteristic time scale of $N_{1/e} = \frac{1}{\ln\frac{L}{(L-3)r}} \approx \frac{1}{\ln\frac{1}{r}}$ depositions, approaching a steady state space-averaged propensity $\pi_* = 3(L-(L-3)r)^{-1}$. In the limit $N \gg N_{1/e}$, there are $\mathcal{O}(L\pi_*)$ sites whose probability of receiving a deposition is $\mathcal{O}(1/L\pi_*)$ (see *Appendix*), while all other sites have probability zero of receiving a deposition. This implies that the effective lattice length is $L\pi_*$ and motivates a definition of time as

$$t = \eta\frac{N}{L\pi_*} = \eta\left(\frac{1}{3} - \frac{L-3}{3L}r\right)N \tag{2.11}$$

for any constant η. In all figures and equations, we define time as in Eq. (2.11) with $\eta = 1$.

2.6 Dynamical exponents

After N depositions, the deposited particles span $\mathcal{O}(\sqrt{N})$ lattice sites and the random walker has performed $\mathcal{O}(N)$ reversals, with each reversal increasing the height by 1. Thus the average height of the interface is $\langle h(N)\rangle \sim c_1\cdot 0\cdot\frac{L-\sqrt{N}}{L} + c_2\cdot N\cdot\frac{\sqrt{N}}{L} \sim \frac{N^{3/2}}{L} = \lambda_{r,L}^{3/2}t^{3/2}/L$ while the mean squared height is $\langle h(N)^2\rangle \sim c_3\cdot 0^2\cdot\frac{L-\sqrt{N}}{L} + c_4\cdot N^2\cdot\frac{\sqrt{N}}{L} \sim \frac{N^{5/2}}{L} = \lambda_{r,L}^{5/2}t^{5/2}/L$ for $\lambda_{r,L} = 1/(L\pi^*)$ and some constants $c_1,\cdots,c_4$. The resulting mean width of the interface becomes

$$w(L,t) \sim \left(\lambda_{r,L}^{5/2}\frac{t^{5/2}}{L} - \delta\lambda_{r,L}^{3/4}\frac{t^{3/4}}{L^2}\right)^{1/2} \tag{2.12}$$

for some constant δ.

In the regime where $\sqrt{N} \ll L$ (i.e., the random walker has yet to span the system), the width of the interface grows with time (the *growth* regime), and its value is dominated by

the first term in Eq. (2.12):

$$w(L,t) \sim \lambda_{r,L}^{5/4} t^{5/4}/\sqrt{L} \propto t^{\beta} L^{-\gamma}. \tag{2.13}$$

This determines the size exponent $\gamma = 1/2$ and growth exponent $\beta = 5/4$, which are in excellent agreement with simulation values, cf. Fig. 2.3, despite finite size effects, and even with r outside the regime $r \ll 1/L$.

In the regime where $\sqrt{N} \gtrsim \mathcal{O}(L)$, the random walker has spanned the lattice, and the surface roughness saturates at

$$w_{\rm sat} \sim N^{5/4}/\sqrt{L} \sim (L^2)^{5/4}/\sqrt{L} = L^2, \tag{2.14}$$

This determines the roughness exponent $\alpha = 2$, which agrees with the simulations, Fig. 2.3.

2.7 Dynamical scaling relation and the scaling law

Since the growth and saturation regimes cross at some time t_c, it follows from Eqs. (2.13) and (2.14) that $\lambda_{r,L}^{-\beta} L^{-\gamma} t_c^{\beta} \sim L^{\alpha}$. The relation $t_c \sim \lambda_{r,L} L^{\frac{\alpha+\gamma}{\beta}} = \lambda_{r,L} L^z$ determines the scaling law $z = \frac{\alpha+\gamma}{\beta} = 2$ in our model. Note that $\lambda_{r,L}$ has a weak dependence on L such that for $L \gg 1$ it is essentially independent of L. The scaling relation becomes

$$w(L,t) \sim L^{\alpha} f(t/L^{\frac{\alpha+\gamma}{\beta}}) \tag{2.15}$$

with f defined as in the Family-Vicsek scaling[38]. Indeed, plotting w/L^{α} against $t/L^{\frac{\alpha+\gamma}{\beta}}$, as in Fig. 2.4, collapses the width vs. time curves plotted for various lattice lengths L in the inset of Fig. 2.4. From this, we conclude that there are three independent exponents α, β and γ that fix the dynamic exponent z.

As a newly arriving particle sticks to the surface following Eq. (2.6), its height is either the same or larger than that of its neighbors. This introduces correlations between neighboring sites. The ensuing height fluctuations spread laterally since particles deposited at nearby sites must have an equal or larger height. This correlation length $\xi_{||}$ can only grow up to the substrate length, i.e., $\xi_{||} \sim L$ for $t \gg t_c$. Replacing L by $\xi_{||}$ in $t_c \sim \lambda_{r,L} L^z$, we

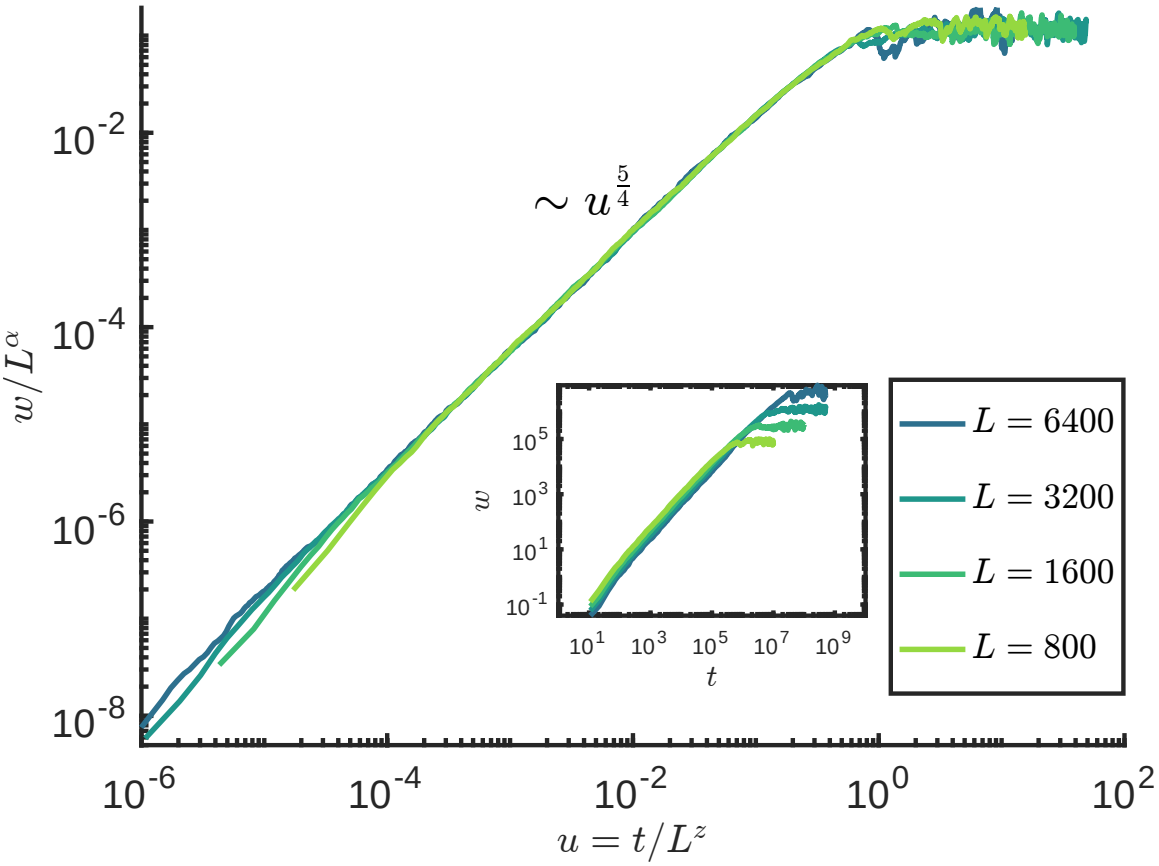

Figure 2.4: Interface width w as a function of time for $r = 0.5$ and systems of different lengths, L. The time axis is scaled by $L^z = L^2$, and the width axis is scaled by $L^\alpha = L^2$, which achieves collapse of all curves. The data come from averaging over 20 independent runs for all L except $L = 6400$, which used 5 independent runs. The inset shows the bare width as a function of the bare time.

find that $\xi_{||} \sim \lambda_{r,L}^{-1/z} t_c^{1/z}$ for $t \gg t_c$. Since $\xi_{||} \sim N^{1/2}$ for $t \ll t_c$, we see that $\xi_{||} \sim \lambda_{r,L}^{-1/z} t^{1/z}$ holds for $t \ll t_c$ as well.

2.8 Varying the memory time scale

As r increases, so does the total size of the randomly moving propensity fingers $L\pi_* = \lambda_{r,L}^{-1}$. Increasing r also decreases linearly the time to saturation t_c, cf. Fig. 2.5. In the limit of $r = 1$, the KPZ exponents as seen in Fig. 2.5 and the standard definition of time $t = N/L$ are recovered. For most of the $r \in [0, 1]$ domain, the surface fluctuations are in the new universality class and are not in the KPZ class, cf. Fig. 2.5. At early times, the Poisson

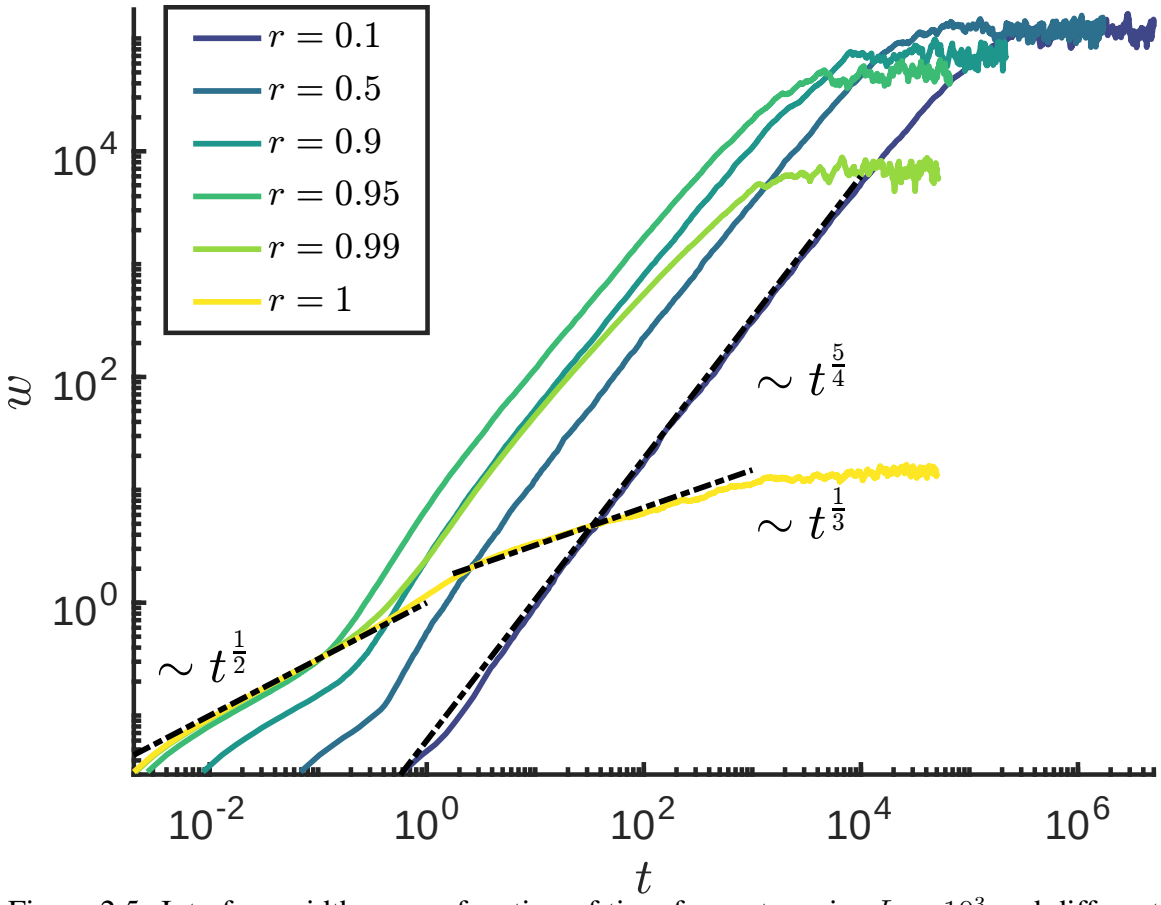

Figure 2.5: Interface width w as a function of time for system size $L = 10^3$ and different propensity decay constants r. The growth exponent changes from the KPZ value of $\beta = 1/3$ to BDM value $\beta = 5/4$ as the propensity decay rate r deviates from $r = 1$. For $r = 1$ and for very small time, the width exhibits the Poisson scaling with $\beta = 1/2$.

regime dominates the growth with a characteristic scale $t^{1/2}$ followed by the KPZ growth with a scale $t^{1/3}$ (effectively, $r \sim 1$) within a moving finger of finite width, and eventual transition to fluctuations with a scale of $t^{5/4}$ (effectively, $r < 1$). For finite L, this transition occurs at $r^*(L) < 1$. However, in the thermodynamics limit $L \to \infty$, the transition value $r^*(L) \to 1$ because the random depositions would cover only a finite part of the lattice. This implies that the KPZ class is an unstable point of the dynamics that occurs only at $r = 1$ if L is infinite.

2.9 Discussion

We numerically and analytically studied a model of surface growth with memory. Introduction of memory breaks the temporal locality of the deposition process, so that deposition at the current time is dependent on the history of prior deposition events. This leads to emergence of a new size exponent γ, which captures the random walk nature of the deposition process at long times, and to changes in the values of the growth exponent ($\beta = 5/4$) and the roughening exponent ($\alpha = 2$). These exponent result in a new scaling law $z = \frac{\alpha+\gamma}{\beta} = 2$, which generalized the classical scaling law $z = \alpha/\beta$. In other words, BDM is a surface growth process that does not belong to the KPZ universality class.

In the standard KPZ and EW universality classes, there is no correlation in the deposition process, so that the position of the next deposition is independent from the positions of all past depositions. By including the propensity, we introduce such correlations. The ensuing universality class differs from the KPZ and the EW classes, and from their various well-known modifications [28, 61, 74, 94, 117, 124, 132]. This is because our novel class changes the dimensionality of the dynamics by introducing additional dynamical variables π in addition to h.

While our model was inspired by biological systems, it remains to be seen if the discovered universality class is relevant to them. In order to verify this, it might be easier to introduce and study a similar memory-enabled extension of the Eden growth model [34, 38]. Further, it is then necessary to explore empirically large spatiotemporal scales that are beyond the typical scales probed in current experiments to test our predicted exponents. We hope that such experiments will provide exciting new insights in interface growth phenomena.

2.10 Appendix

2.10.1 Solving the propensity recursion relation

Multiplying the recursion relation

$$\pi(n+1) = \frac{1}{L}[(L-3)r\pi(n) + 3]$$

by the summing factor $\left(\frac{L-3}{L}r\right)^{-n-1}$ and summing from $n = 0$ to $n = N-1$, we obtain the telescoping sum

$$\sum_{n=0}^{N-1}\left[\left(\frac{L-3}{L}r\right)^{-n-1}\pi(n+1) - \left(\frac{L-3}{L}r\right)^{-n}\pi(n)\right] = \frac{3}{L}\sum_{n=0}^{N-1}\left(\frac{L-3}{L}r\right)^{-n-1}$$

$$\left(\frac{L-3}{L}r\right)^{-N}\pi(N) - \pi(0) = \frac{3}{L}\sum_{n=0}^{N-1}\left(\frac{L-3}{L}r\right)^{-n-1} \quad (2.16)$$

or equivalently,

$$\pi(N) = \left(\frac{L-3}{L}r\right)^{N}\left(\pi(1) + \frac{3}{L}\sum_{n=0}^{N-1}\left(\frac{L-3}{L}r\right)^{-n}\right).$$

This evaluates to

$$\pi(N) = \pi(0)\left(\frac{L-3}{L}r\right)^{N} + \frac{3}{L}\sum_{n=0}^{N-1}\left(\frac{L-3}{L}r\right)^{N-n-1}. \quad (2.17)$$

Notice that the last sum can be rewritten as

$$\frac{3}{L}\sum_{n=0}^{N-1}\left(\frac{L-3}{L}r\right)^{N-n-1} = \frac{3}{L}\sum_{n=0}^{N-1}\left(\frac{L-3}{L}r\right)^{n}. \quad (2.18)$$

Rewriting the solution of the recursion relation in the exponential form yields

$$\pi(N) \quad = \quad \left(\frac{L}{(L-3)r} - \frac{3}{L-(L-3)r}\right)e^{-N\ln\left(\frac{L}{(L-3)r}\right)} \quad + \quad \frac{3}{L-(L-3)r}. \quad (2.19)$$

From the first term, we obtain the propensity decay time scale

$$N_{1/e} = \frac{1}{\ln\left(\frac{L}{(L-3)r}\right)} \approx \frac{1}{\ln\left(\frac{1}{r}\right)}, \text{ for } L \gg 1. \quad (2.20)$$

Thus $\pi(N)$ exponentially decays to

$$\pi_* = \frac{3}{L-(L-3)r} \quad (2.21)$$

on the time scale $\mathcal{O}(N_{1/e})$. In the limit $r \to 1$ and $L \gg 1$, the time scale $N_{1/e} \to \infty$. That is, the propensity remains at the fixed value of $\pi \approx 1$ and no decay occurs, consistent with the regular Ballistic Deposition process.

2.10.2 Computing deposition probabilities

From Eq. (2.5), we deduce that if site j receives a deposition at step N, then the probability that site i will receive a deposition is

$$\mathcal{P}[i, N+1] = \frac{\pi_i(N+1)}{\sum_{i=1}^{L} \pi_i(N+1)} = \frac{\pi_i(N+1)}{L\pi(N+1)} \tag{2.22}$$

$$\approx \begin{cases} \frac{1}{(L-3)r\pi(N)+3}, & \text{for } j-1 \le i \le j+1 \\ \frac{r\pi(N)}{(L-3)r\pi(N)+3}, & \text{otherwise.} \end{cases} \tag{2.23}$$

However, for $N \gg \frac{1}{\ln(1/r)}$, the propensity $\pi(N) \approx \pi_* = \frac{3}{L-(L-3)r}$ and the probability of a deposition at site i becomes

$$\mathcal{P}[i, N+1] \approx \begin{cases} \frac{1}{3} - \frac{L-3}{3L}r, & \text{for } j-1 \le i \le j+1 \\ \frac{r}{L}, & \text{otherwise.} \end{cases} \tag{2.24}$$

As a check, we see that the probability that any site receives a deposition is one

$$\sum_{i=1}^{L} \mathcal{P}[i, N+1] = (L-3)\frac{r}{L} + 3\left(\frac{1}{3} - \frac{L-3}{3L}r\right) = 1.$$

Furthermore, using Eq. (2.24), we see that

$$\sum_{\substack{j \notin \{i-1, i, i+1\} \\ 1 \le j \le L}} \mathcal{P}[j, N+1 | i, N] = \frac{L-3}{L}r \approx r \ll 1 \tag{2.25}$$

for $r \ll 1$. Therefore, the position x_N of the deposition at step N follows an unbiased random walk in this regime.

2.10.3 Expanding the random walk regime

The steady state propensity π_* in practice is not spread out over the entire lattice. Instead $L\pi_* = 3(1 - \frac{L-3}{L}r)^{-1}$ sites have propensity almost one, and the rest of the lattice has zero propensity. This motivates the definition of time as $t = \eta\frac{N}{L\pi_*}$ shown in Eq. (2.11) in

the main text. In the limit $L\pi_* \ll L$, the propensity process is made of multiple fingers of cumulative size $L\pi_*$, which all are performing random walks. These fingers dynamically merge and split as particles are deposited randomly. This has the effect that, outside these fingers, the probability of a deposition is zero and hence our dynamical exponents will hold in the significantly larger regime $L\pi_* \ll L$ or, equivalently, $\frac{3}{1-r} \ll L$.

2.10.4 Extracting exponents from data

To estimate the growth exponent β, we compute the mean width $\langle w(N)\rangle = \frac{1}{n}\sum_{i=1}^{n} w_i(N)$ of n realizations of width $\{w_1(N), \cdots, w_n(N)\}$ obtained from n realizations of the height $\{h_1(N), \cdots, h_n(N)\}$ according to Eq. (2.1). To ensure that the estimated value of the growth exponent remains in the growth regime and is unaffected by cross-over effects, we limit the range of time used in the estimation to $N \in [3\cdot 10^2, 3\cdot 10^4]$. According to Eqs. 2.13 and 2.14, the mean width in the growth regime is $\langle w(N)\rangle \sim L^{-\gamma}N^{\beta}$. Therefore, linear regression obtains the slopes β and γ of the plane $\ln w(N) = \beta \ln N - \gamma \ln L + k$ when regressed against $\ln N$ and $\ln L$ respectively. The slopes obtain values of $\beta \approx 1.25 \pm 0.03$ and $\gamma \approx 0.54 \pm 0.09$.

To determine the size of the fluctuations around the estimated value of β for a fixed substrate length L, we use the covariance matrix $\Sigma_{\beta,\Delta}$ of the parameters β and $\Delta = -\gamma \ln L + k$. The covariance matrix is $\Sigma_{\beta,\Delta} \approx R^{-1}(R^{-1})^T \frac{|\vec{e}_{res}|^2}{df}$ where R is the triangular factor from a QR decomposition of the Vandermonde matrix of $\ln(N)$, $\vec{e}_{res}$ is the vector of residuals between the data and the fitting line, and $df = 2$ is the number of degrees of freedom. The quantity $(\Sigma_{\beta,\Delta})^{1/2}_{\beta,\beta}$ provides the standard deviation on β. A similar procedure is followed when we regress on $-\gamma \ln L + k$ against $\ln L$ to find γ and its standard deviation.

To determine the roughening exponent α, we note that $\langle w(t)\rangle \sim L^{\alpha}$ in the saturated regime. To avoid a bias in the estimate of α due to the transition from growth to saturation, we limit the time range used in the estimate to $N \in I_L = [2L^2, 10L^2]$ for lattice length L.

For a fixed lattice length L, the width in the saturation regime fluctuates over the interval I_L. In this regime, the mean width is obtained from the relation $\langle w_{\rm sat}\rangle_{I_L} = \frac{1}{|I_L|}\sum_{N\in I_L}\langle w(N)\rangle$, where $|I_L|$ is the length of the interval I_L used for the estimate of w_{sat} of the lattice of length L. Using linear regression, the value of the slope of the line $\ln\langle w_{\rm sat}\rangle = \alpha \ln L + \lambda$ gives $\alpha \approx 2.0 \pm 0.18$ as seen in Fig. 2.3, in agreement with analytical results.

Chapter 3 *C. elegans* thermotaxis reveals general mechanisms of extinction and recovery in animal learning

3.1 Author contributions:

This chapter presents the paper [102], whose authors are Ahmed Roman, Konstantine Palanski, Ilya Nemenman, and William S Ryu. William S Ryu, and Ilya Nemenman conceived and designed the project. Konstantine Palanski performed all of the experiments; Ahmed Roman and Konstantine Palanski performed the data analysis; Ahmed Roman and Ilya Nemenman developed and fitted the model. Ahmed Roman, Ilya Nemenman, and William S Ryu wrote the manuscript. William S Ryu and Ilya Nemenman supervised the study.

3.2 Summary

C. elegans is capable of learning and remembering behaviourally relevant cues such as smells, tastes and temperature [7]. This is an example of associative learning, a process where behaviour is modified by making associations between various stimuli [90]. Since the mathematical theory of conditioning [84, 97] does not account for some of its salient aspects, such as spontaneous recovery of extinguished associations, accurate modelling of behaviour of real animals during conditioning has turned out difficult. Here we do this in the context of the dynamics of the thermal preference of *C. elegans*. We quantify *C. elegans* thermotaxis in response to various conditioning temperatures, starvation durations, and genetic perturbations using a high-resolution microfluidic droplet assay. We model this data comprehensively, within a new, biologically interpretable, multi-modal framework. We discover that the strength of the thermal preference is composed of two independent,

genetically separable, contributions and requires a model with at least four dynamical variables. One pathway positively associates the experienced temperature independently of food and the other negatively associates to the temperature when food is absent. The multi-dimensional structure of the association strength provides an explanation for a number of longstanding questions in animal learning [31, 33, 84, 90], including spontaneous recovery, asymmetric response to appetitive vs. aversive cues, latent inhibition, and generalization among similar cues.

3.3 Introduction

Animals modify their behaviour through learning. Conditioning or associative learning, that is, making associations between various stimuli is one of the best studied such process [32, 60, 84, 90, 126]. Nonetheless models of conditioning, quantitatively accounting for intricacies of behaviour are rare [36, 87, 88, 139].

In conditioning, a reward or punishment from an unconditionally appetitive or noxious stimulus (US, or reinforcement) teaches the animal to associate it with conditioned stimuli (CS, or cue), which predict the US. Conditioning is usually modeled within the Rescorla-Wagner (RW) [97] and the reinforcement learning (RL) frameworks [30, 31, 84, 126], where an error an animal makes in predicting a US from CS cues changes the strength of the CS-US associations, decreasing future errors. If the conditioning process is slow, so that many predictive cue-reward pairs happen, and the CSs appear one at a time, then the RW model becomes (see Section 3.6.1 for a derivation):

$$\tau \frac{dV_i}{dt} = f_i(t) \left(\lambda_0 f_{\lambda|i}(t) - V_i(t) \right) . \tag{3.1}$$

Here V_i is the association strength between the ith CS and the US in some arbitrary units set by λ_0, the salience of the US. Further, τ is the time scale of learning, and f_i and $f_{\lambda|i}$, both in the range of $[0, 1]$, are the frequency of the ith CS and the US conditional on it, respectively. The frequencies may depend on the animal's behaviour, which, in turn, is

controlled by the association strengths $\{V_i\}$.

Such conditioning models are simplified and often lack the complexity to serve as *quantitative* models of real life animal behaviour [84]. Even worse, they fail to account for some *qualitative* features of conditioning, including (i) extinguishing an association (not merely unlearning) and subsequent frequent spontaneous recovery [31, 33, 84], (ii) existence of multiple reinforcement systems (some potentially associated with habitual and not reward-driven actions), outputs of which integrate into behaviour [29, 84], (iii) asymmetric responses to appetitive vs. aversive cues [31] and to conditioned association vs. conditioned inhibition (in the latter, a CS predicts absence of the US) [84], (iv) generalization among similar, but distinct cues [84], and so on. The weakness of models is often due to the difficulty in designing informative experiments: measuring behaviour with high precision, quickly, and for long duration is nontrivial; behaviour itself may modify the conditioning contingency; behaviour is noisy, often discrete, and hence is not a reliable readout of a CS-US association; and biology of different reinforcement pathways and mechanisms of their integration are unclear.

To generate more useful data, and to incorporate them into more accurate models of an animal's behaviour during learning, we turn to *C. elegans* as a *quantitative* model system. The worm—one of the simplest organisms exhibiting conditioning [7]—associates the presence of food with various environmental signals, such as salts [105], odors [138] and temperature [45, 78], modifying complex, but measurable behaviours (chemo- and thermotaxis [2, 12, 45, 128]). It is a particularly good model system for studying associative learning because of its simple nervous system [130], short life cycle, and our ability to control the environment and accurately measure its behaviour over long periods of time (hours). In addition, a number of genes affecting learning in *C. elegans* have been identified [50], most notably those in the insulin-like signaling pathway [114].

C. elegans thermotaxis is a well-studied behaviour affected by conditioning: when placed on a thermal gradient, *C. elegans* taxes to their cultivation temperature [45, 78,

104] and, when near this thermal preference, it performs isothermal tracking [45, 78, 104]. A new thermal preference is acquired when worms are placed at a different cultivation temperature [1] [12, 45, 78]. The thermal preference and its temporal dynamics can be measured by tracking worm behaviour on thermal gradients [21, 23, 96, 104]. The thermal preference is established asymmetrically between high and low (above or below $\sim 20^{\circ}$C) temperature conditions [104, 137]. The preferred temperature may depend on the starvation state of the worm [45] and on the steepness of the temperature gradient [54, 96]. Further, the dynamics of the preference is relatively fast (from tens of minutes to a few hours) [12, 96]. This makes designing experiments to assay the dynamics of associations difficult since the preferences can change faster than the worms move in gradients to reveal them. As a result, even some of the most basic questions about *C. elegans* associative learning remain unclear, such as the relative importance of food to the establishment of the preference, precluding accurate mathematical models of the process.

Here we designed a microfluidic assay to monitor the thermotactic preference of individual worms, with the precision and the temporal resolution sufficient to track its dynamics. We showed that the simple model, Eq. (3.1), cannot precisely fit such data. Thus we developed, and then experimentally verified, a more complicated picture of the dynamics of conditional associations in *C. elegans*. We identified multiple pathways affecting thermotaxis: habituation to the experienced temperature, and avoidance of temperature when no food was collected. Using worms with mutations in the insulin signaling pathway—whose behaviour the model predicts with quantitative accuracy—we isolated contributions of these independent pathways to the behaviour. We argue that the developed model solves a variety of long-standing conceptual problems in the field of animal associative learning. We further suggest that such multi-pathway organization of the conditional response dynamics may be optimal for food search in dynamical environments.

[1] cultivation temperature here means the temperature at which the worms were raised.

3.4 Results

3.4.1 Measuring thermal preference

We rear N2 wild type worms at 15°C and 25°C, hereafter called cold and warm worms (cf. Section 3.6.2). Individual worms are then placed in each droplet (4 mm dia.) on a microscope slide with a six droplet array Fig. 3.1 (cf. Section 3.6.3). The slide is sealed, and the swimming patterns of each worm are quantified for up to 4 hours. We first acclimatize the worms for 15 min in the droplet at 20°, and then turn on the thermal gradient of 1°C/cm (from 19.8° to 20.2° across the droplet width). Since the droplet is small and can be traversed by a worm in seconds, the thermotactic bias in the worm's position can be measured in less than a minute, allowing the quantification of its dynamics. Each worm's position along the gradient in the droplet is rescaled to take values between -1 (cold edge) and +1 (warm edge), and the thermotactic index, Θ, is determined as the mean position over some period of time (usually tens of seconds, cf. Section 3.6.4). The average Θ for about 110 worms reared at 15°C and 25°C are shown in Fig. 3.1a (blue and red, respectively). The cold (warm) animals initially show strong preference to the cold (warm) side, $\Theta \approx \pm 0.25$, respectively. For scale, note that the worms move constantly, and if they were to explore just one side of the droplet uniformly—a very strong bias—the thermotactic index would be $1/3$.

3.4.2 Dynamics of thermal preference

After the thermal gradient is turned on, cold (warm) worms perform thermotaxis to the cold (warm) side of the droplet, presumably expecting to find food there, cf. Fig. 3.1(C) and Fig. 3.2. However, the droplet has no food, and the strength of the bias decreases with time, changing sign in about 3h (2h), Fig. 3.2. The warm worms then return back to zero bias by 3.5h in the droplet, while the cold worms remain warm-biased for the remaining time. This observation of the zero crossing in the thermal preference is crucial: the simple dynamical system, Eq. (3.1), cannot oscillate autonomously (cf. Section 3.6.1). Thus the model in

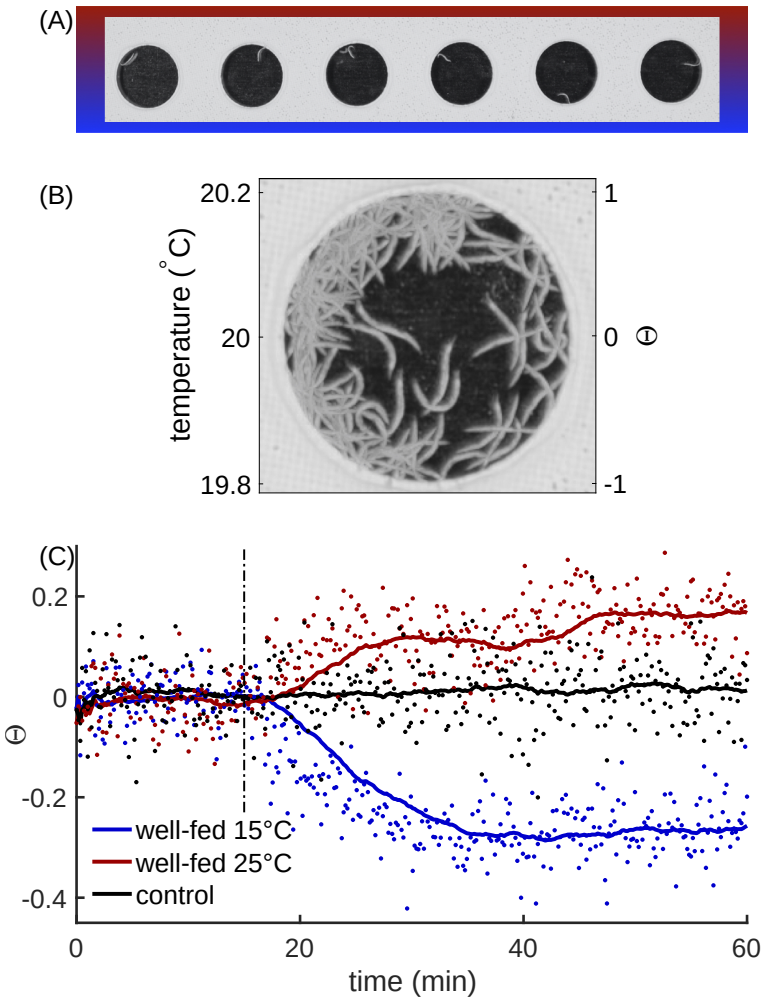

Figure 3.1: **Droplet thermotaxis assay**. (**A**) An array of PTFE constrained droplets (4mm dia.) with single *C. elegans* placed in a thermal gradient. (**B**) Multiple exposures of a single droplet (150 frames captured at 6 frames per minute) superimposed on each other. The thermotaxis index (Θ) is the average position of the worm in the droplet along the thermal gradient. (**C**) Thermotactic response of N2 with a cryophilic and thermophilic preferences (blue and red) and N2 control with no gradient (black). The vertical dashed-dotted line indicates the onset of the temperature gradient in the droplet for the biased worms.

Eq. (3.1) is incomplete. Specifically, to model the observed thermotactic index dynamics, we need at least two interacting dynamical variables in a model, with two distinct time scales.

The dynamics in Fig. 3.2 proceed autonomously, with no food in the droplet, eventually developing *avoidance* of an initially attractive temperature. A possible explanation is a faster decay of the association on the initially attractive side than on the opposite side.

However, the two sides are less than a half a degree apart, precluding large differences in time scales. We thus explain the avoidance by assuming that at least one of the dynamical variables describing the thermotactic index is reinforced by the *absence* of food, encoding avoidance of (rather than preference to) certain temperatures.

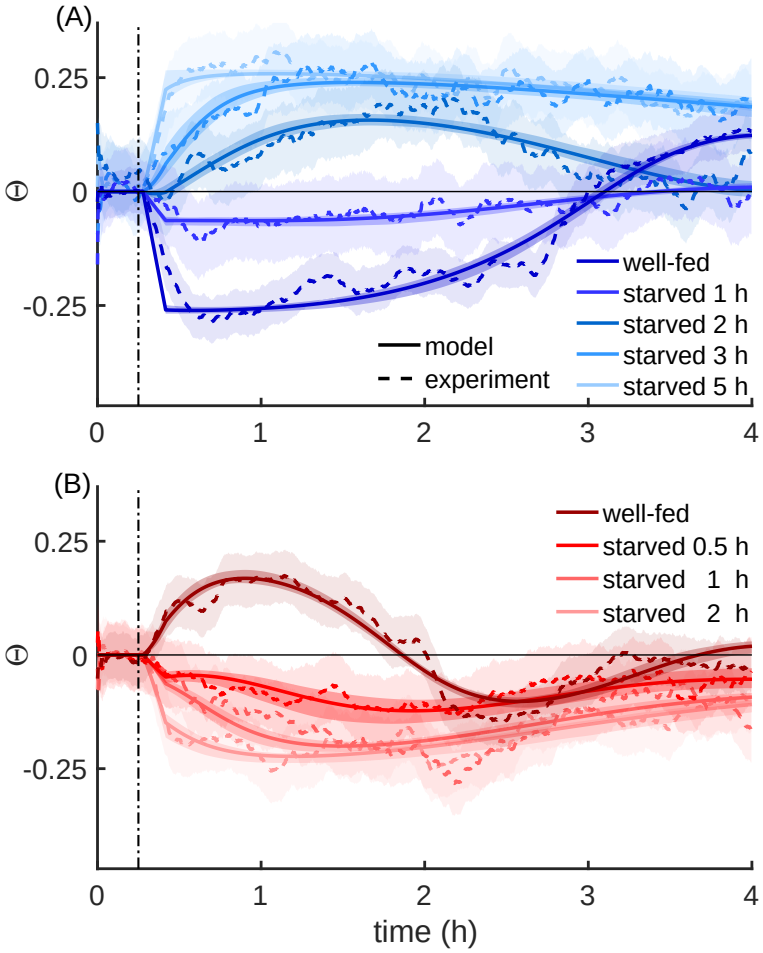

Figure 3.2: **Dynamics of thermal memory for wild type worms.** Thermotactic index of cold (**A**) and warm (**B**) worms with different duration of starvation is shown by different shades of blue (red). The vertical dashed-dotted line indicates the onset of the temperature gradient. Immediate preference of the colder (warmer) side by the non-starved cold (warm) reverts to the preference of the warmer (colder) side as the worm spends more time in the droplet. The initial preference also weakens and reverses to avoidance when the worm is subjected to long duration of starvation before the droplet assay. Data is shown as dashed lines, and model fits are depicted with solid lines. All curves with different pre-assay starvation durations for the same rearing temperature share the same fitting parameters. For presentation purposes only (but not during fitting), both data and model are filtered (cf. Section 3.6.4). Error bands on the data are the 16.5 to 83.5 percentiles. Error bands for the models are obtained using bootstrapping (cf. Section 3.6.13).

To further understand the nature of the dynamical variables involved in the thermotactic preferences, we focus on the *ins-1(nr2091)* strain with a mutation of an insulin-like peptide [91]. This mutant has a more persistent thermal preference [58] and shows a defective negative association in odor and salt learning [65, 121], but presents normal starvation behaviour [58]. We find, cf. Fig. 3.3A, that, while *ins-1* worms reared at either 15- or 25°C initially show the same cryophilic and thermophilic preference as the N2 wild type, the preference is sustained for the duration of the experiment. This persistence makes it possible to interpret these data in the context of models similar to Eq. (3.1), but only if the parts of the thermal memory not affected by the *ins-1* mutation do not decay with time (unlikely for any non-reinforced association) or are reinforced even without food. The latter option suggests that one of the dynamical variables in model of the thermotactic index is likely to be habituation to the current temperature, rather than food-temperature association.

In summary, zero crossing in the N2 data and absence of preference degradation in the *ins-1* mutant collectively suggest that an effective model must include, at least, two dynamical processes: habituation to the current temperature and avoidance of temperature when no food has been observed. Crucially, since *ins-1* worms exhibit no avoidance with no effect on the initial positive association, these two signals must be mediated by distinct biological pathways.

3.4.3 Constructing a model of thermal preference dynamics

To build a mathematical model of the conditioning, we explore the dimensionality of the behavioural dynamics using tools from dynamical systems theory. First, we consider the delayed embedding representation of worm trajectories [119], which bounds the number of dynamical variables by at least 3 (cf. Section 3.6.5). Second, our data sets include worms that were starved after initial rearing and before being assayed, cf. Fig. 3.2. Crucially, for some starvation durations, Θ does not approach its saturating values. In this linear regime,

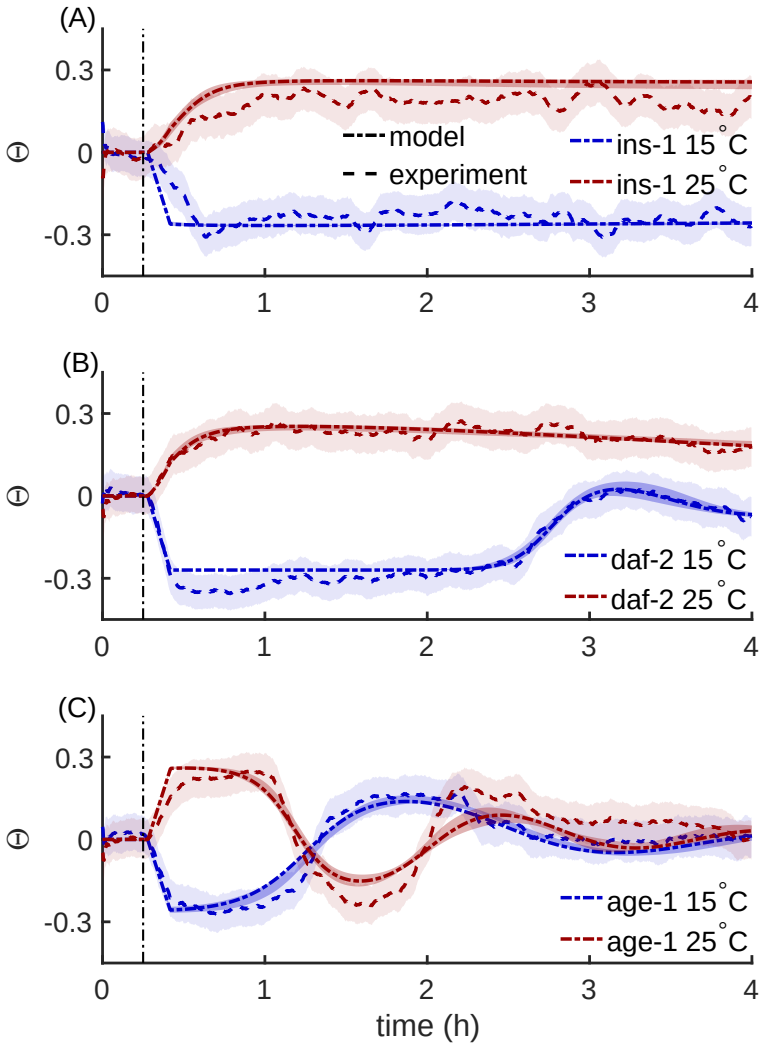

Figure 3.3: **Dynamics of thermal memory for mutants.** Thermotactic index of the *ins*-1 (**A**), *daf*-2 (**B**) and *age*-1 (**C**) mutants, reared at 15°C (blue) and 25°C (red). Worms are well fed before assaying the thermotactic response. The vertical dashed-dotted line indicates the onset of the temperature gradient. Plotting conventions are as in Fig. 3.2.

multiple time scales of oscillations suggest that at least four dynamical variables contribute to the conditioning dynamics (cf. Section 3.6.5).

In other words, we seek to describe the dynamics of Θ as a combination of habituation to the current temperature, its avoidance if not reinforced by food, and two additional variables. To shed light on the latter, we note that the worms are reared and assayed at different temperatures. Thus, for example, initial preference to the cold side at 19.8°C is a *generalization* [45] of the rearing at 15°C. Thus one needs separate dynamical variables—both the avoidance and the habituation—to model the temperature preferences at the rearing and at

the assaying temperatures.

Putting all of this together, we model the thermotactic bias $\Theta(t)$ as a a nonlinear function (here we choose $\tanh$) of the sum of the habituation and the avoidance in the droplet, $h(t)$ and $a(t)$. In turn, these integrate the experienced temperature (independent of and without food, respectively), and relax to zero (no preference or avoidance) with time if no additional reinforcement is present. (Note that this is distinct from the RW model, Eq. (3.1), where the association strength decreases only when a CS is observed, but not reinforced.) Similarly, by $h_{\rm r}(t)$ and $a_{\rm r}(t)$, we represent the habituation and the avoidance of the temperature, at which the worm was reared. The generalization is modeled by letting the rearing and the in-droplet variables to become similar to each other with time. This results in a new, four-variable, generalization of the RW model, Eq. (3.1) (see Section 3.6.6 for more details). While many similar models can be written, we found only one that could quantitatively fit the experimental data:

$$\tau_h \partial_t h = A_h \hat{\Theta}(t) - h + g_h h_{\rm r}, \tag{3.2}$$

$$\tau_a \partial_t a = (1 - F(t)) A_a \hat{\Theta}(t) - a + g_a a_{\rm r}, \tag{3.3}$$

$$\tau_{h,{\rm r}} \partial_t h_{\rm r} = -h_{\rm r}, \tag{3.4}$$

$$\tau_{a,{\rm r}} \partial_t a_{\rm r} = -a_{\rm r}, \tag{3.5}$$

$$\hat{\Theta}(t) = \Theta_0 \tanh(h(t) - a(t) + c). \tag{3.6}$$

Here $\tau_{h/a}$ are the time scales of the relaxation of the in-droplet habituation and avoidance, and $\tau_{h,{\rm r}/a,{\rm r}}$ denote these time scales at the rearing temperature. $A_{h/a}$ are the strength of conditioning of the corresponding internal states. $F(t)$ indicates the presence/absence of food at time t; it is always zero in the droplet by design. $\hat{\Theta}(t)$ is the thermotactic index predicted by the model, which is also the proxy for the warm/cold temperature that the worms experience. $g_{h/a}$ are the strengths, with which the preferences to the rearing temperature generalize to the preferences in the droplet. We set the scale of $g_{h/a}$ by rescaling $h_{\rm r}$ and $a_{\rm n}$, respectively. Then $g_{h/a} = +1$ (or -1) for the warm (cold) worms. Finally, $\Theta_0 < 1$ is the

maximum thermotactic index that the worm can experience due to the droplet geometry, and c is the intrinsic thermal bias (see Section 3.6.6 for a more detailed description)

Now we can develop intuition for how this model can account for the thermotactic dynamics data. For example, the worm reared at 15°C initially is habituated to cold and hence starts with $h_r < 0$, which generalizes to $h < 0$; it is well fed, so that $a = a_r = 0$. The worm then starts with $\hat{\Theta} < 0$. However, since the droplet contains no food, avoidance of the cold accumulates, $a < 0$. At some point $a \approx h$, and $\hat{\Theta} \approx 0$. The worm then experiences both sides of the droplet nearly equally, and no biased reinforcement happens. However, if now $\tau_h < \tau_a$, the preference to the cold degrades faster than its avoidance, and Θ crosses zero and becomes positive. In contrast, if the assay starts following a short starvation period, then $a_r \neq 0$, which generalizes to $a \neq 0$, and the worm's preference to the cold side starts from a smaller value.

3.4.4 Fitting the model to data

The following parameters of the constructed model must be fitted: four time scales, two reinforcement strengths, and initial conditions at the start of the droplet assay for every preference at the reared and the assay temperatures. We fit the same parameters *globally* for all starvation durations at the same rearing temperature. We expect most parameters to be different for different rearing temperature and different mutants since worms develop into substantially different animals under these conditions [42, 135]. However, not all of the parameter values are possible. For example, we expect that the bias decays for long times, and that the initial conditions, $a(t = 0)$, depend monotonically on the starvation duration. We enforce such constraints (cf. Section 3.6.7) into the optimization (cf. Section 3.6.10) to fit the model to the data.

Empirically, while many parameter values are not determined tightly by the data (i.e. have small effects on the quality of fit), some combinations of parameters are more strongly constrained. We use this observation and Bayesian model selection [70] to reduce the

number of model parameters, and hence to increase the generalization and decrease the optimization cost (cf. Section 3.6.11). Most notable reductions include the same initial conditions for avoidance values at similar starvation durations, zero value of the thermotactic bias c, and the same ratio of habituation to avoidance time scales at the rearing and the assay temperatures (see Section 3.6.11 and Tbls. 3.3, 3.4). Figure 3.2 shows that this results in a model with excellent quantitative fits, with χ^2/f of 1.17 and 1.12 including all starvation conditions, for the cold and the warm worms, respectively. The fitted model confidence intervals (cf. Section 3.6.13) fall within the confidence intervals of the data (cf. Section 3.6.4). Crucially, the model also reproduces salient qualitative features of the data, including rapid initial saturation, slower crossover, and—for 25°C—the second reversal to zero bias. Further, the decay and the eventual reversal of the initial Θ as a function of the starvation duration is also modeled well, Fig. 3.2. While most inferred model parameters, Tbls. 3.3 and 3.4 are not illustrative, a few parameter combinations are notable. First, the ratio τ_h/τ_a is reliably small, ≈ 0.28 and ≈ 0.41 for the cold and the warm worm, respectively. That the habituation has faster dynamics than the avoidance is what allows for overshoots and oscillations within our model, and should guide future experiments. Second, the bias c is zero for the cold worm and is negative for the warm worm. This is consistent with the worms avoiding warmer, but not colder temperatures in isothermal tracking [137], and is an independent confirmation of inferences from our model.

3.4.5 Thermotactic dynamics in mutants

To test the generality of our model, we investigated worms with mutations in the insulin-like signaling pathway, which is involved in olfactory, chemotactic, and thermotactic avoidance learning [20, 58, 121]. We chose three mutants that affect insulin-like signaling at different stages (see Section 3.6.2). These show qualitative differences in their thermotactic dynamics compared to the wild-type worm. As discussed earlier, *ins-1(nr2091)* worms learn a thermal preference, and this preference persists for the duration of the experiment

(Fig. 3.3A). *daf-2(e1370)* have a mutation in an insulin-like receptor which produces a temperature-sensitive dauer phenotype [41], disrupting the pathway predominantly at warm temperatures (Fig. 3.3B). Consistent with this, warm *daf-2* worms show a persistent thermal preference like *ins-1* worms, while cold worms have a reduced avoidance response, but with a similar timing. *age-1(hx546)* mutants, with a degraded function of a kinase downstream of the insulin-like peptide receptor, have accelerated rates of negative association in their thermal learning assay [58]. Indeed, these mutants show faster crossover for both rearing temperatures (Fig. 3.3C).

If our model is biologically realistic, we expect the mutants to have have different parameters values in its avoidance branch, Eqs. (3.3, 3.5), but not the habituation branch, Eqs. (3.2, 3.4). Therefore, we fit the mutant data by keeping all but the avoidance parameters to be the exactly those inferred for the N2 worm (cf. Section 3.4.4). We only fit the avoidance parameters for each mutant and each rearing temperature (see Section 3.6.8 and Section 3.6.12). Further, we do not fit *any* parameters for the *ins-1* worm, and simply remove the avoidance branch for this mutant, setting $a(t) = a_r(t) = 0$, consistent with the avoidance effectively removed by the *ins-1* mutation. The agreement between the model and the data for all mutants in Fig. 3.3 is generally within the error bars (cf. Section 3.6.13). This is better than could have been expected, considering that mutant worms are very different from the wildtype and implies a biological relevance rather than purely statistical prowess of the model. The inferred values of the parameters, Tbl. 3.5, are consistent with the expected effects of the mutations. Crucially, the ratio τ_a/τ_h remains large for all datasets.

To further explore the biological relevance, we fit an alternative model where only the habituation (rather than the avoidance) parameters are fitted for mutants, contradicting the known biology. Crucially, this alternative model, Tbl. 3.6, fits the data *quantitatively* worse than the primary model, cf. Tbl. 3.7, so that *a posteriori* odds are about $10^{16} : 1$ in favor of the primary model. Such agreement between the known biology and the quantitative

analysis again signals that the primary model captures the relevant biological mechanisms.

3.5 Discussion

We combined novel experimental tools for comprehensive behavioural assaying, mutants for genetic perturbations, and mathematical modelling to understand the thermal preference dynamics of *C. elegans*. We discovered that (i) the dynamics is governed by two independent pathways: positive association (habituation) to the experienced temperature independent of food, and negative association (avoidance) of the temperature with no food; that (ii) the avoidance time scale is slower than the habituation one; that (iii) the dynamics is quantitatively different in warm and in cold worms; that (iv) the full dynamical description requires at least four dynamical variables; and that (v) these dynamical variables are likely related to the worm's ability to generalize across similar temperatures. We *quantitatively* fit all of the available data (different life histories and genetic background) with high precision, indicating that the model has captured the underlying biological mechanisms.

While many fundamental genetic [42, 77] and neurophysiological [44, 57, 78] factors underlying *C. elegans* thermotaxis and thermal preference have been identified, small differences in experimental protocols (assay duration, gradient shape, range, and steepness, rearing protocols, etc.) cause large variation in *C. elegans* thermotaxis behaviour, which sometimes produce contradictory results [45, 54, 56, 68, 137]. This makes it difficult to definitively address open questions, such as the effect of food on the thermotactic response [21, 67, 137]. Nevertheless, our model's multiple reinforcement pathways, with the fastest of them operating on scales of less than half an hour, is consistent with many of these conflicting observations. Our discoveries support the existence of an active avoidance of thermal memory when food is absent [45]. They also demonstrate that food is not necessary for the establishment of the temperature preference consistent with other studies [12, 21, 23, 45, 67, 104, 137]. Further, our multidimensional framework can explain the previously ignored non-monotonic dependence of the temperature preference on the exposure

duration [12, 96]. Finally, it may account for the limited operating range of thermotaxis [96] as weaker generalization over larger temperature ranges.

Our model, Eqs. (3.2-3.6), is inspired by the animal learning literature [84, 97], and may seem foreign in the field of *C. elegans* thermal memory. We suggest that the aforementioned difficulties in explaining diverse experimental data come from the focus of traditional models on *one* dynamical preferred temperature, while our data demonstrate that there are, at least, two separate conditioning pathways. The traditional "one preference" model can be augmented to account for this. For example, the animal may be modeled to have a food-independent preferred temperature and a food-dependent temperature to avoid, with the joint push-pull of the two establishing the observed thermotactic pattern. Another possibility is that the animal may have a single set-point temperature, but also an internal state (or states) describing the strength and the nature (attraction or repulsion) of the temperature drive. The simplest versions of either of these models would have too few dynamical variables to explain our data, (cf. Section 3.6.5). To attain the necessary complexity, we would need to consider multi-scale temperature averaging, or more than two preference/avoidance temperatures and complicated generalization rules. With little known biology to build upon, such models would be indistinguishable from our functional, Rescorla-Wagner style model. Thus the next crucial experiments should move away from the "single preference" paradigm and focus on characterizing the dimensionality and the nature of thermal preference states, and the ability to generalize among them. Such experiments will likely benefit from modern interpretable machine learning techniques [25, 26], which can automate the search for a correct multi-dimensional model.

To explain spontaneous recovery after extinction, Pavlov speculated [90] that conditioned responses are produced by an interplay between excitation and inhibition. Extinction happens when inhibition overpowers excitation, and recovery when their magnitudes reverse again. Our findings can be viewed as a mathematical implementation of the idea, with the habituation channel playing the role of excitation and the avoidance channel the

role of inhibition. They are also consistent with his suggestion [90] that the strength of inhibition can be quantified by the time delay between extinction and spontaneous recovery, similar to the time delay between lobes of the same sign in Figs. 3.2, 3.3. Remarkably, with no further adjustments, separate excitation and inhibition also explain latent inhibition (CS that is common before appearance of the US establishes avoidance and hence requires a longer time to then develop a net positive association). They also establish asymmetric responses to appetitive vs. aversive cues and to conditioned association vs. conditioned inhibition, and (asymmetric) generalization among similar, but not equivalent cues. Crucially, two distinct mechanisms for positive and negative associations are also biologically plausible in larger animals (activation and excitation generally are mediated by different neurons and neuropeptides), but infrequently explored across the animal kingdom. In particular, while dopamine is implicated in associative learning in vertebrates [110], very little is known about what can mediate negative associations [109]. We thus hope that our discoveries in the worm, and especially identification of the two independent learning mechanisms, will spur a search for similar phenomenology and similar mechanisms in larger animals. Similarly, we hope that these ideas can make their way into improving artificial intelligence systems built on reinforcement learning [40].

Finally, it is intriguing to speculate about the functional importance of two independent thermal preference pathways, with the avoidance being reliably slower. When searching for food, animals start where they expect food to be found based on the conditioned stimuli. However, if the search there is unsuccessful, they turn to searching *elsewhere* (avoiding the already searched area) rather than *everywhere* (including the area just searched). Switching to *elsewhere* involves complex statistical strategies, such as Levy flights [10], which the worm exhibits as well [107]. A system with a single association strength would result in no preference and thus in searching *everywhere* as the association decays. In contrast, a distinct avoidance channel with a long time scale would guide the animal away from the condition with no food. Thus we expect that the multi-modal and multi-scale thermal

preference is optimal for fast search in temporally varying environments.

3.6 Methods

3.6.1 Nonlinear dynamics of animal learning

Here our goals are two-fold. First, we would like to show that traditional animal learning models map into the dynamics similar to Eq. (3.1). Second, we want to show that such dynamics, with realistic mapping between measurable variables and variables internal to the animal cannot produce oscillations in our experimental systems (and, in particular, zero crossings that we see in Fig. 3.2).

We start with the Rescorla-Wagner model of conditional associations [84, 97]:

$$V_i(t+\Delta t) = V_i(t) + \eta\delta_i(t)\left(\lambda_0\delta_\lambda(t) - \sum_j V_j(t)\delta_j(t)\right), \tag{3.7}$$

where $V_i(t)$ is the strength of association between the ith Conditioned Stimulus (CS) and the Unconditioned Stimulus (US), η is the learning rate, λ_0 the magnitude or salience of a single US reward, and Δt is the duration of one experimental epoch. Further, $\delta_i(t) = 1$ if the ith CS is present at time t, and 0 otherwise. Similarly, $\delta_\lambda(t) = \{1, 0\}$ depending on if the US was or was not present at time t. In other words, $\sum_j V_j(t)\delta_j(t)$ is the US prediction based on the strength of all CSs present at the time. Thus if a CS is present, then Eq. (3.7) posits that its US association strength is changed in proportion to how well all of the CSs predict the realized US reward, $\lambda_0\delta_\lambda(t)$.

We define the frequency of the CS and US presentations, $f_i = \langle\delta_i\rangle$ and $f_\lambda = \langle\delta_\lambda\rangle$, as well as the frequency of US given the ith CS $f_{\lambda|i} = \langle\delta_i\delta_\lambda\rangle/f_i$ and one CS given another $f_{j|i} = \langle\delta_i\delta_j\rangle/f_i$. We then introduce a time interval dt, which is small enough so that the association strength does not change much over its duration, $\eta\lambda_0 f_i f_{\lambda|i}\frac{dt}{\Delta t} \ll 1$, and which nonetheless is large enough so that it contains a lot of CS and US presentations,

$f_i f_{\lambda|i} \frac{dt}{\Delta t} \gg 1 \langle \delta_i \rangle \langle \delta_\lambda \rangle \Delta t \gg 1$. We integrate Eq. (3.7) over this interval to obtain

$$\sum_{t' \in (t, t+dt)} (V_i(t' + \Delta t) - V_i(t')) = dt \frac{\eta/\Delta t}{dt/\Delta t} \sum_{t' \in (t, t+dt)} \left(\lambda_0 \delta_\lambda(t') \delta_i(t') - \sum_j V_j(t') \delta_i(t') \delta_j(t') \right), \tag{3.8}$$

$$dV_i(t) = dt \frac{\eta}{\Delta t} f_i \left(\lambda_0 f_{\lambda|i}(t) - \sum_j V_j(t) f_{j|i}(t) \right), \tag{3.9}$$

$$\tau \frac{dV_i}{dt} = f_i(t) \left(\lambda_0 f_{\lambda|i} - \sum_j V_j(t) f_{j|i}(t) \right), \tag{3.10}$$

where $\tau = \Delta t / \eta$ is the learning time scale.

Sometimes all CSs are exclusive: for example, while many temperatures can serve as predictors of food, an animal can only experience one temperature at a time. In this case, $f_{j|i} = 0$ for $j \neq i$, and $f_{i|i} = 1$. Then Eq. (3.11) becomes

$$\tau \frac{dV_i}{dt} = f_i(t) \left(\lambda_0 f_{\lambda|i}(t) - V_i(t) \right). \tag{3.11}$$

This is Eq. (3.1) in the main text. Note that, for Pavlovian associations, $f_i(t)$ is under the control of an experimenters, while the animal can influence it in the operant conditioning protocols, so that $f_i(t) = f_i(V(t))$. Also note that in traditional analysis of conditioning, the CS must precede (and hence predict) the US for the association to form. In the context of our experiments, this temporal contingency structure can be disregarded because the CS temperature signals are experienced by the worm not episodically, but constitutively.

We now turn to showing that, in the case of our experiments, this version of the Rescorla-Wagner model cannot produce the experimentally observed crossings of the zero thermal bias line. First, notice that, because of the finite temperature resolution by the worm (cf. Section 3.6.6), there are only about two discernible CSs in the droplet: warm ($+$) and cold ($-$). Further, when the worm is not in the warm, it is in the cold, so that $f_+ = 1 - f_-$. In its turn, the worm's position and hence the experienced CS is affected by the strength of the associations, so that $f_+ = f_+(V_+ - V_-)$. In the droplet, there is no US (no food), so

that $f_{\lambda|\pm} = 0$. Thus in the droplet, the model in Eq. (3.11) becomes:

$$\tau \frac{dV_+}{dt} = -f_+(V_+ - V_-)V_+, \tag{3.12}$$

$$\tau \frac{dV_-}{dt} = -(1 - f_+(V_+ - V_-))V_-. \tag{3.13}$$

We now note that $f_+ \approx 0.5$ corresponds to a small thermal bias $f_+ - f_- = 1 - 2f_+ \approx 0$. The small bias is only possible when $V_+ \approx V_-$. In this regime, Eqs. (3.12, 3.13) further simplify:

$$\tau \frac{dV_\pm}{dt} = -0.5V_\pm. \tag{3.14}$$

These are equations for a simple exponential relaxation $V_\pm \to 0$, and both association strengths will decay with the same rate $0.5/\tau$. Thus if the thermal bias is near zero at some point in time, it must remain so indefinitely, and substantial oscillations are not possible.

One can create a single crossing of the zero thermal bias line using Eq. (3.14) if the time scales of the dynamics of the two associations are not the same, $\tau_+ \neq \tau_-$. However, this is likely insufficient to our experiment for three reasons. First, some of the mutant worms exhibit more than one zero crossing, cf. Fig. 3.3. Second, the cold and the warm sides of the droplet are less than a half a degree apart, which is much smaller than the range of temperatures the worm tolerates. Hence one would expect differences in time scales, if present, to be similarly relatively small. In this case, the overshoots of the zero line would be tiny and would take a long time to develop, while in all of our experiments the overshoots happen on the same time scales of the overall dynamics, cf. Figs. 3.2 and 3.3. Third, the worm is only reared at one temperature. Thus in the simplest model, we would have either $V_+ = 0$ or $V_- = 0$ at zero time. Thus a small temperature bias would require not just $V_+ = V_-$, but also $V_+ = V_- = 0$. In this case, even two distinct dynamical scales would not produce zero crossings. Collectively, these arguments suggest that a simple Rescorla-Wagner style model cannot account for the worm's thermal preference dynamics, and more complicated models are needed.

3.6.2 Strains and preparation

The mutant strains used in this study were as follows: We obtained strains (N2, *ins-1(nr2091)*, *daf-2(e1370)*, and *age-1(hx546)*) from the Caenorhabditis Genetics Center at the University of Minnesota.

All experiments used young adult animals cultivated at 15°C and $25°\mathrm{C}$ on nematode growth medium (NGM: 50 mM NaCl, 15 g/L agar, 20 g/L peptone, 1 g/L g/L, 1mM cholesterol, 1mM $\mathrm{CaCl_2}$, 1mM $\mathrm{KH_2PO_4}$ agar plates seeded with *Escherichia coli* strain OP50 under standard conditions [13]. M9 Buffer (3g $\mathrm{KH_2PO_4}$, 5 g NaCl, 6 g $\mathrm{Na_2HPO_4}$, 1 ml 1 M $\mathrm{MgSO_4}$, $\mathrm{H_2O}$) used for strain washes and assay. Animals were stage-synchronized using a standard bleach synchronization protocol [115]. We washed synchronized young adult animals with 1ml of M9 buffer into a 15mL Falcon tube, added an additional 10mL of M9 buffer, and pelleted animals by spinning at 0.4 RCF for 1 minute. We aspirated the supernatant and repeated the wash. In the case of 'well-fed' state experiments, we re-suspended the animals in 2ml of M9 and poured them onto a 5cm NGM plate. We then picked individuals into 2ml of buffer on a second 5cm NGM plate and transferred them into the μDroplet assay for observation. In the case of 'starved state' experiments, we decanted animals onto a 10cm NGM plate and allowed them to starve at their rearing temperatures as per protocol durations. We then washed animals off the plate using 2ml of M9 and decanted them onto a 5cm NGM plate. We picked individuals into 2ml of buffer on a second 5cm NGM and the transferred them into the μDroplet assay for observation.

3.6.3 μ droplet assay

Individual worms are picked into a grid of 3μL M9 buffer droplets on glass-printed 4mm diameter hydrophilic spots (Electron Microscopy Sciences Inc. Item #63430-04) surrounded by a hydrophobic PTFE (Teflon) surface. A coverslip placed on top of a 127 μm-thick silicone gasket bonded to the assay slide with VALAP (1:1:1 Vaseline, Lanolin, Paraffin).

The assay is centered and clamped on a temperature controlled aluminium stage. The stage is preset to either $15°$C or $25°$C depending on rearing temperature. Animals are acclimated for 5 minutes before imaging commences. The black-anodized Aluminium stage measures 165x58x3mm. Waterblocks (Swiftech MCW30) are secured to the ends of the stage through 40x40x3mm aluminium spacers coated with thermal paste. A 40x40mm peltier element (MCTE1-19908L-S) is secured under one spacer and waterblock. A waterbath circulator (Fisher Scientific IsoTemp 3016) is used to control the initial stage temperature. The peltier element is used to establish and program the temperature gradient. A 15-$25°$C stage temperature translates to a $1°$C/cm assay steepness and thus a temperature range of 19.8 to $20.2°$C on the gradient-aligned extremes of the droplet confirmed through direct IR camera observation and COMSOL Multiphysics modelling (data not shown).

Illumination for image capture is provided by two red LED light strips (3W 48-LED 180-Lumen Aluminum alloy light strip) positioned $\sim$ 9cm above the imaging stage. Raw monochrome images are captured via a DSLR camera (Nikon D7000) with a fixed macro lens (Nikon AF Micro Nikkor 60mm f/2.8D) controlled by Nikon's Camera Control Pro software. In order to minimize the correlation of the worms' position through time, capture rate is 6 frames per minute during long-term observation.

Image capture begins at the 5-minute mark in the absence of a temperature gradient for 100 frames ($\sim$ 16 minutes), after which a 15-$25°$C stage gradient is applied and recording continues to 1440 frames (4 hours). The absence of a gradient for the first 100 frames was used as a built-in control for each assay trial.

Experiments were performed successively 2-3 times per day. No randomization procedures were undertaken to control for time-of-day effects.

3.6.4 Data processing

We post-processed images and computed behavioural metrics using MATLAB. For this, we located regions of interest (ROI) around each droplet, subtracted the background, and

produced a binary image of the worm. We used the measure of worm area to filter out potential segmentation artifacts before calculating the center of mass position of the worm for each frame. The ROI determines the extremes of possible motion and is used to normalize the animal movement along a linear index from -1 ($19.8°$C) to +1 ($20.2°$C). When artifacts occur, such as in the identification of two objects in a droplet, we do not record a position in that frame. If a worm track is less than 95% complete, we discard it from the data set. A final thermal preference metric, the thermotaxis index, Θ, is calculated by summing all normalized values of an animal's movement in a certain time and dividing by the total number of observations.

Specifically, for each worm type (rearing temperature, mutation), all data are indexed by three indices: μ stands for the condition (i. e., the starvation duration), n stands for the nth individual worm in that condition, and t represents the time in hours since the beginning of assaying in the droplet. Since occasionally a worm is not tracked for some times due to image processing artifacts, the number $N_\mu(t)$ of individual worms tracked at time t in condition μ is time dependent.

The trajectory on the gradient of the n^{th} individual worm in condition μ at time t is $\Theta_{\mu,n}(t)$. Therefore, the average thermotactic index at time t is

$$\bar{\Theta}_\mu(t) = \sum_{n=1}^{N_\mu(t)} \frac{\Theta_{\mu,n}(t)}{N_\mu(t)}. \tag{3.15}$$

Similarly, the variance of thermotactic index at time t is

$$\mathrm{v}_\mu(t) = \sum_{n=1}^{N_\mu(t)} \frac{(\Theta_{\mu,n}(t) - \bar{\Theta}_\mu(t))^2}{N_\mu(t) - 1}. \tag{3.16}$$

Then the standard error of the mean for worms in condition a at time t,

$$s_\mu(t) = \left(\frac{\mathrm{v}_\mu(t)}{N_\mu(t)}\right)^{1/2}, \tag{3.17}$$

defines the experimental error bars on the thermotactic trajectory for the time and the condition.

As a measure of the overall noise of the data for worms in condition μ, we use the time averaged variance of the data trajectories over all worms in the condition μ

$$\bar{v}_\mu = \frac{1}{T}\sum_{t=0}^{T} v_\mu(t), \tag{3.18}$$

where $T = 4$ hours is the assay duration in the droplet.

Finally, for presentation purposes only (but not for the fitting), to remove rapid fluctuations, in all figures the thermotactic index and its experimental error are filtered through a causal exponential filter with the time scale of 6 min.

3.6.5 Bounding the dimensionality of the thermal memory dynamics

C. elegans thermal memory is a dynamical system. To model the behaviour, we first must estimate the number of dynamical variables in this system. Suppose that the dynamics, averaged over all worms, can be fully described by a d-dimensional time dependent vector $\vec{x}(t)$ that evolves according to an unknown, but continuous and deterministic dynamics. Suppose further that the one-dimensional observable $\Theta(t)$, the thermotactic index indicating the current temperature the worm guides itself to, is a smooth function of $\vec{x}(t)$, which may depend on all components of $\vec{x}(t)$. We estimate the dimensionality of $\vec{x}(t)$ using the following arguments.

Thermotactic memory dynamics is multidimensional Suppose that $d = 1$. Then, $x(t)$ and $\Theta(t)$ are functions of each other, and, for the dynamics to be well defined, the velocity $\Theta'(t)$ should be a single valued function of the position of the worm $\Theta(t)$. This is not the case. For example, at $\Theta(t) = \mp 0.1$, the thermotactic index $\Theta(t)$ for well-fed wild type worms at $15, 25°$C in Fig. 3.2 has two different thermotactic velocities $\Theta'(t)$ early and late in the experiment. Even simpler, the thermotactic index crosses a zero multiple times. Thus $d = \dim(\vec{x}(t)) > 1$.

Thermotactic memory dynamics has $d \geq 2$ To determine if $d = 2$ would be sufficient to capture the behaviour, we consider the delayed embedding representation of the

dynamics [119] as a k dimensional vector $\vec{\Theta}_k(t) = (\Theta(t), \Theta(t-\tau), \cdots, \Theta(t-(k-1)\tau)$. The Takens theorem [119] allows reconstruction of a d dimensional attractor from such $k = 2d+1$ dimensional embedding. However, if τ is much smaller than the characteristic time scale of the dynamics, then the d dimensional dynamics can be reconstructed uniquely from even smaller sequences, $\vec{\Theta}_k$ with $k = d$, provided the trajectories in the $\vec{\Theta}_k$ space do not self-intersect (to see this, notice that, at $\tau \to 0$, $\vec{\Theta}_k$ maps one-to-one onto $(\Theta(t), \Theta'(t), \Theta''(t), \ldots, \Theta^{(k-1)}(t))$. Thus to bound the dimensionality of $\vec{x}$, we seek the minimal k, for which delayed embedding trajectories at small τ show no self-intersections. The characteristic time scale of the thermal memory dynamics is $\mathcal{O}(1\mathrm{h})$ (cf. Fig. 3.2), and the trajectories show meaningful changes, statistically distinguishable from noise, for $\tau > \mathcal{O}(1\mathrm{s})$. This allows for a broad range of τ for our analysis. In what follows, we choose $\tau = 6.67$min (every 40th data point), but results are qualitatively the same for similar values. To quantify the uncertainty in the delayed embedding trajectories due to the experimental noise, we generate bundles of $n_\mathrm{t} = 20$ trajectories that are different by their starting time, $t_0 = 15\mathrm{min} + \frac{\tau}{2n_\mathrm{t}} i$, $i = 0, \ldots, n_\mathrm{t} - 1$, and we look for self-intersections of these trajectory bundles.

Figure 3.4 shows the trajectory bundles with $k = 2, 3$ for the 25°C worms starved for 1 hour. The 2-d embedding shows a clear intersection of the bundles, while the 3-d embedding does not. Performing the same analysis for all other starvation durations and both rearing temperatures (15°C and 25°C), we observe that $k = 3$ is always sufficient to avoid self-intersections of the bundles. Thus we conclude that $d \geq 3$.

Linear analysis suggests $d \geq 4$ For the well-fed and the 30 minutes starved worms raised at $25°$C, the observed thermotactic index $\Theta(t)$ is always far from the maximum value of the index, $\Theta_0 \approx 0.27$, observed for $15°$C (cf. Fig 3.2 B). We thus treat these trajectories as far from saturated, and assume that they can be produced by the dynamics linearized around $\Theta = 0$. These trajectories have different oscillation periods: about 3h for the well-fed

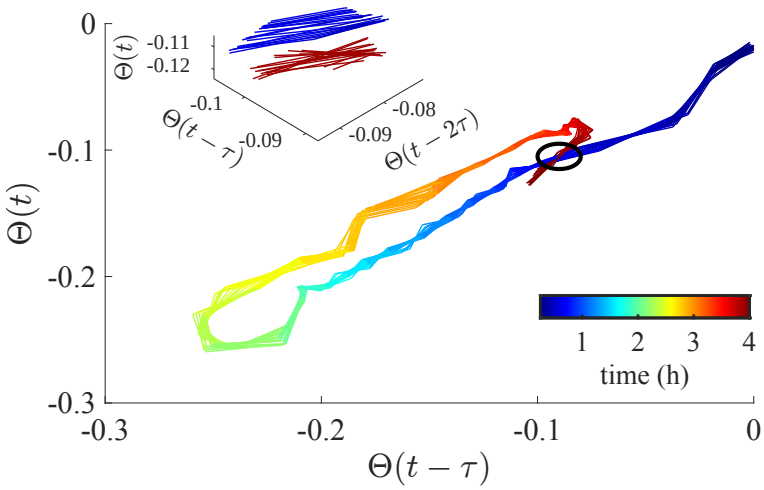

Figure 3.4: **Delayed embedding analysis of the thermotactic preference dynamics.** Average thermotactic index of the N2 worm reared at 25°C and starved for half an hour before assaying the thermotactic response in the droplet is analyzed using the delayed embedding coordinates. The time within the experiment is indicated by the color bar starting from 15 minutes (dark blue) and ending at 4 hours (dark red). The delay time $\tau \approx 6.67$ min for both the two and three dimensional embedding. Bundles of curves contain 20 curves each, corresponding to different starting offsets. The two dimensional embedding (main plot) has a self-intersection, indicated by a black circle. This confirms that the underlying dynamics is, at least, three dimensional. Zooming in on the relevant region in a three dimensional embedding (inset), we find no self-intersection.

worm and $>$ 4h for the starved ones. For a linear dynamical model, eigenvalues describing the oscillations come in complex conjugate pairs, each pair sharing the frequency. Thus two complex conjugate pairs are required to model the two distinct oscillation frequencies. We thus conclude that $d \geq 4$, which is consistent with all of the previous analyses.

3.6.6 Model

The same dynamical system can be represented in many different ways. Our focus is on finding a description that is interpretable, biologically plausible, and can be probed through realistic experiments. We aim for a parsimonious description, and hence look for a model with $d = 4$ dynamical variables, the smallest number of variables consistent with the analysis above.

Before we begin the model construction, we first comment on the difficulties that arise when modeling individual worms. In the experiment, the worm takes about 8 seconds

to go around the droplet. The time scales on which the worm measurably changes its bias are at minimum 15 minutes for the average worm. In those fifteen minutes there are about 100 measurements of the thermal position of the worm in the gradient. The error in determining the mean position of the worm scales as the inverse square-root of the number of measurements. Thus the error in the mean position over a 15 minute measurement period is of order 1/10. Given that the thermotactic index is maximum at 0.27 in our experiments, an error of 1/10 is about 37%. This is too large an error prevents any effective description of thermal learning of an individual worm.

The identity of the relevant internal states of the worm that store its thermal preference remains unclear [21]. While the only measurable quantity in our experiments is the thermotactic index, it is not an internal state. In diverse experiments, the worm develops either a preference to or avoidance of certain temperatures [45], and there has been a lively debate in the literature [12, 21, 23, 45, 67, 104, 137] regarding whether the avoidance or the preference is primary, and if either is conditioned on the presence of food or not. Knowing from the analysis above that the thermal preference dynamics is, in fact, multidimensional allows us to avoid the controversy, and to try various forms of the internal states. We explored many such models, and the only model that we found, which was able to fit the entire corpus of our data *quantitatively*, has separate dynamical variables representing which temperature the worm likes, and which temperature it avoids. Crucially, in this model, the attractive temperatures are not conditioned by food, but describe habituation: the worm likes the temperature that it has experienced before, which can be viewed as conditioning on prior survival. In contrast, the avoidance is conditioned by the *absence* of food: the worm does not like the temperature where the food has been unavailable previously. We stress that this is different from most models in the literature, which usually have only one dynamical variable, representing preference conditioned on the presence of food [12].

In principle, within the conditioned associations paradigm, one may model the habit-

uation and the avoidance as functions of the temperature field, so that every temperature has its own habituation and avoidance values, and the *generalization* of preferences to the nearby temperature ensures that these functions are smooth. However, the μ droplet assay establishes the gradient of only $0.4°$C across the entire droplet. The worms move stochastically, and their length is comparable to the droplet diameter. Thus geometrically they cannot stay at the poles of the droplet, where $\Theta = \pm 1$. Because of this, Θ usually spans only about a quarter of the possible range, between about $\pm 0.25 \ldots \pm -0.3$. Thus the worms typically experiences only $\sim 0.1°$C temperature range, which is comparable to their thermal sensitivity [23, 57, 78]. In other words, the worms cannot reliably distinguish more than about two temperature values in the droplet: warm and cold. In this situation, modelling the habituation and the avoidance as functions of the temperature field is unnecessarily complex, and we model the worm's internal states with two scalar variables. First, h (unconditioned positive preference, or habituation) takes positive values when the worm likes the warm side. Second, a (conditioned negative preference, or avoidance) is positive when the worm avoids the warm side. We then define the overall temperature preference as the combination of the habituation and the avoidance, passed through a saturating nonlinearity, as in Eq. (3.6).

With this, we identify our four dynamical variables with habituation and avoidance at the rearing temperature and at the assay temperature. The rearing variables change autonomously during the droplet assay, while the assay temperature variables are affected by the animal's behaviour. The dynamics is given by Eqs. (3.2-3.6).

Note that the rearing temperature preferences decay autonomously in Eqs. (3.4, 3.5) since these temperatures are not experienced in the assay. However, during the rearing time, the dynamics of h_r and a_r are expected to have the conditioning terms similar to h and a in Eqs. (3.2, 3.3). Finally, for parsimony, we do not model the generalization from the droplet to the rearing temperature.

It is useful to explain the dynamics described by Eqs. (3.2-3.6) narratively. A worm's

preference of a certain temperature emerges from a combination of an unconditioned habituation and a conditioned avoidance, the latter driven by the absence of food. Worms habituate to their rearing temperature. If the period of rearing is followed by starvation before being placed in the droplet, then additionally the avoidance of the rearing temperature sets in, roughly in proportion to the starvation time. Experiences during rearing are fully under the control of the experimenters, and thus represent classical or Pavlovian conditioning [90]. In contrast, the habituation and the avoidance of one side of the droplet are initially determined by generalizing the relevant preferences from the rearing temperature. However, as the worm begins to explore the droplet, it controls the temperature that it experiences, setting up an operant conditioning protocol, in which it habituates to the side of the droplet that it visits the most. At the same time, since there is no food in the droplet, the worm also develops the avoidance of the side it frequents. Crucially, this happens in the background of the decaying habituation and avoidance to the rearing temperature (and hence of the generalization effects), which cannot be reinforced while in the droplet and hence decay as solutions of the first order linear differential equation—exponentially with time.

3.6.7 Constraints on the model parameters

Here we discuss constraints on the model parameters, which we use in model fitting.

Constraints on initial conditions All worms in the experiment were raised for days, well-fed at their rearing temperature. In contrast, the typical time scales of the thermal preference is hours (cf. Fig. 3.2). This implies that the habituation to the rearing temperature is likely saturated for all worms. In other words, for all starvation conditions, $h_{\mathrm{r}}(0)$ is the same, depending only on the rearing temperature and the worm genetic background. The dynamics of habituation in the droplet in Eq. (3.2) is driven by the generalization for the entire duration of the rearing, and is likely to be saturated as well, at some value $h(0)$, the same for all starvation conditions, but also likely different for different rearing temper-

ature and the genetic background. Further, $h(0)$ is potentially different from A_h (saturation value in the droplet) since the maximum strength of conditioning may depend on the ambient temperature.

In our model, avoidance is a non-decreasing function of the starvation duration s. A typical starvation duration is a few hours long, so the avoidance states are probably not saturated. Thus the magnitude of the initial avoidance of the rearing temperature for different s, $|a_{s,\mathrm{r}}(0)|$, for worms raised in the same temperature, should be a non-decreasing function of s, $|a_{s_1,\mathrm{r}}(0)| \leq |a_{s_2,\mathrm{r}}(0)|$ for $s_1 < s_2$. Since avoidance is caused by starvation, the initial avoidance for a well-fed worm is zero, $|a_{0,\mathrm{r}}(0)| = 0$. The initial avoidance in the droplet emerges from the generalization of the avoidance at the rearing temperature, and thus should follow the same law: $|a_{s_1}(0)| \leq |a_{s_2}(0)|$ for $s_1 < s_2$, and $a_0(0) = 0$. Further, since the definition of the thermotactic index in the droplet sets $\Theta(0) = 0$ at $20°\mathrm{C}$, all initial conditions for worms raised at $15°\mathrm{C}$ will be non-positive and for those trained at $25°\mathrm{C}$ will be non-negative.

In summary, there are two initial conditions (rearing and droplet) for the habituation for each rearing temperature. There are 4 different starvation durations (and the well-fed worm) for the $15°\mathrm{C}$ worm, which corresponds to 8 ordered initial conditions for rearing and droplet avoidance, and 3 different starvation durations (in addition to the well-fed) for the $25°\mathrm{C}$ worm, resulting in 6 ordered avoidance initial conditions.

Constraints on dynamical parameters For parsimony, the habituation and avoidance time scales τ_h, τ_a, $\tau_{h,\mathrm{r}}$, $\tau_{a,\mathrm{r}}$, the maximum conditioning strengths A_h and A_a, and the bias c are assumed to be common among all starvation durations reared at the same temperature. From Fig. 3.2, we observe that the mean thermotactic index never goes above ≈ 0.27, so we choose $\Theta_0 = 0.27$ for all starvation and rearing temperatures to simplify the multidimensional fitting.

Since worms generally avoid high temperatures [137], the thermotactic index for starved

worms at $25°$C have a negative bias after ~ 3 hours, cf. Fig. 3.2. Such bias is not present at $15°$C. Thus we choose to set $c = 0$ for $15°$C worms, and to limit $c \leq 0$ for $25°$C worms.

3.6.8 Constraints on parameters for mutants

The *ins-1* mutation suppresses the starvation avoidance behaviour [58]. Thus in our main analysis, we make an assumption that it does not affect any of the parameters except τ_a and A_a (there is no starvation and hence no avoidance at the rearing temperature for any of the mutant worms). Since *ins-1(nr2091)* is a putative null mutation, we chose to model its effect as completely removing the avoidance behaviour, while keeping all other parameters inferred from the N2 worm unchanged. If this assumption is correct, then no additional fitting is needed to model the *ins-1* worms data after we set $A_a = 0$.

DAF-2 and AGE-1 are homologs of an insulin receptor tyrosine kinase and a phosphoinositide 3-kinase, respectively. Mutations in these genes are known to change the avoidance behaviour [58], and hence the avoidance time scale τ_a, strength A_a, and the bias c, as well as the avoidance parameters to the rearing temperature, may be changed by either mutation. Thus in our main analysis we fit these parameters to the mutant data (separately for the warm and the cold worms), while keeping the rest of the dynamical and the initial conditions parameters as inferred from the N2 worms reared at the appropriate temperature.

Note that we do not have starvation data for any of the mutants. Hence initial conditions for avoidance at the rearing and the assaying temperatures are set to 0, and $\tau_{a,15/25}$ do not enter the model.

3.6.9 Constructing the loss function

To fit models to the data we must define a loss function, which is minimized to find the optimal parameters. Designing the loss function for physical problems is, generally, complicated since it also must avoid guiding the optimization to nonphysical regimes. The following considerations went into choosing the loss in our problem.

We use a simple mean square difference between model prediction for the worm bias at

time t for the worm condition μ, $\hat{\Theta}_\mu(t)$, and the corresponding individual-averaged worm trajectories, $\bar{\Theta}_\mu(t)$, as a measure of the quality of fit

$$L_{\text{fit}} = \frac{1}{2}\sum_{\mu\in M}\sum_{n=0}^{T/\Delta t}\frac{1}{v_\mu(t)}(\hat{\Theta}_\mu(t_n) - \bar{\Theta}_\mu(t_n))^2. \tag{3.19}$$

$T = 4$ h is the duration of the experiment, $\Delta t = t_n - t_{n-1} = 10$ s and v_μ is the variance found in Eq. (3.16). The condition set $\mathcal{M}$ with size M consists of different starvation durations: $\{0, 1, 2, 3, 5\}$ h for 15°C, $M = 5$; and $\{0, 0.5, 1, 2\}$ h for 25°C, $M = 4$. Since each of the three mutants is well fed, $M = 1$ for each of them.

Removing non-biological behaviour Some parameter combinations that minimize L_{fit} result in the thermotactic behaviour that we consider biologically unrealistic: persistently oscillating bias at $t \gg T$. We expect the bias to converge to a constant value at long times because we are not aware of any animal that maintains non-diminishing amplitude oscillatory behaviour forever. The long-term constant bias potentially differs from zero due to rearing-temperature dependent asymmetry in the thermotactic behaviour [56, 68, 137], and it is potentially different from the value of the temperature dependent bias c in our model Eq. (3.6). We impose this by adding the following term to the loss function

$$L_{\text{far}} = \frac{\gamma}{2}\sum_{\mu\in\mathcal{M}}\sum_{n=0}^{T/\Delta t}\frac{M}{\sum_{\nu\in\mathcal{M}}\bar{v}_\nu}(\hat{\Theta}_\mu(t_n + T_{\text{far}}) - \bar{\hat{\Theta}})^2. \tag{3.20}$$

This term suppresses the mean squared difference between the model predicted trajectories $\hat{\Theta}_\mu(t+T_{\text{far}})$ starting at time $T_{\text{far}} = 16$h (long after the 4h of assaying in the droplet) and the time and the condition averaged model trajectories $\bar{\hat{\Theta}} = \frac{\Delta t}{M\cdot T}\sum_{\mu\in\mathcal{M}}\sum_{n=0}^{T/\Delta t}\hat{\Theta}_\mu(t_n + T_{\text{far}})$ after time T_{far}. Here $\Delta t = 10$s is the sampling interval, and the ratio $T/\Delta t = 4\text{h}/10\text{s} = 1440$ counts the number of sample points per trajectory. We use $\gamma = 0.1$ as a dimensionless weight in what follows. This is about the smallest value of γ that is able to prevent sustained oscillations in our fits. There is a wide range of $\gamma > 0.1$ that similarly suppresses the oscillations, and still does not result in the degradation of the quality of fit to the experimental data.

Removing flat regions in the loss function Empirically, our optimization landscape abounds with flat regions, where the model fits have low sensitivity to large correlated parameter changes. To break this near-invariance and allow gradient-based methods to find minima faster, we penalize very small and very large parameter values with

$$L_{\text{param}} = \lambda \sum_{k=1}^{K} (\ln \theta_k)^2, \tag{3.21}$$

where θ_k, $k = 1, \ldots, K$ are all parameters in the model. We use $\lambda = 0.1$ is what follows, which is about the smallest value of the constraint that still allows the gradient descent to converge. The quality of the fits evaluated with and without the constraint at this λ does not change significantly.

Normalization of the loss function In Eq. (3.19), L_{fit} scales as T, the total number of time points in the data. However, the data points in the experimental thermotactic trajectories are correlated because of the slow time scales of the worm behaviour. To reflect that there is an effective number of independent points in the thermotactic trajectory, we must normalize L_{fit} to scale linearly not with the total experiment duration, but with the number of independent measurements over the duration. For this, we compute the auto-correlation time T_{corr} of the residual $\hat{\Theta}_\mu(t) - \bar{\Theta}_\mu(t)$ over all conditions μ for worms raised at the specific rearing temperature. The number of independent measurements is then $n = T/T_{\text{corr}}$, or, alternatively, L_{fit} must be rescaled by $\Delta t/T_{\text{corr}}$, where $\Delta t = 10$s is the temporal resolution of the experiment. The auto-correlation time $T_{\text{corr}} \approx 13.3$ min for $15°C$, and it is $T_{\text{corr}} \approx 16.7$ min for $25°C$.

The overall loss function Overall, the considerations above yield the following loss function

$$L = \frac{\Delta t}{T_{\rm corr}} \left(\frac{1}{2} \sum_{\mu \in \mathcal{M}} \sum_{n=0}^{T/\Delta t} \frac{1}{v_\mu(t)} (\hat{\Theta}_\mu(t_n) - \bar{\Theta}_\mu(t_n))^2 \right.$$
$$\left. + \frac{\gamma}{2} \sum_{\mu \in \mathcal{M}} \sum_{n=0}^{T/\Delta t} \frac{M}{\sum_{\nu \in \mathcal{M}} \bar{v}_\nu} (\hat{\Theta}_\mu(t_n + T_{\rm far}) - \bar{\hat{\Theta}})^2 + \lambda \sum_{k=1}^{K} (\ln \theta_k)^2 \right), \quad (3.22)$$

which is what we optimize to fit parameters of the model to data.

3.6.10 Optimization and parameter values

To fit the model to the data we vary the optimization parameters $\vec{\theta}$ in order to minimize to the loss function. For each worm type (rearing temperature and mutation), we start from $\sim$ 1000 initial values of the parameters being optimized. Since some parameters take on a certain sign for biophysical reasons, we enforce these constraints by casting $\theta_i = \text{sign}(\theta_i) \exp(\theta_i')$. This means that $\vec{\theta'}$ is unconstrained and allows us to sample the parameters $\vec{\theta}$ across many orders of magnitude. To obtain parameter values which are $\mathcal{O}(1)$, components of the vector $\vec{\theta'}$ are sampled from a uniform distribution on the unit interval.

We then use the Quasi-Newton method to minimize the loss function starting from the initial value. To determine the quality of the fit we utilize the χ^2 per degree of freedom:

$$\chi^2/f = \frac{L_{\rm fit}}{Mf}. \quad (3.23)$$

This is a rescaled part of the of the overall loss function, responsible only for the quality of the fit. A value $\chi^2/f \sim 1$ implies an excellent fit of the model to all data points in the fitted condition.

The effects of γ on the quality of fit The loss function, Eq. (3.22), includes the term $L_{\rm far}$, which penalizes nonphysical sustained oscillations in the thermotactic index at long times. The term is weighted by a coefficient γ, compared to the quality of fit term. The

γ	.1	1	10	10^6
χ^2/f	1.17	1.17	1.17	1.25

Table 3.1: χ^2/f as a function of the dimensionless weight γ, which penalizes long-term oscillations. Quality of fit for the $15°$C worms with all starvation durations is tabulated at $\lambda = 0.1$. The impact of $\gamma > 0.1$ on the quality of fit is negligible.

λ	.1	1	10	100	1000
χ^2/f	1.17	1.18	1.18	1.19	1.76

Table 3.2: χ^2/f as a function of the dimensionless weight λ, which penalizes large and small parameter values. Quality of fit for the $15°$C worms with all starvation durations is tabulated at $\gamma = 0.1$. As long as λ stays below ~ 100, the effect on χ^2/f is minimal.

choice of γ is unclear *a priori*. In Tbl. 3.1, we report the dependence of χ^2/f on γ. The change in the quality of fit is $\sim 1\%$ in response to many orders of magnitude changes in γ. We conclude that L_{far} is able to dampen the oscillations without significantly altering the fit quality. We thus choose $\gamma = 0.1$ for all fits reported in this work, which is about the smallest value of γ able to suppress the oscillations.

The effects of λ on the quality of fit The third term of the loss function, L_{param}, penalizes very small and very large values of parameters. It is weighted by the parameter λ relative to the goodness of fit. In other words, the larger λ is, the smaller the region of parameters explored in fitting. Just like for γ above, setting the value of λ is impossible based on first principles. We explore the dependence of χ^2/f on λ in Tbl. 3.2. As long as $\lambda < 1000$, the dependence is minimal. We thus choose to work with the smallest value $\lambda = 0.1$, which enabled an effective parameter search, for all results reported here.

3.6.11 Model reduction and fitted values

We choose to use fitted parameter combinations that are strongly constrained by the data as a hard algebraic constraint on the parameter values, thus reducing the number of parameters by one per constraint. We measure the effects on such reduction on the quality of fit using the Bayesian Information Criteria (BIC) [111]. That is, the original model and

all of the reduced model are assigned a Bayesian score, which, in our case, is

$$BIC = k\ln(nM) + 2\left(\hat{L}_{\text{fit}}\frac{\Delta t}{T_{\text{corr}}}\right). \tag{3.24}$$

Here, as always, $n = T/T_{\text{corr}}$, and M is the number of different starvation conditions. Further, $\hat{L}_{\text{fit}}$ is the optimal value of the first (quality of fit) term in the loss function, Eq. (3.22), k is the total number of parameters in the model (including the initial condition that must be fit). The score balances the complexity of the model (the first term in Eq. (3.24)) with the quality of the fit (the second term). According to BIC, the reduced model is statistically better than the full, unconstrained model if its score is lower. More precisely, the posterior odds of two models are given by $P_1/P_2 \approx \exp(BIC_2 - BIC_1)$.

We emphasize that there are no first-principle reasons for the model reduction. In fact, they may not correspond to realistic biophysical constraints, may not be interpretable, and are only useful to the extent that they simplify the fitting, while not decreasing the quality of the fits significantly. Thus we do not push the reduction to the extreme, and stop when the algebraic constraints coming from the reduction become uninterpretable.

With this, for $15°$C wild type worms, we identify the following candidate parameter reductions. First, $a_{1,15}(0) = a_{0,15}(0) = 0$ and $a_{5,15}(0) = a_{3,15}(0)$, so that there are only two independent initial conditions for the avoidance at the rearing temperature. In other words, the rearing temperature avoidance for a well-fed worm and a 1 h starved worm are the same, and so are the avoidances for 3 and 5 h of starvation (the former presumably because 1 h of starvation is not enough to excite large avoidance, and the latter presumably because the avoidance gets saturated). Second, we expect $c = 0$, so that there is no long-term thermotactic bias. Third, we further verify if $\tau_{h,15}/\tau_{a,15} = \tau_h/\tau_a$. If true, this would signify that the ratio of habituation to avoidance time scales is independent of the temperature where it is measured (rearing or assaying).

There are 17 parameters in the unconstrained model for $15°$C data: four assaying and four rearing temperature avoidance initial conditions $a_s(0)$ and $a_{s,15}(0)$ for different starvation times (we remind the reader that $a_0(0) = a_{0,15}(0) = 0$ for this case by construction),

habituation and avoidance time scales for the memory at the assaying (τ_h, τ_a) and the rearing ($\tau_{h,15}$, $\tau_{a,15}$) temperatures, rearing and assay temperature habituation initial conditions ($h_{15}(0)$, $h(0)$) (both are assumed to be saturated and hence independent on the starvation duration), maximum possible values of avoidance (A_a) and habituation (A_h) at the assay temperature during assaying, and finally the long term bias c. The 17-parameter model has the BIC score of $BIC_{17} = 176.6$. We progressively reduce the parameters one at a time and calculate the BIC scores for the reduced models. First, $c = 0$ (16 parameters), we have $BIC_{16} = 173.2$. We then set $a_{1,15}(0) = a_{0,15}(0) = 0$ (15 parameters) to get $BIC_{15} = 168.6$. Further, setting $a_{5,15}(0) = a_{3,15}(0)$, we obtain $BIC_{14} = 165.0$. Finally, requesting that $\tau_p/\tau_n = \tau_{p,\mathrm{train}}/\tau_{n,\mathrm{train}}$ results in $BIC_{13} = 161.2$. This is the final model we use. Each of the models in the sequence is progressively more likely than the previous one. Specifically, the BIC scores give the odds $P_{13} : P_{17} \approx 2200 : 1$. The inferred parameters for this final model are in Tbl. 3.3. The best fit value for this model is $\chi^2/f = 1.17$.

For 25°C wild-type worms, we identify the following putative model reductions. First, we set $a_1(0) = a_{0.5}(0)$, so that the initial conditions for the avoidance at the droplet temperature are the same for starvation durations of 0.5 and 1.0 h. This is reasonable since the transfer of the worm from the rearing plate to the plate with no food for starvation itself takes ~ 0.25h, making the two starvation durations similar to each other. We then explore $A_h = A_a$, so that the maximum value of the habituation and avoidance are the same in the droplet. Finally, like for the cold worm, we try $\tau_h/\tau_a = \tau_{h,25}/\tau_{a,25}$, so that, while the actual time scales for the memories at the rearing and the experimental temperature may differ during the assaying in the droplet, the ratios of habituation and avoidance time scales stay the same.

The 25°C model starts with 15 parameters (we have one fewer starvation duration than for the cold worm, and hence two fewer initial conditions). For this model, $BIC_{15} = 123.6$. Setting $a_1(0) = a_{0.5}(0)$ results in $BIC_{14} = 120.2$. Adding $A_h = A_a$ gives $BIC_{13} = 116.6$. Finally, with $\tau_h/\tau_a = \tau_{h,25}/\tau_{a,25}$, we have $BIC_{12} = 113.2$. Again, every next model in

Parameter	τ_h/τ_a	A_h	$h(0)$	τ_a	A_a	$a_0(0)^*$	$a_1(0)$	$a_2(0)$	$a_3(0)$	$a_5(0)$
value	0.28	6.37	-1.92	1.39	6.44	0	-1.68	-1.86	-2.02	-2.91
Σ_{ii}	0.14	2.3	0.88	0.58	5.1		0.84	0.84	0.85	1.2
$h_{ii}^{-1/2}$	0.009	0.2	0.08	0.06	0.2		0.06	0.08	0.2	0.8
Parameter	$\tau_{h,15}/\tau_{a,15}$	$h_{15}(0)$	$\tau_{a,15}$	$a_{0,15}(0)^*$	$a_{1,15}(0)$	$a_{2,15}(0)$	$a_{3,15}(0)$	$a_{5,15}(0)$	c	Θ_0^*
value	τ_h/τ_a	-1.15	6.66	0	$a_{0,15}(0)$	-0.51	-1.41	$a_{3,15}(0)$	0	0.27
Σ_{ii}		0.48	4.9			0.42	0.84			
$h_{ii}^{-1/2}$		0.05	0.4			0.09	0.2			

Table 3.3: Optimal parameter values and their uncertainties (quantified by Σ and h, see text) for the final, 13-parameter model, describing the thermotactic dynamics of the cold N2 worm (reared at 15°C). BIC favors this model over a full 17-parameter model with odds $\approx 2200:1$. The quality of the fit is $\chi^2/f = 1.17$. Σ_{ii} are the estimates of the upper bound on the parameter uncertainty, accounting for variation of the other parameters, and $h_{ii}^{1/2}$ are the estimates of the lower bound on the uncertainty; see text for details. Parameters indicated by * are set a priori and are not fitted. Parameters with values relating them to other parameters are set by the model reduction. Parameter values are defined in Eqs. (3.2-3.6). Briefly: τ_h/τ_a – ratio of time scales for habituation and avoidance dynamics at the droplet temperature during the droplet assay; A_h – the maximum possible value of habituation at the droplet temperature during the assay; $h(0)$ – the initial value of the habituation at the droplet temperature at the beginning of the assay (hours); τ_a – the time scale of the avoidance at the droplet temperature during the assay; A_a – the maximum possible value of the avoidance at the droplet temperature during the assay; $a_s(0)$ – initial values of the avoidance at the droplet temperature at the beginning of the assay for starvation durations s; $\tau_{h,15}/\tau_{a,15}$ – the ratio of the time scales for the habituation and avoidance at the rearing temperature of 15°C during the assay; $h_{15}(0)$ – initial value of the habituation at the rearing temperature at the beginning of the assay; $a_{s,15}(0)$ – initial values of the avoidance at the rearing temperature at the beginning of the assay for starvation durations s; c – thermotactic bias; Θ_0 – saturating value of the worm bias.

the sequence is more likely than the previous one, and the odds $P_{12} : P_{15} \approx 181 : 1$. The inferred parameters for this final model are in Tbl. 3.4. The best fit value for this model is $\chi^2/f = 1.12$.

3.6.12 Mutant fits

As explained in Section 3.6.8, we inherit most of the parameters for the mutants from the N2 fits. We do not need to fit any additional parameters for *ins-1* mutants. For *age-1* and *daf-2*, we only fit the avoidance parameters τ_a and A_a, as well as the long-term bias c. While we try to keep $c = 0$ for the cold worms, as for N2, the obtained fit for the 15°C

Parameter	τ_h/τ_a	A_h	$h(0)$	τ_a	A_a	$a_0(0)^*$	$a_{0.5}(0)$	$a_1(0)$	$a_2(0)$
value	0.41	A_a	0.25	0.75	7.37	0	0.35.	$a_{0.5}(0)$	0.68
Σ_{ii}	0.29		0.27	0.39	5.2		0.28		0.46
$h_{ii}^{-1/2}$	0.02		0.1	0.06	0.6		0.1		0.3
Parameter	$\tau_{h,25}/\tau_{a,25}$	$h_{25}(0)$	$\tau_{a,25}$	$a_{0,25}(0)^*$	$a_{0.5,25}(0)$	$a_{1,25}(0)$	$a_{2,25}(0)$	c	Θ_0^*
value	τ_h/τ_a	0.68	2.66	0	0.63	1.2	1.42	-0.1	0.27
Σ_{ii}		0.36	1.4		0.37	0.49	0.66	0.14	
$h_{ii}^{-1/2}$		0.09	0.4		0.1	0.2	0.2	0.03	

Table 3.4: Optimal parameter values for the final, 12-parameter model, describing the thermotactic dynamics of the warm N2 worm (reared at $25°$C). BIC favors this model over a full 15 parameter model with odds $\approx 181:1$. The quality of the fit is $\chi^2/f = 1.12$. Notation used is the same as in Tbl. 3.3.

daf-2 worm is bad ($\chi^2/f > 2$), and so $c \neq 0$ is also allowed here. The fitted values of the parameters and the quality of fits are in Tbl. 3.5.

We additionally argue that fitting just the habituation (rather than the avoidance) parameters does not produce good fits, so that mutations, indeed, affect the avoidance and not the habituation pathway. The fitted values of the parameters and the quality of fit for this model are in Tbl. 3.6.

3.6.13 Parameter and model trajectory error bars

In Tbls. 3.3 and 3.4, we report the best fit parameter values for the reduced models for the cold and the warm worms. Since, even after the model reduction, there are still many parameter combinations that result in similar dynamics, we need to quantify the uncertainty on both the fitted parameter values, and on the dynamics themselves. Due to the near-degeneracy of the loss function for different parameter values, we consider the uncertainty on the dynamics a more important characteristic of the model fit than the parameter uncertainty.

Parameter error bars For estimating parameter uncertainties, we make the usual assumption that the loss function is quadratic in the parameter values around the optimum, $\vec{\theta}^*$. While not strictly true, the inaccuracies introduced by this assumption are not critical,

Temperature		$15°C$			$25°C$		
	Parameter	τ_a	A_a	c	τ_a	A_a	c
ins-1	value	0.39	0	0	6.9	0	1.5
	Σ_{ii}						
	$h_{ii}^{-1/2}$						
	Parameter	τ_a	A_a	c	τ_a^*	A_a^*	$c^\dagger$
daf-2	value	7.5	156	-9.7	1.2	3.3	0.026
	Σ_{ii}	6.6	260	11	0.64	1.1	0.19
	$h_{ii}^{-1/2}$	0.27	4.7	0.34	0.53	0.28	0.046
	Parameter	τ_a	A_a	c	τ_a	A_a	c
age-1	value	3.4	35	0	6.9	120	1.5
	Σ_{ii}	1.6	16		9.0	170	0.78
	$h_{ii}^{-1/2}$	0.08	0.23		0.32	5.5	0.22

Table 3.5: Optimal parameter values and their uncertainties for the parameters that change between the N2 worm and the mutant worms (*daf*-2 and *age*-1) in the model that assumes that the mutations affect the avoidance pathway only. There are 3 or 2 such parameters depending on the mutant and the rearing temperature. All notations are as in Tbl. 3.3. The quality of fit values are listed in Tbl. 3.7. Note that increases (decreases) in the parameter indexed by $*$ (†) within $+h_{ii}^{-1/2}$ ($-h_{ii}^{-1/2}$) result in nonphysical long-term sustained oscillations in the model. The *ins*-1 mutant is not fitted, instead A_a is set to zero while τ_a and c are inherited from Tbls. 3.3, and 3.4.

precisely because we consider the trajectories, and not the parameter values, as the important properties of the model fits. The Hessian around the optimal value $H_{ij} = \partial_{\theta_i}\partial_{\theta_j}L|_{\vec{\theta}^*}$ is estimated for us during the optimization by MATLAB's `fminunc()` using the finite differences method.

Parameters have correlated effects on model fits. Thus we report two measures of their uncertainty. The first is a lower bound on the error bar of each θ_i, obtained as $1/\sqrt{H_{ii}}$. This quantity measures how much the parameter can change and not affect the loss function significantly, while keeping all other parameters fixed. The second is the upper bound on the uncertainty in the quadratic approximation, $\Sigma_{ii} = \sqrt{(H^{-1}|_{\theta^*})_{ii}}$. This quantity measures how much a specific parameter can change without affecting the quality of the fit, while allowing variability in other parameters to compensate for the effects of changes in the explored parameter. Both errors are shown for the reduced model of the N2 worms in

Temperature		$15°C$			$25°C$		
	Parameter	τ_h	A_h	c	τ_h	A_h	c
ins-1	value	113.9	158.1	0	0.59	10.9	-0.1
	Σ_{ii}	19.9	20.1		0.17	0.87	
	$h_{ii}^{-1/2}$	19.8	20.0		0.12	0.59	
	Parameter	τ_h	A_h	c	τ_h	A_h	c
daf-2	value	0.11	5.5	-0.5	0.39	11.8	-0.014
	Σ_{ii}	0.04	0.14	0.07	0.16	0.86	0.12
	$h_{ii}^{-1/2}$	0.02	0.14	0.03	0.14	0.51	0.08
	Parameter	τ_h	A_h	c	τ_h	A_h	c
age-1	value	0.06	3.8	0	0.12	5.4	0.08
	Σ_{ii}	0.03	0.30		0.01	0.20	0.09
	$h_{ii}^{-1/2}$	0.02	0.25		0.01	0.18	0.08

Table 3.6: Optimal parameter values and their uncertainties for the parameters that change between the N2 worm and the mutant worms (*daf-2* , *age-1*, and *ins-1*) in an alternative model, where the mutations affect the habituation pathway only. All notations are as in Tbl. 3.3. The quality of fit values are listed in Tbl. 3.7. As the data in Tbl. 3.7 shows, the fits overall are worse than for our primary model.

.

Temperature	Avoidance		Habituation	
15°C	χ^2/f	BIC	χ^2/f	BIC
ins-1	1.28	23.1	1.19	27.2
daf-2	1.26	31.4	1.76	40.4
age-1	1.20	27.4	2.37	48.4
25°C	χ^2/f	BIC	χ^2/f	BIC
ins-1	1.59	22.9	0.89	18.2
daf-2	1.06	23.3	1.06	25.1
age-1	1.98	36.5	2.36	42.0
Total BIC		164.5		201.3

Table 3.7: χ^2 per degree of freedom and BIC scores for the model fits appearing in Tbl. 3.5 and 3.6. Overall, the BIC scores imply that the model where mutations affect the avoidance pathway rather than the habituation pathway explains the data much better, with the odds of $9.6 \times 10^{15}:1$ in its favor.

Tbls. 3.3 and 3.4, and for the mutant worms in Tbls. 3.5 and 3.6.

Wildtype trajectory error bands To estimate the uncertainty in the fitted dynamics, we use bootstrapping [49]. We use a single worm (rather than a single time point in

a worm's dynamics) as a single unit of data [108]. We sample with replacement N_μ worms from N_μ worms in the experimental condition μ (here μ defines the rearing temperature and the starvation duration). From the resampled data, for each condition, we calculate the new four-hour average thermotactic index curve $\bar{\Theta}_\mu(t)$. We repeat this a 100 times, thus obtaining $\{\bar{\Theta}_{\mu,k}(t)\}, k = 1, \ldots, 100$. Each of the bootstrapped trajectories is fit by the model, Eqs. (3.2-3.6), producing parameter fits $\{\hat{\vec{\theta}}_k\}$ and obtaining predictions $\{\hat{\Theta}_{\mu,k}(t)\}, k = 1, \ldots, 100$. For each time and condition, we report the error bands as the $16.5 \ldots 83.5$ range within the set $\{\hat{\Theta}_{\mu,k}(t)\}$. These bands would correspond to a one standard deviation confidence band if the statistics of the trajectories were Gaussian (which they are not).

Mutant trajectory error bands Mutants reared at warm or cold temperature inherit most parameters from their respective N2 worms, except for a handful parameters (those describing avoidance in the primary model, and habituation in the alternative model), which are fitted to the mutant data. Thus they inherit the bootstrapped N2 data as well, and the associated best fit parameters, $\{\hat{\vec{\theta}}_k\}$, and model predictions, $\{\bar{\Theta}_{\mu,k}(t)\}, k = 1, \ldots, 100$. If additional mutant-specific parameters are fitted, for such mutants, we generate the bootstrapped data with 100 sets of worms resampled with replacement, similarly to the wildtype worms. We then pair, at random, each mutant resampled data set with one of the N2 resampled data sets, inherited from estimating uncertainty of the N2 models. For each such pair, we use its fitted N2 parameters and fit the mutant parameters with these N2 ones. The 16.5 to 83.5 percentile range of the fits is then reported as the error band.

3.6.14 Data and material availability:

For the thermotactic index data please see [98].

Chapter 4 Entropy Estimation for under-sampled discrete distribution

4.1 Summary

[1] A fundamental problem in analysis of complex systems is getting a reliable estimate of entropy of their probability distributions over the state space. This is difficult because unsampled states can contribute substantially to the entropy, while they do not contribute to the Maximum Likelihood estimator of entropy, which replaces probabilities by the observed frequencies. Bayesian estimators overcome this obstacle by introducing a model of the low-probability tail of the probability distribution. Which statistical features of the observed data determine the model of the tail, and hence the output of such estimators, remains unclear. Here we show that well-known entropy estimators for probability distributions on discrete state spaces model the structure of the low probability tail based largely on few statistics of the data: the sample size, the Maximum Likelihood estimate, the number of coincidences among the samples, the dispersion of the coincidences. We derive approximate analytical entropy estimators for undersampled distributions based on these statistics, and we use the results to propose an intuitive understanding of how the Bayesian entropy estimators work.

4.2 Introduction

Estimating entropy – that is, the measure of uncertainty [24, 112] – of a random variable from its samples is often a key question in analysis of complex systems. This estimation from a finite (and often small) set of samples is a hard problem, especially for high dimensional systems, where the number of states that a variable can take quickly overwhelms

[1] This chapter presents the paper [48] which is a collaboration with Damián G. Hernández, Ilya Nemenman.

the number of samples N. Then many of the states, hereafter called *low probability states*, have probability $< 1/N$. Collectively, we refer to all of these states as the *tail* of the probability distribution. While there may be a lot of samples in the tail, each low probability state will not be sampled typically, or will be sampled at most once. Because of the tail, the entropy estimator that replaces probabilities of states by their empirical frequencies (the so called *naive* or *Maximum Likelihood* estimator [116]) has a large sample size dependent bias [89]. Corrections have been derived to overcome this bias [11, 43, 75], but these tend to be valid only in the well-sampled regime. Outside of this regime, Bayesian [6, 82, 134] and some non-parametric [16–18] estimators may still result in low bias estimates by imposing *a priori* assumptions on the probabilities of the low-probability states.

Although these Bayesian and non-parametric estimators perform well on some data sets, it is known that no estimator can be universally unbiased in this regime [4, 89]. Thus it is crucial to understand how these estimators extract information about entropy from data, and hence when they will fail. Unfortunately, such theoretical understanding is missing for many estimators. Ma was the first to point out that estimation of entropy is possible for poorly-sampled uniform distributions by analysing a particular statistics of the data: *coincidences* [69]. Nemenman extended the theoretical idea that coincidences determine entropy to non-uniform distributions obeying some Bayesian priors [80]. However, a similar theoretical understanding is still missing in a broader context, and it remains unclear which statistics of data, in addition to the number of coincidences, may contribute to entropy estimation and why.

In this paper, we analytically investigate two Bayesian estimators: that of Nemenman, Shafee and Bialek [81, 82] and of Archer and Pillow [6]. We focus on the regime, which is arguably the most important for real life applications, where the number of states with at least one sample, K_1, is similar to the total number of samples, $K_1 \sim N \gg 1$, and yet $K_1 < N$, so that there are coincidences in the data. Outside of this regime, the probability distribution is either well-sampled (so that many different methods for entropy estimation

would work), or there are no coincidences at all (so that entropy estimation is impossible). In our regime of interest, we show that the result of the estimation by the studied estimators depends on the Maximum Likelihood entropy estimate S_0, the number of coincidences, and also on two measures of *dispersion of coincidences*. The first of these, K_2, is the number of states with at least two samples. The second, which we call Q_1, characterizes the spread of coincidences over states with three or more samples.

We show that values of these statistics are related to the structure of the tails of the probability distribution that is assumed by the estimators. Specifically, a short, exponential, tail is more likely to be inferred by the estimators when there many coincidences or they are dispersed. If the number of coincidences is intermediate, and the coincidences are concentrated, then the estimators infer a long tail [2]. In between these two regions, a mixed tail dominates. We show that the studied estimators correct Maximum Likelihood, and that the correction is larger when there are fewer coincidences and they are concentrated, which in turn happens with a large exponential tail or a slowly-decaying long tail. This understanding relates the observable data statistics to assumptions that Bayesian estimators make about the underlying probability distributions (see Fig. 4.1), and hence provides an intuitive explanation for how these estimators work and, crucially, when they fail.

4.3 Overview of Bayesian entropy estimation

Given a probability distribution $\{q_x\} = \boldsymbol{q}$ for a discrete one-dimensional random variable X, its entropy is defined as [112]

$$S(\boldsymbol{q}) = -\sum_x q_x \log q_x. \tag{4.1}$$

Note that we use the natural logarithm throughout this paper, and hence entropy is measured in *nats*. One is often faced with a problem when S must be estimated for unknown q_x from a set of N samples $\{x_1, \ldots, x_N\}$ from the probability distribution. The Maximum

[2] a long tail is power law tail

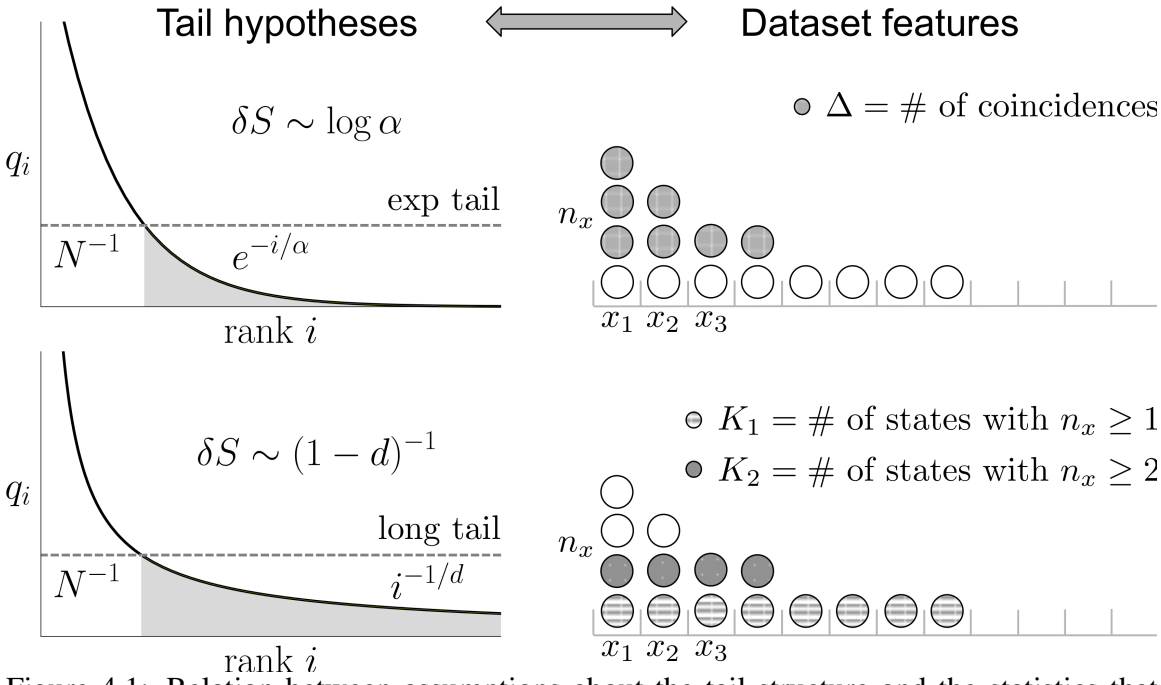

Figure 4.1: Relation between assumptions about the tail structure and the statistics that determine entropy estimation. The set of unsampled states, $q_i \leq 1/N$, which we refer to as the *tail*, may contribute substantially to the entropy. However, the Maximum Likelihood estimation overlooks this contribution. If the rank ordered plot of the tail is exponential with the scale α (top panel), then the tail has effectively α states, which contribute $\delta S \sim \log \alpha$ to the entropy. While the tail cannot be observed directly, it pulls samples from the head of the distribution, so that the number of coincidences, Δ, in the head decreases as α grows. Thus one can estimate α and hence the entropy itself from Δ. Alternatively, if the rank-ordered plot of the tail has a power law structure with the exponent $-1/d$, then the tail does not have a finite effective size (bottom panels). Then its contribution to entropy depends on d as $\delta S \sim (1-d)^{-1}$. In this case, one can estimate d, and hence the entropy, from the dispersion of the coincidences, which depends, in part, on how many samples happen once or more, K_1, or twice or more, K_2, in the dataset.

Likelihood estimator of entropy, S_0, is then defined by replacing the probabilities with frequencies $q_x \to \hat{q}_x = n_x/N$,

$$S_0 = S(\hat{\boldsymbol{q}}) = -\sum_x \frac{n_x}{N} \log \frac{n_x}{N}. \tag{4.2}$$

States with zero frequencies in the sample do not contribute to S_0 resulting typically in underestimation of the entropy [89]. In general, because of this low probability tail, estimation of entropy from data is very hard when the number of samples is smaller than the number of effective states of the variable, $N \ll \exp(S)$.

Bayesian estimators address the problem by imposing various *a priori* assumptions

$p(\boldsymbol{q})$. One then uses Bayes theorem to infer the *a posteriori* distribution of $\boldsymbol{q}$, and finally integrates over $\boldsymbol{q}$ to get the *a posteriori* distribution or moments of entropy. Specifically, the mean posterior entropy $\hat{S} = \langle S|\boldsymbol{n}\rangle$ given the counts $\boldsymbol{n} = \{n_x\}$ of how many times state x was sampled is given by

$$\begin{aligned}\hat{S} = \langle S|\boldsymbol{n}\rangle &= \int S(\boldsymbol{q})p(S|\boldsymbol{q})p(\boldsymbol{q}|\boldsymbol{n})d\boldsymbol{q} \\ &= \int S(\boldsymbol{q})\delta\left(S + \sum_x q_x \log q_x\right) p(\boldsymbol{q}|\boldsymbol{n})d\boldsymbol{q},\end{aligned} \tag{4.3}$$

where $p(\boldsymbol{q}|\boldsymbol{n})$ is the posterior over $\boldsymbol{q}$ under some prior $p(\boldsymbol{q})$,

$$p(\boldsymbol{q}|\boldsymbol{n}) = \frac{p(\boldsymbol{n}|\boldsymbol{q})p(\boldsymbol{q})}{p(\boldsymbol{n})} = \frac{\prod_x q_x^{n_x} p(\boldsymbol{q})}{p(\boldsymbol{n})}. \tag{4.4}$$

For distributions with known finite size $\mathcal{A}$ of the space of the possible outcomes (aka the *alphabet size*), the Dirichlet distribution is often chosen as a prior due to its conjugacy with the categorical distribution:

$$p(\boldsymbol{q}) = \text{Dirichlet}(\boldsymbol{q}|\lambda) \propto \prod_{i=1}^{\mathcal{A}} q_i^{\lambda}, \tag{4.5}$$

where λ is known as the concentration parameter.

Note that any chosen prior $p(\boldsymbol{q})$ implicitly imposes assumptions on the structure of the low probability tail (and hence its contribution to the entropy) based on the observed statistics of the well-sampled part of the probability distribution. However, these implicit assumptions usually are not made explicit, and they remain mysterious even for most commonly used Bayesian estimators. Lifting this veil is the goal of this work.

4.3.1 The Nemenman-Shafee-Bialek (NSB) Estimator

Nemenman et al. [82] showed that, for variables with the finite alphabet size $\mathcal{A}$, Dirichlet priors on $\boldsymbol{q}$ with a fixed value for the concentration parameter λ correspond to highly concentrated *a priori* distribution on entropy, which persists for large sample sizes. This bias induces incorrect entropy estimates, which nonetheless have low variance and hence

are certain about their outputs. To address this issue, Ref. [82] suggested a Dirichlet-mixture prior

$$p_{\rm NSB}(\boldsymbol{q}) = \int \text{Dirichlet}(q|\lambda) p_{\rm prior}(\lambda) d\lambda, \tag{4.6}$$

where $p(\lambda)$ are the mixture weights determined by

$$p_{\rm prior}(\lambda) \propto \partial_\lambda \langle S|\lambda\rangle = \mathcal{A}\psi_1(\mathcal{A}\lambda+1) - \psi_1(\lambda+1), \tag{4.7}$$

and where $\langle S|\lambda\rangle$ is the *a priori* expected entropy under the Dirichlet($\boldsymbol{q}|\lambda$) prior, and $\psi_1(\cdot)$ is the tri-gamma function [1]. This choice of weights implies a nearly uniform *a priori* distribution for the entropy S on the interval $[0, \log \mathcal{A}]$. The resulting entropy estimate is then

$$\begin{aligned} \hat{S}_{\rm NSB} = \langle S|\boldsymbol{n}\rangle &= \int\int S(\boldsymbol{q}) p(\boldsymbol{q}|\boldsymbol{n},\lambda) p(\lambda|\boldsymbol{n}) d\boldsymbol{q} d\lambda \\ &= \int \langle S|\boldsymbol{n},\lambda\rangle \frac{p(\boldsymbol{n}|\lambda) p_{\rm prior}(\lambda)}{p(\boldsymbol{n})} d\lambda. \end{aligned} \tag{4.8}$$

Here $\langle S|\boldsymbol{n},\lambda\rangle$ is the posterior mean entropy under the prior Dirichlet($\boldsymbol{q}|\lambda$), and $p(\boldsymbol{n}|\lambda)$ is the evidence (which has a Polya distribution) [76],

$$\begin{aligned} p(\boldsymbol{n}|\lambda) &= \int p(\boldsymbol{n}|\boldsymbol{q}) p(\boldsymbol{q}|\lambda) d\boldsymbol{q} \\ &= \frac{N!\,\Gamma(\mathcal{A}\lambda)}{\Gamma(\lambda)^{\mathcal{A}}\Gamma(N+\mathcal{A}\lambda)} \prod_{i=1}^{\mathcal{A}} \frac{\Gamma(n_i+\lambda)}{n_i!} \end{aligned} \tag{4.9}$$

where $\Gamma(\cdot)$ is the gamma function [1]. Using the analytical expressions for the first two moments of posterior mean entropy $\langle S|\boldsymbol{n},\lambda\rangle$ (available from Refs. [82, 134]), one then uses one-dimensional numerical integration over λ to obtain $\hat{S}_{\rm NSB}$.

4.3.2 The Dirichlet and the Pitman-Yor Processes

When the size of the state space is unknown or infinite, the standard NSB construction does not work. Then one commonly uses one of the following two stochastic processes to construct a prior $p(\boldsymbol{q})$ over a countably infinite state space: the Pitman-Yor Process (PYP) [92] and its special case, the Dirichlet Process (DP) [39]. To specify these processes, one

requires two inputs: a parameter vector and a base distribution. Parameters of the Pitman-Yor process are known as the discount parameter d, $0 \leq d < 1$, and the concentration parameter α. The parameters control the shape of typical distributions generated by the process. Specifically, d controls the structure of the low probability tail of $\boldsymbol{q}$, so that the tail typically decays as $q_x \propto x^{-1/d}$. The concentration parameter α control the probability mass near the head of the distribution. In the limit $d \to 0$, PYP(d, α) becomes the Dirichlet Process, DP(α). In other words, the Dirichlet Process generates distributions with short tails.

When the base distribution is the Beta distribution, one draws samples $q_x \sim$ PYP(d, α) via the so called *stick-breaking process* [53], which uses an infinite sequence of independent Beta-distributed random variables $\beta_x \sim$ Beta$(1 - d, \alpha + xd)$, so that

$$\tilde{q}_x = \beta_x \prod_{y=1}^{x-1} (1 - \beta_y). \tag{4.10}$$

Thus obtained $\tilde{\boldsymbol{q}}$ are not strictly decreasing with x, and so one obtains a strictly non-increasing distribution $\boldsymbol{q}$ from them by rank ordering.

4.3.3 Expectations over DP and PYP Posteriors

Previous studies [52] showed that PYP priors (for multinomial observations) yield a posterior $p(\boldsymbol{q}|\boldsymbol{n}, \alpha, d)$, which consists of two parts: probability of K_1 states that exist in the sample with the counts of, at least, one, and probability of states that are not sampled. We will denote the set of states with nonzero counts as $\mathbb{K}$, and its cardinality is $K_1 = ||\mathbb{K}||$. Then the first term of the posterior is given by the Dirichlet distribution, $p(\boldsymbol{q} \in \mathbb{K}|\boldsymbol{\mu}) \propto \prod_x q_x^{\mu_x}$, where $\boldsymbol{\mu}$ is a concentration vector $\boldsymbol{\mu} = (n_1 - d, \cdots, n_{K_1} - d, \alpha + K_1 d)$. This leaves the probability of $q_* = 1 - \sum_{x \in \mathbb{K}} q_x$ for the unobserved states. In other words, the states

with nonzero counts contribute the following to the posterior:

$$\begin{aligned} p(\boldsymbol{q} \in \mathbb{K}|\boldsymbol{n}) &= p(q_1, \cdots, q_{K_1}, q_*|\boldsymbol{n}) \\ &= \text{Dirichlet}(n_1 - d, \cdots, n_{K_1} - d, \alpha + K_1 d) \\ &\propto q_*^{\alpha + K_1 d} \prod_{i=1}^{K_1} q_i^{n_i - d}. \end{aligned} \tag{4.11}$$

For the states that have no samples, the posterior is equal to the prior. Thus their contribution to the posterior is the Pitman-Yor Process, normalized by their total probability being q^*:

$$p(\boldsymbol{q} \notin \mathbb{K}) = p(q_{K_1+1}, q_{K_1+2}, \cdots) = q^* \text{PYP}(d, \alpha + K_1 d). \tag{4.12}$$

Overall, this yields a closed form solution for the posterior mean and variance of the entropy S. Specifically, the resulting posterior mean $\langle S|\boldsymbol{n}, \alpha, d\rangle$ is

$$\begin{aligned} \langle S|\boldsymbol{n}, \alpha, d\rangle = \psi(\alpha + N + 1) - \frac{\alpha + K_1 d}{\alpha + N}\psi(1 - d) \\ -\frac{1}{\alpha + N}\left(\sum_{x=1}^{K_1}(n_x - d)\psi(n_x - d + 1)\right), \end{aligned} \tag{4.13}$$

where $\psi(x) = \partial_x \log \Gamma(x)$ is the di-gamma function [1]. Unfortunately, this is usually not a good estimate of entropy since, for fixed α and d, the prior PYP(d, α) on $\boldsymbol{q}$ corresponds to a highly concentrated *a priori* distribution on entropy, just like was noted before in the context of the NSB estimator. To counter this, Archer and Pillow [6] followed the NSB prescription and introduced a prior (mixture) over the parameters of $PYP(d, \alpha)$, $p_{\text{prior}}(\alpha, d)$, which uniformized the induced prior over entropy (with the caveat that, for a distribution on a countable alphabet, the entropy may be infinite, and hence strict uniform distribution over entropy is impossible). Specifically, they used

$$p_{\text{prior}}(\alpha, d) = p(\gamma) = e^{-10/(1-\gamma)}, \qquad \text{where} \tag{4.14}$$

$$\gamma = (\psi(1) - \psi(1 - d))/(\psi(\alpha + 1) - \psi(1 - d)), \tag{4.15}$$

and then they confirmed numerically that this choice of the prior leads to good estimates of entropy for various test data sets. In other words, they proposed a new estimate of entropy,

the Pitman-Yor Mixture (PYM):

$$\hat{S}_{PYM} = \langle S|\boldsymbol{n}\rangle = \int \langle S|\boldsymbol{n},\alpha,d\rangle p_{\text{posterior}}(\alpha,d|\boldsymbol{n})d(\alpha,d)$$
$$= \int \langle S|\boldsymbol{n},\alpha,d\rangle \frac{p(\boldsymbol{n}|\alpha,d)p_{\text{prior}}(\alpha,d)}{p(\boldsymbol{n})}d(\alpha,d), \tag{4.16}$$

where $\langle S|\boldsymbol{n},\alpha,d\rangle$ is given in Eq. (4.13). The evidence $p(\boldsymbol{n}|\alpha,d)$ is then given by (see Ref. [6] for a detailed derivation)

$$p(\boldsymbol{n}|\alpha,d) = \frac{\Gamma(1+\alpha)\prod_{l=1}^{K_1}(\alpha+ld)\prod_{x=1}^{K_1}\Gamma(n_x-d)}{\Gamma(1-d)^{K_1}\Gamma(\alpha+N)}. \tag{4.17}$$

Note that taking $d \to 0$ in Eqs. (4.16 and 4.17) and making the identification $\alpha = \mathcal{A}\lambda$ in the limits $\lambda \to 0$ and $\mathcal{A} \to \infty$ such that α is finite, result in a countably-infinite analogue of the NSB estimator.

4.4 Determining data statistics that define entropy estimates

In the section, we approximate the likelihood function of the Pitman-Yor process, Eq. (4.17), analytically in terms of coincidence-based data statistics. We then numerically show that the resulting analytical entropy estimates are close to the exact Pitman-Yor Mixture estimator. We focus on the regime where the Maximum Likelihood entropy estimator fails dramatically. For this, we study random variables with many accessible states in the regime where the number of unique samples, K_1, is of the order of the total sample size N. This regime corresponds to $K_1 \lesssim N \leq \exp(S)$, where N is the number of samples and S is the true entropy.

We start by considering the log-likelihood function, which is the logarithm of the evidence $p(\boldsymbol{n}|\alpha,d)$ in Eq. (4.17):

$$\mathcal{L}(\mathbf{n}|\alpha,d) = \log\Gamma(1+\alpha) - \log\Gamma(N+\alpha) + \log\Gamma\left(\frac{\alpha}{d}+K_1\right)$$
$$- \log\Gamma\left(\frac{\alpha}{d}+1\right) + \sum_{i=1}^{K_1}\log\Gamma(n_i-d) - K_1\log\Gamma(1-d). \tag{4.18}$$

We now define K_m as the number of states with at least m counts in the total sample of size N, $K_m = \sum_{n_i \geq m} 1$. We denote by m_f the largest occupancy of any state in the sample. Further, we define $\mathcal{K}$ as the vector, whose mth element is K_m. We note that, for any function $f(n)$,

$$\sum_i f(n_i) = \sum_m (K_m - K_{m+1}) f(m). \tag{4.19}$$

Thus, in particular, the log-likelihood $\mathcal{L}(\mathbf{n}|\alpha, d)$ can be viewed as $\mathcal{L}(\mathcal{K}|\alpha, d)$. With this, we can expand Eq. (4.18) around $d = 0$ to get (see Appendix 4.7.1 for details):

$$\begin{aligned}\mathcal{L}(\boldsymbol{n}|\alpha, d) \approx \mathcal{L}_a(\mathcal{K}|\alpha, d) \equiv \log\Gamma(1+\alpha) \\ - \log\Gamma(N+\alpha) + \log\Gamma\left(\frac{\alpha}{d} + K_1\right) - \log\Gamma\left(\frac{\alpha}{d} + 1\right) \\ + (K_1 - 1)\log d + K_2 \log(1-d) - Q_1 d + \mathcal{O}(d^2),\end{aligned} \tag{4.20}$$

where

$$Q_1 = \sum_{m=3}^{m_f} \frac{K_m}{m-1}, \tag{4.21}$$

and the subscript a denotes the $d \to 0$ asymptotic nature of the expression.

By rewriting the Maximum Likelihood estimate S_0 of Eq. (4.2) in terms of coincidences (see Appendix 4.7.2), using the identity Eq. (4.19), and approximating certain terms that are finite in the limit $d \to 1$ via a Taylor expansion around $d \ll 1$, the mean posterior entropy, Eq. (4.13), results in (see Appendix 4.7.3):

$$\begin{aligned}\langle S|\mathbf{n}, \alpha, d\rangle \approx \langle S|\mathcal{K}, \alpha, d\rangle_a \equiv \psi(N+\alpha+1) \\ - \left(\frac{\alpha + K_1}{\alpha + N}\right)\psi(1-d) + \frac{1}{\alpha+N}\Bigg[.N(S_0 - \log N) - K_1 \\ + K_2(\log 4 - 1 - \psi(2-d)) + Q_1 d \\ + \mathcal{O}\left(d^2, \sum_{m=3} \frac{K_m}{(m-1)^2}\right)\Bigg],\end{aligned} \tag{4.22}$$

where $\mathcal{O}(d^2, \sum_{m=3} K_m/m^2)$ means that we kept terms that are at most linear in d and at most proportional to $\sum_{m=3} \frac{K_m}{(m-1)}$. Interestingly, within this approximation, the log-likelihood and the posterior mean entropy depend on the sample size N, the Maximum

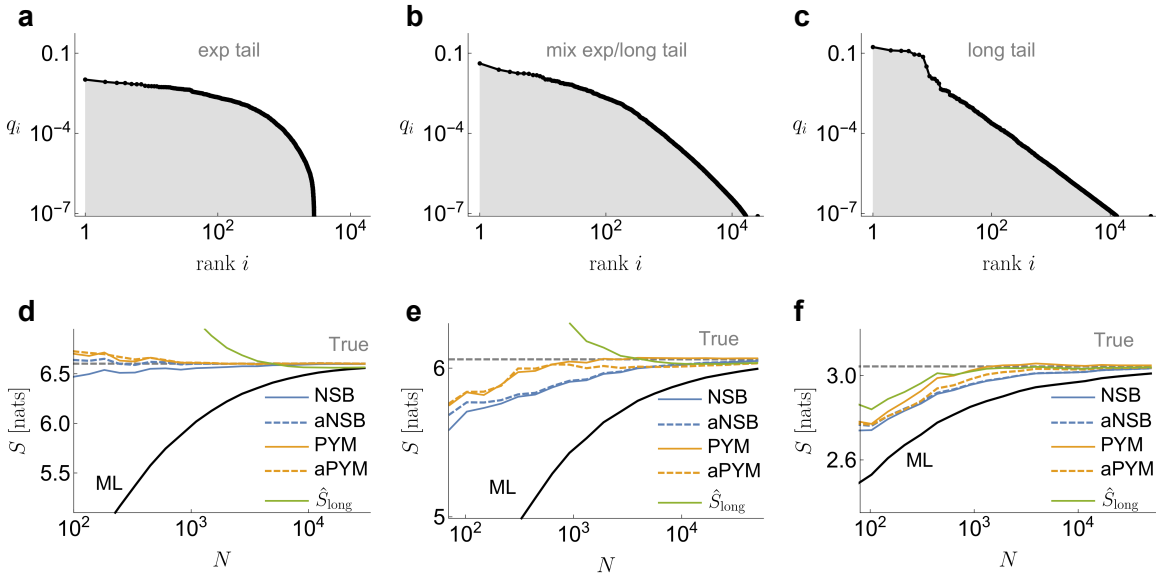

Figure 4.2: Comparison between PYM and related estimators and their approximations for distributions with different tails. The upper panels (**a**-**c**) show the distributions, whose entropy is being estimated. The lower panels (**d**-**f**) show the corresponding entropy estimates as a function of the number of samples, averaged over ten sets of samples. The full estimators, PYM and NSB (with a large alphabet size $\mathcal{A} = 20K_1$), almost overlap with our approximations, aPYM and aNSB. In all panels, we show results for Maximum Likelihood (black), NSB (blue), aNSB (dashed blue), PYM (orange), aPYM (dashed orange), and $\hat{S}_{\rm long}$ (green) estimators. The dashed gray line represents the true value of entropy for each of the studied distributions.

Likelihood entropy estimate S_0, and the three characteristics of the coincidence vector: K_1, K_2 and Q_1.

The final step in approximating the estimator $\hat{S}_{PYM}$, Eq. (4.16), is to integrate the expected entropy for fixed hyper-parameters $\langle S|\mathcal{K},\alpha,d\rangle_a$ over the posterior $p_{\rm posterior}(\alpha,d|\mathbf{n}) \propto p(\mathbf{n}|\alpha,d)p_{\rm prior}(\alpha,d)$ to form the Pitman-Yor mixture. Then the variance of the resulting estimator is dominated by the contribution from the uncertainty in the posterior distribution of the parameters α, d, which is about 80% of the total variance in our simulations.

This procedure of replacing $\langle S|\boldsymbol{n},\alpha,d\rangle$ with the asymptotic expression $\langle S|\mathcal{K},\alpha,d\rangle_a$ in Eq. (4.16) leads to a new estimator of entropy, which we call *approximate PYM* estimator, or aPYM. This estimator is fully determined by just few data statistics, N, S_0, K_1, K_2, and Q_1. There are also two limiting cases of this estimator. First, by taking $d \to 0$ in Eqs. (4.20, 4.22), we define the approximate version of the NSB limit of the PYM estimator on a

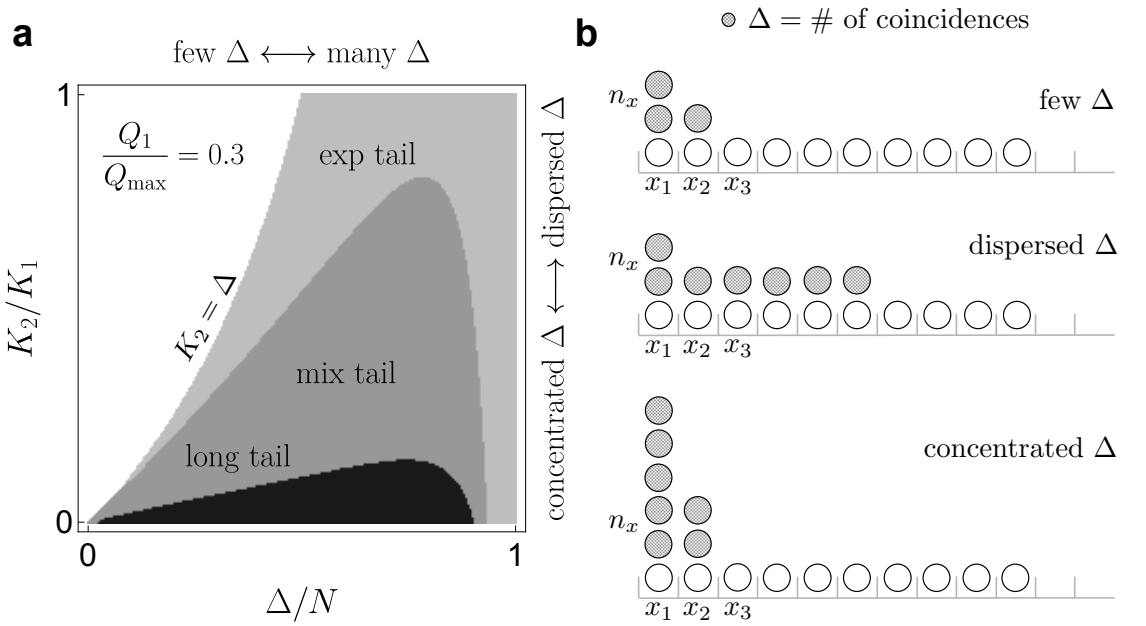

Figure 4.3: **a**: Phase diagram of the dominant tail hypothesis selected by the PYM estimator as a function of various statistics of the data sample. The explored statistics are the fraction of coincidences in the sample, Δ/N, and dispersion of the coincidences, K_2/K_1. This diagram is evaluated at the third crucial data statistics set at $Q_1 = 0.3\,Q_{\text{max}} = 0.3(\Delta - K_2)/2$. **b**: Schematic diagram that illustrates how sample sets with different Δ, K_1, and K_2 may look like. An empty or gray circle above a state x_i represent a single sample for that state. Gray circles denote coincidences.

countably infinite number of possible outcomes, which we denote as aNSB. At the other extreme, taking $\alpha \to 0$ in Eqs. (4.20, 4.22), corresponds to a prior that favors distributions with long tails. We denote the corresponding estimator as $\hat{S}_{\text{long}}$.

The above observation that, in the undersampled regime where $\exp(S/2) < N < \exp(S)$, the PYM entropy estimator and its relatives are determined approximately by just few statistics of the data, $\{N, S_0, K_1, K_2, Q_1\}$, is the main result of our paper. To corroborate this, we explore the quality of the approximation numerically for different distributions $\boldsymbol{q}$. Figure 4.2 presents results for three distributions with different structures of tails, generated from the Pitman-Yor Process: a distribution with an exponential tail (Fig. 4.2**a**, $\text{PYP}(d = 0, \alpha = 400) = \text{DP}(400)$), one with a mixed tail (Fig. 4.2**b**: $\text{PYP}(d = 0.4, \alpha = 100)$), and one with a long tail (Fig. 4.2**c**: $\text{PYP}(d = 0.6, \alpha = 0)$). In the lower panels we show the results of estimating entropy for different dataset sizes

using the ML estimator, the PYM estimator, the NSB estimator with a large alphabet size $\mathcal{A} = 20K_1$, and the three approximations: aPYM, aNSB, and $\hat{S}_{\text{long}}$. All results are averaged over ten sets of random samples. In all cases, the differences between NSB and aNSB on the one hand, and PYM and aPYM on the other are negligible, supporting the accuracy of the approximation. All four of these estimators produce high quality estimates for all sample sizes. Further, we also checked that the approximation of the posterior error of the estimators is close to that of the full versions (not shown). In contrast, $\hat{S}_{\text{long}}$ only performs well when the distribution has a long tail, and the Maximum Likelihood never works well.

4.5 Tail-hypothesis and entropy estimation phase diagrams

The above discussion shows that the PYM estimator and its relatives work by first estimating the most likely α and d from the sampled data, and then using these estimated parameters to approximate the structure of the low probability tail (from short, to long) and hence of its contributions to the entropy. We further showed that, in the regime of interest, the log-likelihood of α and d is dominated by just few statistics: N, S_0, K_1, K_2, and Q_1. It is thus illustrative to understand, which combinations of these statistics select which hypothesis on the structure of the tail. Building the corresponding phase diagram of the selected tail structure as a function of the data statistics is the goal of this Section.

We will consider three classes of tails: exponential ($d = 0$ selected, denoted as hypothesis $H = 1$), long tail ($\alpha = 0$ selected, denoted as hypothesis $H = 2$), and a mixed tails (arbitrary α and d, denoted as $H = 3$). Our goal is then to evaluate which of the three tail hypotheses has a higher probability given the data. Long and short tail hypotheses have one parameter each, while the mixed tail hypothesis has two parameters and contains the other two hypotheses as special cases. Thus when evaluating the log-likelihoods of each of the hypotheses, we must penalize them for having a different number of parameters, which

we do using Bayesian Information Criterion [111]. To do this, we evaluate the likelihoods

$$\mathcal{L}_H = \log p(\mathcal{K}|\hat{\alpha}, \hat{d}) + \log p_{\text{prior}}(\hat{\alpha}, \hat{d}) - \frac{n_H}{2} \log N, \tag{4.23}$$

where $\hat{\alpha}$ and $\hat{d}$ are the maximum likelihood values of the parameters within each hypothesis, and n_H is the number of parameters for the hypothesis ($n_H = 2$ for $H = 3$, and $n_H = 1$ otherwise). We remind the reader that, by construction, $\hat{\alpha} = 0$ for the long tail hypothesis, $H = 2$, and $\hat{d} = 0$ for the short tailed hypothesis, $H = 1$.

We determine the regions of the N, S_0, K_1, K_2, Q_1 space, where one of the three $\mathcal{L}_H$ dominates, and plot the slice of this phase diagram in Fig. 4.3. Specifically, in the Figure, we vary the total number of *coincidences*, $\Delta = N - K_1$, and the number of *states with coincidences*, that is, the number of states with more than two counts, K_2. By sampling many distributions, we empirically observe that the value $Q_1 \sim 0.6(\Delta - K_2)/2$ is when the rest of the $\Delta - K_2$ counts are uniformly dispersed, and Q_1 tends to zero when the rest of the counts are concentrated in a single state. Note that the maximum value Q_1 can take is $Q_{\text{max}} = \frac{\Delta - K_2}{2}$. For this reason, we choose the intermediate representative value $Q_1 = 0.3 Q_{\max} = 0.3 \frac{\Delta - K_2}{2}$.

To simplify the presentation, we plot the winning tail hypothesis as a function of Δ/N and K_2/K_1. Normalized in this way, the diagram is constrained to a square of size 1, as $0 \leq \Delta/N, K_2/K_1 \leq 1$. In addition, $K_2 \leq \Delta$, which means that the upper left corner is not accessible. The ratio Δ/N determines how common are the coincidences, and the ratio K_2/K_1 describes whether the coincidences in the data are concentrates in a few states, or dispersed over many states (see Figure 4.3b).

Figure 4.3**a** show that the exponential tail hypothesis dominates when there are many coincidences, $\Delta/N \sim 1$, or when the coincidences are dispersed, that is $K_2/K_1 \sim 1$ or $K_2/\Delta \sim 1$. Both cases can be explained as corresponding to distributions that are relatively uniform on some fixed number of states, and have zero probability elsewhere. A long tail only dominates when the fraction of coincidences has an intermediate value, but the coincidences are highly concentrated, $K_2/K_1 \ll 1$. In other words, in this case, there

are dominant states, but a lot of samples still fall outside of them. For other values of Δ/N and K_2/K_1, the mixed tail hypothesis dominates.

Equipped with this picture of which tail hypothesis is selected by the PYM estimator as a function of data statistics, we now can calculate how the estimator corrects the ML entropy value S_0 for different data statistics. Integrating the mean posterior entropy $\langle S|\mathcal{K}, \alpha, d\rangle_a$, Eq. (4.22), over our approximation of the posterior, $p_a(\alpha, d|\mathcal{K})$, which we obtain by exponentiating Eq. (4.20), we get the approximate PYM estimator $\hat{S}_{PYM,a}$. The Maximum Likelihood estimate S_0 enters linearly in the posterior mean entropy, Eq. (4.22). Thus we write

$$\langle S|\mathcal{K}, \alpha, d\rangle_a = b_{\alpha,d}\, S_0 + \delta S_{\alpha,d}, \tag{4.24}$$

where $b_{\alpha,d}$ and $\delta S_{\alpha,d}$ can be read off from Eq. (4.22). Performing the integral over the approximate posterior, this becomes:

$$\hat{S} = \delta S + b\, S_0, \tag{4.25}$$

where δS and b are averages of the corresponding α- and d-dependent quantities. Thus *independent* of the Maximum Likelihood entropy value, within our approximation, the PYM estimator obtains the entropy estimate by decreasing the ML contribution from the well-sampled head of the distribution and adding an offset that comes from the low probability tail. This is similar to so-called partition-based entropy estimators, [18, 81, 83, 113], which divide the state space into sub-spaces, estimate entropy in each sub-space, and then add the estimates weighted by the probability of being in a corresponding sub-space. However, here this partitioning arises naturally from the Bayesian framework within our approximations.

Both the scale factor and the offset depend on the dominant α and d contributing to the estimator, and hence on the usual statistics of the data, Δ, K_1, K_2, and Q_1. Specifically, we numerically observe that the value of b obtained from Eq. (4.25) satisfies

$$b = \langle N/(\alpha + N)\rangle \leq 1, \tag{4.26}$$

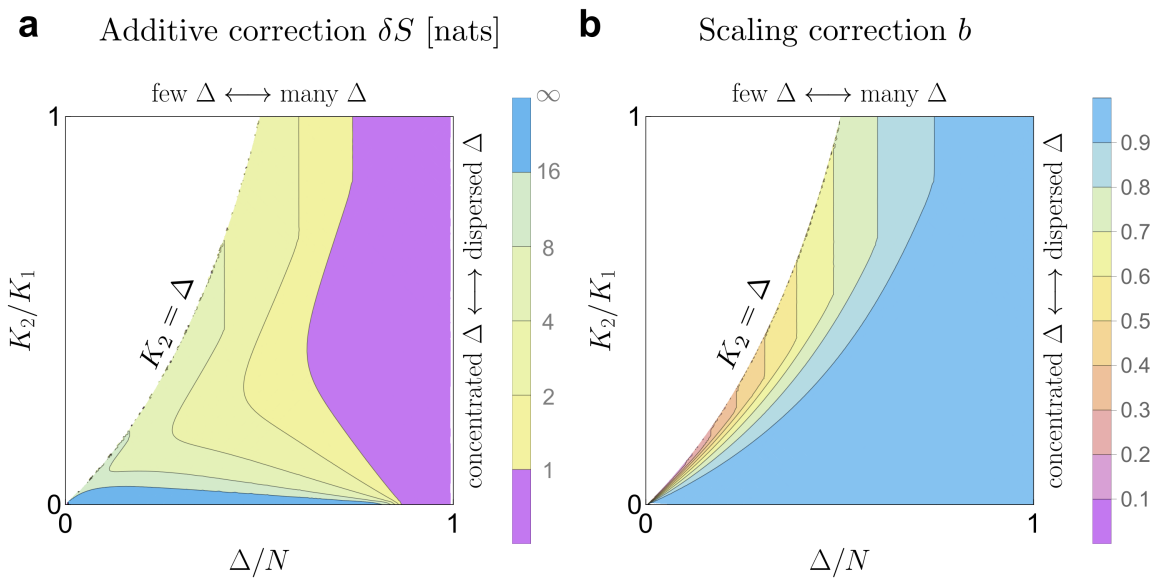

Figure 4.4: Corrections to entropy estimation as a function of determining data statistics. We break down the final estimation for entropy in two parts, as $\hat{S} = \delta S(\Delta/N, K_2/K_1) + b(\Delta/N, K_2/K_1)\, S_0$, where δS is the additive correction and b is scaling factor or weight for the Maximum Likelihood estimate. Well-sampled distributions are located in the upper-right corner where $\delta S = 0$ and $b = 1$. As in the previous plots, we leave $Q_1 = 0.3\,(\Delta - K_2)/2$. **a**: Additive correction to entropy. **b**: Scaling correction to entropy.

where the average is over the product of the approximate posterior obtained by exponentiating Eq. (4.20) and the prior $p(\gamma) = e^{-7\gamma/100}$ with γ defined in Eq. (4.15). Note that α is a measure of how much probability is concentrated in the tail. Thus the ratio $N/(\alpha + N)$ approximates the overall weight of the the well-sampled head of the distribution, requiring to decrease the contribution to the entropy from the head by this factor. This matches our assertion that the aPYM estimator is a partition-based estimator, separating the head from the tail.

In Figure 4.4 we show results of numerical estimation of the offset δS and the scaling factor b as a function of the fraction of coincidences, Δ/N, and the dispersion of coincidences, K_2/K_1. As in the previous case, we keep $Q_1 = 0.3Q_{\max}$. We also set $N = 10^4$. Figure 4.4(a) shows that the additive term grows when the fraction of coincidences Δ/N decreases, and when K_2/K_1 is small, so that coincidences are concentrated. Both of these cases correspond to a lot of mass in the tail (see corresponding long tail region in Fig-

ure 4.3(a). The largest values of δS occur along the boundary strip $(\Delta/N, K_2/K_1 \ll 1)$ and the boundary $K_2 = \Delta$. Panel **b** shows that the scaling factor b is close to 1 in most areas, except near the boundary edge $K_2 = \Delta$. Along this boundary, the scaling factor becomes the largest when the number of coincidences decreases, $\Delta/N \ll 1$. Figure 4.4 clearly highlights when Bayesian corrections to the ML estimation of entropy are essential: regions of few and concentrated coincidences.

4.6 Discussion

The major finding of this work is an excellent approximation for the PYM estimator, one of the best Bayesian entropy estimators, and its various relatives (such as NSB). The approximation simplifies the numerics considerably. Crucially, the approximation also shows that the output of the PYM entropy estimator depends on just a few statistics of the data, namely the maximum likelihood (ML) entropy estimate, the fraction of coincidences Δ/N, and the dispersion of coincidences K_1/K_2, and Q_1. We showed that that workflow of the estimator can be interpreted as first estimating the parameters d and α based on the aforementioned statistics, and with them the tail structure and the total weight of the tail. Then the estimator rescales the ML entropy estimate by the weight of the well-sampled head of the distribution, and adds to it the estimated entropy of the tail. The phase diagrams of which tail structure the estimator selects, Fig. 4.3, and how it corrects the ML estimate, Fig. 4.4, illustrate these points.

Early work of Ma [69] showed that when states are equiprobable, in the under-sampled regime, the coincidences in counts can help with the inference of the entropy of a system. Later Nemenman [80] showed that in the severely under-sampled regime (K_1 close to N), entropy estimation depends on the number of coincidences K_1. Further, he pointed out how reliable entropy estimates may be obtained by partitioning the overall state space of the variable into sub-spaces with similar sampling properties [83]. Here we extend these results to the whole regime where entropy estimation is challenging for multinomial ob-

servations, $\exp(S/2) < N < \exp(S)$, by approximating the more general PYM estimator. Our identification of the small set of statistics, which define the output of the estimator, lifts the veil from its inner workings, allowing for a simple, semi-analytical estimation procedure. In particular, this allows us to predict if a particular estimator will be biased simply by looking at the values of the select statistics of the data.

How to match *a priori* assumptions about the underlying distributions to the data to allow for an unbiased estimation of quantities of interest—such as entropy [6, 81] or the mutual information [46]— is an open problem [47]. It requires understanding the relation between the *a priori* assumptions and the data features that influence the inference. In this work, we build such a link for entropy estimation, and we hope that similar links might exist for other difficult estimation problems.

4.7 Appendix

4.7.1 Marginal likelihood approximation for a Pitman-Yor process

In this Appendix we show how to approximate the marginal posterior of a Pitman-Yor process in the regime $K_1 \lesssim N \leq \exp(S)$. We start by manipulating each term in the logarithm of the evidence $\mathcal{L} = \log p(\mathbf{n}|\alpha, d)$ from Eq. (4.17),

$$\mathcal{L}(\mathbf{n}|\alpha, d) = \sum_{l=1}^{K_1-1} \log(\alpha + ld) + \sum_{i=1}^{K_1} \log\Gamma(n_i - d) - K_1 \log\Gamma(1-d) + \log\Gamma(1+\alpha) - \log\Gamma(N+\alpha). \quad (4.27)$$

To simplify the first term in Eq. (4.27), we rewrite it in terms of coincidences K_1 as

follows:

$$I_1 = \sum_{l=1}^{K_1-1} \log(\alpha + ld) \tag{4.28}$$

$$= \sum_{l=1}^{K_1-1} \left[\log d + \log\left(\frac{\alpha}{d} + l\right)\right] \tag{4.29}$$

$$= (K_1 - 1)\log d + \sum_{l=1}^{K_1-1} \left[\log\Gamma\left(\frac{\alpha}{d} + l + 1\right) - \log\Gamma\left(\frac{\alpha}{d} + l\right)\right] \tag{4.30}$$

$$= (K_1 - 1)\log d + \log\Gamma\left(\frac{\alpha}{d} + K_1\right) - \log\Gamma\left(\frac{\alpha}{d} + 1\right). \tag{4.31}$$

In order to rewrite the rest of the terms of Eq. (4.27) in terms of various coincidence statistics, we use the identity Eq. (4.19). Joining the second and third terms in Eq. (4.27) and rewriting them in terms of count multiplicities yields

$$\begin{aligned}
&\sum_{i=1}^{K_1} \log\Gamma(n_i - d) - K_1 \log\Gamma(1 - d) \\
&= -K_2 \log\Gamma(1 - d) + \sum_{m=2} (K_m - K_{m+1}) \log\Gamma(m - d) \\
&= \sum_{m=2} K_m \left[\log\Gamma(m - d) - \log\Gamma(m - 1 - d)\right] \\
&= \sum_{m=2} K_m \log(m - 1 - d) \\
&= K_2 \log(1 - d) + Q(d),
\end{aligned} \tag{4.32}$$

where

$$Q(d) = \sum_{m=3}^{m_f} K_m \log(m - 1 - d). \tag{4.33}$$

where m_f denotes the largest occupancy of any state in the sample. Since the domain of $0 \leq d < 1$ is small, $Q(d)$ is approximately linearly varying with d, so that we can expand it around $d = 0$:

$$\begin{aligned}
Q(d) &= Q(0) - \sum_{j=1} \left[\sum_{m=3} \frac{K_m}{(m-1)^j}\right] \frac{d^j}{j} \\
&\approx Q(0) - \left[\sum_{m=3} \frac{K_m}{m-1}\right] d - \frac{1}{2}\left[\sum_{m=3} \frac{K_m}{(m-1)^2}\right] d^2 + \mathcal{O}(Q_3), \\
&= Q_0 - Q_1 d - \frac{1}{2} Q_2 d^2 + \mathcal{O}(Q_3).
\end{aligned} \tag{4.34}$$

where

$$Q_j = \sum_{m=3} \frac{K_m}{(m-1)^j} \tag{4.35}$$

for $j \geq 1$. As d approaches 1, the term $K_2 \log(1-d)$ goes to infinity, which renders any error in the Taylor expansion of $Q(d)$ irrelevant. This makes the approximations above useable even if we ignore $\mathcal{O}(d^2)$ terms.

Putting all of the approximations above together, the ensuing approximate logarithm of the evidence $\mathcal{L}(\mathbf{n}|\alpha, d)$ is

$$\begin{aligned}\mathcal{L}(\mathbf{n}|\alpha, d) \approx (K_1 - 1)\log d + \log\Gamma\left(\frac{\alpha}{d} + K_1\right) - \log\Gamma\left(\frac{\alpha}{d} + 1\right) + \log\Gamma(1+\alpha) \\ - \log\Gamma(N+\alpha) + K_2 \log(1-d) - Q_1 d + \mathcal{O}\left(d^2 \sum_{m=3} \frac{K_m}{(m-1)^2}\right),\end{aligned} \tag{4.36}$$

up to an additive constant. This is Eq. (4.20) in the main text.

4.7.2 Maximum likelihood Entropy in terms of coincidences

To relate the conditional entropy, Eq. (4.13), to the Maximum Likelihood entropy estimator S_0, we need to rewrite the latter in terms coincidences. Utilizing the identity Eq. (4.19), we write

$$N\left[S_0 - \log N\right] = -\sum_i n_i \log n_i \tag{4.37}$$

$$= -\sum_{m=2} (K_m - K_{m+1}) m \log m \tag{4.38}$$

$$= -K_2(2\log 2) - \sum_{m=3} K_m [m\log m - (m-1)\log(m-1)]. \tag{4.39}$$

Rewriting the expression in brackets as

$$m\log m - (m-1)\log(m-1) = 1 + \psi(m) + \mathcal{O}(m^{-2}). \tag{4.40}$$

and plugging this into Eq. (4.39), we finally obtain,

$$N\left[S_0 - \log N\right] = -K_2 \log 4 - (N - K_1 - K_2) - \sum_{m=3} K_m \psi(m) + \mathcal{O}\left(\sum_m K_m/m^2\right). \tag{4.41}$$

4.7.3 Mean posterior entropy approximation for the Pitman-Yor Process

Similar to Appendix 4.7.1, here we approximate the posterior entropy, Eq. (4.13), in the limit of small d. To simplify the notation, we use the shorthand $S = \langle S|\boldsymbol{n}, \alpha, d\rangle$ in this Appendix. Rearranging Eq. (4.13), we obtain

$$(\alpha + N)\,[S - \psi(N + \alpha + 1)] = \\ - \alpha\,\psi(1-d) - K_1\,d\,\psi(1-d) - \sum_i (n_i - d)\psi(n_i + 1 - d). \quad (4.42)$$

We now again use Eq. (4.19) and a Taylor expansion in small d to rewrite the last term on the right hand side of Eq. (4.42):

$$\begin{aligned} & K_1\,d\,\psi(1-d) - \sum_i (n_i - d)\psi(n_i + 1 - d) \\ &= K_1\,d\,\psi(1-d) - \sum_{m=1} (K_m - K_{m+1})(m-d)\psi(m+1-d) \\ &= -\sum_{m=1} K_m\,[(m-d)\psi(m+1-d) - (m-1-d)\psi(m-d)] \\ &= -\sum_{m=1} K_m\,[1 + \psi(m-d)] \\ &= -\sum_{m=1} K_m - \sum_{m=1} K_m\psi(m-d) \\ &= -N - K_1\psi(1-d) - K_2\psi(2-d) - \sum_{m=3} K_m\psi(m-d). \end{aligned} \quad (4.43)$$

where we used $\psi(m+1-d) = \left(\psi(m-d) + \frac{1}{m-d}\right)$.

Since $m \geq 3$, we can Taylor expand the sum in this last term around $d = 0$ to obtain

$$\sum_{m=3} K_m\psi(m-d) \approx \sum_{m=3} K_m\psi(m) + d\sum_{m=3} K_m\psi'(m) + \mathcal{O}(d^2 \sum_m K_m\psi''(m)). \quad (4.44)$$

Now using the relations $\psi'(m) = \frac{1}{m-1} + \mathcal{O}(m^{-2})$ and the expression for $\sum_{m=3} K_m\psi(m)$ in Eq. (4.41), we rewrite Eq. (4.44) as

$$\sum_{m=3} K_m\psi(m-d) \approx K_2 \log 4 + (N - K_1 - K_2) - N\,[S_0 - \log N] \\ + d\sum_{m=3} \frac{K_m}{m-1} + \mathcal{O}(d^2, \sum_{m=3} K_m/m^2), \quad (4.45)$$

where $\mathcal{O}(d^2, \sum_{m=3} K_m/m^2)$ means that we kept terms that are at most linear in d and whose summands are at most proportional to $\sum_{m=3} K_m/m$. Plugging these approximation in Eq. (4.43) and noticing that $Q_1 = \sum_{m=3} \frac{K_m}{m-1}$, we obtain

$$(\alpha + N)\,[S - \psi(N + \alpha + 1)] = N(S_0 - \log N) - \alpha\,\psi(1 - d) + K_1\,[-1 - \psi(1 - d)] + K_2\,[-1 - \psi(2 - d) + \log 4] - Q_1\,d + \mathcal{O}(d^2, \sum_{m=3} K_m/m^2), \quad (4.46)$$

which after isolating S becomes Eq. (4.22) of the main text.

Chapter 5 Conclusion

In the previous chapters, we discovered new physics from our analysis of biology and data. We learned a new universality class of surface growth that is inspired by the behavior of *Dictyostelium discoideum* cells [103]. By analyzing experiments of thermal learning in *C. elegans*, we learned a new model of conditioning. This led to the understanding of a new mechanism that explains previously unexplained phenomena in associative learning [84, 90, 98]. By analyzing Bayesian entropy estimators, we found that few emergent data statistics reflect the tail hypothesises assumed by the estimators. In all of these cases, new physics emerged from the analysis. Here we discuss possible extensions of the work done in this thesis.

5.1 Surface Growth

In Chapter 2, we discussed a biologically inspired extension of the ballistic deposition model. We found a new universality class that requires three dynamical exponents. It remains unclear how general this new class is and how relevant it is to the biological systems that inspired it. One way to test the generality of this class is by conducting experiments that probe large spatiotemporal scales that are beyond the typical scales studied in current experiments to test our predicted exponents. Experimental cultures on the length scale of meters and time scale of many days would be needed to observe and measure the relevant exponents. Another way is to write down a system of stochastic partial differential equations that describe the growth phenomena observed in our simulations. Phenomena that can be described by the ensuing equations should fall into the same universality class. We also expect many other universality classes to come about as a consequence of introducing internal states to the constituents that make up the interface. We hope that our new rule will

inspire many growth rules and universality classes.

5.2 Animal Learning

The prevailing models of the thermotactic behavior of *C. elegans* are models of exponential decay of the thermal preference of the worm to the cultivation temperature. These models assume that the worms possess a preference for a single temperature: the cultivation temperature. The more complex model we developed in Chapter 3, was able to explain behavioral phenomena such as the spontaneous recovery of associations that were extinct. It was also able to fit the behavior of four-hour long thermotactic index trajectories, which previous models could not do. We expect our model to describe learning in many other animals as well. However, all of these models described above preclude the possibility of the worm possessing multiple temperature preferences at the same time.

In the wild, animals encounter multiple environmental cues that animals associate with predators. The variety of these environmental cues selects for animals that are able to recognize and remember multiple cues; animals that do not recognize their predator's cues are destined for extinction. The ability to store and recognize multiple cues is therefore essential to the survival and fitness of animals. While higher organisms do possess the ability to form multiple associations, it would be intriguing to observe it in worms with brains consisting of only 302 neurons. It would also lead to novel and realistic models of learning in animals.

Worms do not live in an environment where food and temperature are the only stimuli. Worms in their natural environments interact with other worms. Intuitively, there should be many advantages to being in a group instead of living alone. In the context of learning, is there a learning advantage to being in a group? Can two worms that are reared at different temperatures with food influence each other's behavior on a thermal gradient? Can a learned worm that associates certain temperatures with punishment (heat from laser zap) teach a novice worm to avoid such temperatures, thereby accelerating the learning pro-

cess? Can a learned rat that knows how to reach the cheese in a complex maze lower the time it takes another rat to reach the cheese? Through experiments, quantitative answers to these questions will spur theoretical advances that could uncover new physics that explains how an individual learner is fundamentally different from collective learners. This could have an impact on our understanding of how *C. elegans* interacts with other species on an ecological level [64] and how learning influences ecology in general [79, 99–101]. It will also shed light on how animals at large learn from each other instead of learning from the environment and the communication channels through which the teaching occurs.

5.3 Data statistics

In Chapter 4, we showed that a well-known Bayesian entropy estimator, the PYM estimator, relies on a few data statistics to correct the Maximum Likelihood estimate in the undersampled regime. It would be interesting to repeat the analysis of Chapter 4 for more Bayesian entropy estimators that allow for a larger class of tail hypothesis. This would clarify if tail hypothesises are generally summarized by a few data statistics or is that just a property of a certain class of tail hypothesises? Another useful extension of this work would be to estimate mutual information between two random variables where their joint distribution is undersampled. This is a likely scenario in modern data analysis and may provide new methods for estimating mutual information.

Another possible extension is to calculate analytically the PYM entropy estimate and write it in terms of the data statistics mentioned in the chapter. This can be done by using asymptotic techniques to approximate the integral of the product of the expected entropy 4.22, the posterior (the exponential of Eq. 4.20), and the mixing prior Eq. 4.14. The resulting formula would make estimating entropy for multinomial observations devoid of any numerical calculation of quadrature.

Entropy is not special in the sense that it is but one property of a distribution. The techniques used in our work can be applied to any other property of a distribution. It would

be interesting to see if estimating moments of an undersampled discrete distribution also relied on a few data statistics similar to what we found in entropy estimation.

Bibliography

1. Abramowitz, M. & Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* ninth Dover printing, tenth GPO printing (Dover, New York City, 1964).
2. Amano, H. & Maruyama, I. N. Aversive olfactory learning and associative long-term memory in Caenorhabditis elegans. *Learn Mem* **18,** 654–65 (2011).
3. Amaral, L. A. N., Barabási, A. L., Buldyrev, S. V., Havlin, S. & Stanley, H. E. New exponent characterizing the effect of evaporation on imbibition experiments. *Physical Review Letters* **72,** 641–644 (1994).
4. Antos, A. & Kontoyiannis, I. *Estimating the entropy of discrete distributions* in *Proceedings. 2001 IEEE International Symposium on Information Theory (IEEE Cat. No.01CH37252)* (), 45.
5. Archer, E., Park, I. M. & Pillow, J. W. Bayesian and quasi-Bayesian estimators for mutual information from discrete data. *Entropy* **15,** 1738–1755 (2013).
6. Archer, E., Park, I. M. & Pillow, J. W. Bayesian entropy estimation for countable discrete distributions. *The Journal of Machine Learning Research* **15,** 2833–2868 (2014).
7. Ardiel, E. L. & Rankin, C. H. An elegant mind: Learning and memory in Caenorhabditis elegans. en. *Learning & Memory* **17,** 191–201 (01/2010).
8. Bachman, P., Hjelm, R. D. & Buchwalter, W. *Learning Representations by Maximizing Mutual Information Across Views* in *Advances in Neural Information Processing Systems* (eds Wallach, H. *et al.*) **32** (Curran Associates, Inc., 2019).
9. Barabási, A. L. & Stanley, H. E. *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995).
10. Bartumeus, F., Luz, M. G. E. d., Viswanathan, G. M. & Catalan, J. Animal search strategies: a quantitative random-walk analysis. *Ecology* **86,** 3078–3087. ISSN: 1939-9170 (2005).
11. Berry II, M. J., Tkačik, G., Dubuis, J., Marre, O. & da Silveira, R. A. A simple method for estimating the entropy of neural activity. *Journal of Statistical Mechanics: Theory and Experiment* **2013,** P03015 (2013).
12. Biron, D. *et al.* A diacylglycerol kinase modulates long-term thermotactic behavioral plasticity in C. elegans. *Nat Neurosci* **9,** 1499–505 (2006).
13. Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77,** 71–94 (05/1974).
14. Byrne Rodgers, J. & Ryu, W. S. Targeted thermal stimulation and high-content phenotyping reveal that the C. elegans escape response integrates current behavioral state and past experience. *PLOS ONE* **15,** 1–22 (03/2020).

15. Cavagna, A. & Giardina, I. Bird Flocks as Condensed Matter. *Annual Review of Condensed Matter Physics* **5,** 183–207 (2014).

16. Cerquetti, A. Exact Good-Turing characterization of the two-parameter Poisson-Dirichlet superpopulation model. *arXiv preprint arXiv:1901.09665* (2019).

17. Chao, A. & Shen, T.-J. Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environmental and ecological statistics* **10,** 429–443 (2003).

18. Chao, A., Wang, Y. & Jost, L. Entropy and the species accumulation curve: a novel entropy estimator via discovery rates of new species. *Methods in Ecology and Evolution* **4,** 1091–1100 (2013).

19. Chechik, G., Globerson, A., Tishby, N. & Weiss, Y. Information Bottleneck for Gaussian Variables. *Journal of Machine Learning Research* **6,** 165–188 (2005).

20. Chen, Z. *et al.* Two insulin-like peptides antagonistically regulate aversive olfactory learning in C. elegans. *Neuron* **77,** 572–85 (2013).

21. Chi, C. A. *et al.* Temperature and food mediate long-term thermotactic behavioral plasticity by association-independent mechanisms in C. elegans. *J Exp Biol* **210,** 4043–52 (2007).

22. Chi, C. A. *et al.* Temperature and food mediate long-term thermotactic behavioral plasticity by association-independent mechanisms in C. elegans. eng. *J Exp Biol* **210,** 4043–4052 (11/2007).

23. Clark, D. A., Biron, D., Sengupta, P. & Samuel, A. D. The AFD sensory neurons encode multiple functions underlying thermotactic behavior in Caenorhabditis elegans. *J Neurosci* **26,** 7444–51 (2006).

24. Cover, T. M. & Thomas, J. A. *Elements of information theory* (John Wiley & Sons, 2012).

25. Daniels, B. C. & Nemenman, I. Automated adaptive inference of phenomenological dynamical models. *Nature Communications* **6,** 8133 (2015).

26. Daniels, B. C., Ryu, W. S. & Nemenman, I. Automated, predictive, and interpretable inference of Caenorhabditis elegans escape dynamics. *Proc Natl Acad Sci (USA)* **116,** 7226–7231. ISSN: 0027-8424 (2019).

27. Darling, D. The influence of the maximum term in the addition of independent random variables. *Transactions of the American Mathematical Society* **73,** 95–107 (1952).

28. Das Sarma, S. & Tamborenea, P. A new universality class for kinetic growth: One-dimensional molecular-beam epitaxy. *Physical Review Letters* **66,** 325–328 (1991).

29. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neurosci* **8,** 1704–1711 (12/2005).

30. Dayan, P. in *Steven's Handbook of Experimental Psychology* (ed Gallistel, C. R.) 1–35 (John Wiley and Sons, New York, NY, 06/2001).

31. Dayan, P. & Niv, Y. Reinforcement learning: the good, the bad and the ugly. *Curr Opin Neurobiol* **18,** 185–196 (04/2008).

32. De Houwer, J., Thomas, S. & Baeyens, F. Associative learning of likes and dislikes: a review of 25 years of research on human evaluative conditioning. *Psychol Bull* **127,** 853–69 (2001).

33. Dunsmoor, J. E., Niv, Y., Daw, N. & Phelps, E. A. Rethinking Extinction. *Neuron* **88,** 47–63 (10/2015).

34. Eden, M. *A two-dimensional growth process* in *Proc Fourth Berkeley Symp Mathematics, Statistics, and Probability* **4** (Berkeley: UC Press, 1961), 223–239.

35. Evertsz, C. Self-affine nature of dielectric-breakdown model clusters in a cylinder. *Phys Rev A* **41,** 1830–1842 (1990).

36. Faghihi, F., Moustafa, A. A., Heinrich, R. & Wörgötter, F. A computational model of conditioning inspired by Drosophila olfactory system. *Neural Netw* **87,** 96–108 (2017).

37. Family, F. Scaling of rough surfaces: effects of surface diffusion. *Journal of Physics A: Mathematical and General* **19,** L441–L446 (1986).

38. Family, F. & Vicsek, T. Scaling of the active zone in the Eden process on percolation networks and the ballistic deposition model. *Journal of Physics A: Mathematical and General* **18,** L75–L81 (1985).

39. Ferguson, T. S. A Bayesian Analysis of Some Nonparametric Problems. *The Annals of Statistics* **1,** 209–230. ISSN: 00905364 (1973).

40. François-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G. & Pineau, J. An Introduction to Deep Reinforcement Learning. *Found Trends Machine Learning* **11,** 219–354 (2018).

41. Gems, D. *et al.* Two pleiotropic classes of daf-2 mutation affect larval arrest, adult behavior, reproduction and longevity in Caenorhabditis elegans. eng. *Genetics* **150,** 129–155 (09/1998).

42. Goodman, M. B. & Sengupta, P. How Caenorhabditis elegans Senses Mechanical Stress, Temperature, and Other Physical Stimuli. *Genetics* **212,** 25–51 (05/2019).

43. Grassberger, P. Entropy estimates from insufficient samplings. *arXiv preprint physics/0307138* (2003).

44. Hawk, J. D. *et al.* Integration of Plasticity Mechanisms within a Single Sensory Neuron of C. elegans Actuates a Memory. English. *Neuron* **97,** 356–367.e4 (01/2018).

45. Hedgecock, E. M. & Russell, R. L. Normal and mutant thermotaxis in the nematode Caenorhabditis elegans. *Proc Natl Acad Sci (USA)* **72,** 4061–5 (1975).

46. Hernández, D. G. & Samengo, I. Estimating the Mutual Information between Two Discrete, Asymmetric Variables with Limited Samples. *Entropy* **21,** 623 (2019).

47. Hernández, D. G. & Samengo, I. Inferring a property of a large system from a small number of samples. *Entropy* **24,** 125 (2022).

48. Hernández, D. G., Roman, A. & Nemenman, I. *Low probability states, data statistics, and entropy estimation* 2022.

49. Hesterberg, T. Bootstrap. *Wiley Interdisciplinary Reviews: Computational Statistics* **3,** 497–526. ISSN: 1939-5108 (2011).

50. Hobert, O. Behavioral plasticity in C. elegans: paradigms, circuits, genes. eng. *Journal of Neurobiology* **54,** 203–223 (01/2003).

51. Holmes, C. M. & Nemenman, I. Estimation of mutual information for real-valued data with error bars and controlled bias. *Phys. Rev. E* **100,** 022404 (2 08/2019).

52. Ishwaran, H. & James, L. F. GENERALIZED WEIGHTED CHINESE RESTAURANT PROCESSES FOR SPECIES SAMPLING MIXTURE MODELS. *Statistica Sinica* **13,** 1211–1235. ISSN: 10170405, 19968507 (2003).

53. Ishwaran, H. & James, L. F. Gibbs Sampling Methods for Stick-Breaking Priors. *Journal of the American Statistical Association* **96,** 161–173 (2001).

54. Jurado, P., Kodama, E., Tanizawa, Y. & Mori, I. Distinct thermal migration behaviors in response to different thermal gradients in Caenorhabditis elegans. *Genes Brain Behav* **9,** 120–7 (2010).

55. Kardar, M., Parisi, G. & Zhang, Y.-C. Dynamic Scaling of Growing Interfaces. *Physical Review Letters* **56,** 889–892 (1986).

56. Kimata, T., Sasakura, H., Ohnishi, N., Nishio, N. & Mori, I. Thermotaxis of C. elegans as a model for temperature perception, neural information processing and neural plasticity. *Worm* **1,** 31–41 (2012).

57. Kimura, K. D., Miyawaki, A., Matsumoto, K. & Mori, I. The C. elegans thermosensory neuron AFD responds to warming. *Curr Biol* **14,** 1291–5 (2004).

58. Kodama, E. *et al.* Insulin-like signaling and the neural circuit for integrative behavior in C. elegans. *Genes Dev* **20,** 2955–60 (2006).

59. Kolchinsky, A. & Tracey, B. D. Estimating mixture entropy with pairwise distances. *Entropy* **19,** 361 (2017).

60. Krypotos, A.-M., Effting, M., Kindt, M. & Beckers, T. Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience* **9** (2015).

61. Lai, Z. W. & Das Sarma, S. Kinetic growth with surface relaxation: Continuum versus atomistic models. *Physical Review Letters* **66,** 2348–2351 (1991).

62. Lamperti, J. A contribution to renewal theory. *Proceedings of the American Mathematical Society* **12,** 724–731 (1961).

63. Lesne, A., Blanc, J.-L. & Pezard, L. Entropy estimation of very short symbolic sequences. *Phys. Rev. E* **79,** 046208 (4 04/2009).

64. Lev, I. & Zimmer, M. Predator–prey interactions: Strategic biting. *Current Biology* **32,** R367–R370 (2022).

65. Lin, C. H. A. *et al.* Insulin Signaling Plays a Dual Role in Caenorhabditis elegans Memory Acquisition and Memory Retrieval. *The Journal of Neuroscience* **30,** 8001–8011 (06/2010).

66. Lombardi, D. & Pant, S. Nonparametric k-nearest-neighbor entropy estimator. *Phys. Rev. E* **93,** 013310 (1 01/2016).

67. Luo, L., Clark, D. A., Biron, D., Mahadevan, L. & Samuel, A. D. Sensorimotor control during isothermal tracking in Caenorhabditis elegans. *J Exp Biol* **209,** 4652–62 (2006).

68. Luo, L. *et al.* Bidirectional thermotaxis in *Caenorhabditis elegans* is mediated by distinct sensorimotor strategies driven by the AFD thermosensory neurons. *Proc Natl Acad Sci (USA)* **111,** 2776–2781 (2014).

69. Ma, S.-k. Calculation of entropy from data of motion. *Journal of Statistical Physics* **26,** 221–240 (1981).

70. MacKay, D. J. C. *Information theory, inference, and learning algorithms* (Cambridge UP, 2003).

71. MacKay, D. J. Bayesian interpolation. *Neural computation* **4,** 415–447 (1992).

72. Martinez-Corral, R., Liu, J., Süel, G. M. & Garcia-Ojalvo, J. Bistable emergence of oscillations in growing *Bacillus subtilis* biofilms. *Proceedings of the National Academy of Sciences* **115,** E8333 (2018).

73. Meakin, P. & Family, F. Diverging length scales in diffusion-limited aggregation. *Phys Rev A* **34,** 2558–2560 (1986).

74. Medina, E., Hwa, T., Kardar, M. & Zhang, Y.-C. Burgers equation with correlated noise: Renormalization-group analysis and applications to directed polymers and interface growth. *Physical Review A* **39,** 3053–3075 (1989).

75. Miller, G. Note on the bias of information estimates. *Information theory in psychology: Problems and methods* (1955).

76. Minka, T. *Estimating a Dirichlet distribution* 2000.

77. Mohri, A. *et al.* Genetic Control of Temperature Preference in the Nematode Caenorhabditis elegans. *Genetics* **169,** 1437–1450 (03/2005).

78. Mori, I. & Ohshima, Y. Neural regulation of thermotaxis in Caenorhabditis elegans. *Nature* **376,** 344–8 (1995).

79. Mowlaei, S., Roman, A. & Pleimling, M. Spirals and coarsening patterns in the competition of many species: a complex Ginzburg–Landau approach. *Journal of Physics A: Mathematical and Theoretical* **47,** 165001 (2014).

80. Nemenman, I. Coincidences and estimation of entropies of random variables with large cardinalities. *Entropy* **13,** 2013–2023 (2011).

81. Nemenman, I., Bialek, W. & van Steveninck, R. d. R. Entropy and information in neural spike trains: Progress on the sampling problem. *Physical Review E* **69,** 056111 (2004).

82. Nemenman, I., Shafee, F. & Bialek, W. *Entropy and inference, revisited* in *Advances in neural information processing systems* (2002), 471–478.

83. Nemenman, I., Wall, M. E. & Strauss, C. E. Of fishes and birthdays: Efficient estimation of polymer configurational entropies. *arXiv preprint arXiv:1502.02364* (2015).

84. Niv, Y. Reinforcement learning in the brain. *J Math Psych* **53,** 139–154 (06/2009).

85. Nuttley, W. M., Atkinson-Leadbeater, K. P. & Kooy, D. v. d. Serotonin mediates food-odor associative learning in the nematode Caenorhabditis elegans. en. *Proc Natl Acad Sci (USA)* **99,** 12449–12454 (09/2002).

86. Ódor, G. Universality classes in nonequilibrium lattice systems. *Reviews of Modern Physics* **76,** 663–724 (2004).

87. Otto, A. R., Gershman, S. J., Markman, A. B. & Daw, N. D. The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychol Sci* **24,** 751–61 (2013).

88. Pamir, E., Szyszka, P., Scheiner, R. & Nawrot, M. P. Rapid learning dynamics in individual honeybees during classical conditioning. *Frontiers in Behavioral Neuroscience* **8** (2014).

89. Paninski, L. Estimation of entropy and mutual information. *Neural Computation* **15,** 1191–1253 (2003).

90. Pavlov, I. *Conditioned Reflexes* (Oxford University Press, 1927).

91. Pierce, S. B. *et al.* Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse C. elegans insulin gene family. eng. *Genes & Development* **15,** 672–686 (03/2001).

92. Pitman, J., Yor, M., *et al.* The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *The Annals of Probability* **25,** 855–900 (1997).

93. Prindle, A. *et al.* Ion channels enable electrical communication in bacterial communities. *Nature* **527,** 59–63 (2015).

94. Rácz, Z., Siegert, M., Liu, D. & Plischke, M. Scaling properties of driven interfaces: Symmetries, conservation laws, and the role of constraints. *Physical Review A* **43,** 5275–5283 (1991).

95. Ramot, D., MacInnis, B. L. & Goodman, M. B. Bidirectional temperature-sensing by a single thermosensory neuron in C. elegans. en. *Nature Neuroscience* **11,** 908–915 (08/2008).

96. Ramot, D., MacInnis, B. L., Lee, H.-C. & Goodman, M. B. Thermotaxis is a robust mechanism for thermoregulation in Caenorhabditis elegans nematodes. *J Neurosci* **28,** 12546–12557. ISSN: 1529-2401 (2008).

97. Rescorla, R. & Wagner, A. in *Classical conditioning II: Current research and theory* (eds Black, A. & Prokasy, W.) 64–99 (Appleton-Century-Crofts, New York, NY, 1972).

98. Roman, A., Palanski, K., Nemenman, I. & Ryu, W. *Thermal learning in C. elegans reveals a general multi-dimensional model of conditioning figshare* `https://figshare.com/articles/dataset/C_elegans_Data_zip/19907110`. 2022.

99. Roman, A., Dasgupta, D. & Pleimling, M. A theoretical approach to understand spatial organization in complex ecologies. *Journal of Theoretical Biology* **403,** 10–16 (2016).

100. Roman, A., Dasgupta, D. & Pleimling, M. Interplay between partnership formation and competition in generalized May-Leonard games. *Phys. Rev. E* **87,** 032148 (3 03/2013).

101. Roman, A., Konrad, D. & Pleimling, M. Cyclic competition of four species: domains and interfaces. *Journal of Statistical Mechanics: Theory and Experiment* **2012,** P07014 (2012).

102. Roman, A., Palanski, K., Nemenman, I. & Ryu, W. S. *Multi-dimensional structure of C. elegans thermal learning* 2022.

103. Roman, A., Zhu, R. & Nemenman, I. *Ballistic deposition with memory: a new universality class of surface growth with a new scaling law* 2022.

104. Ryu, W. S. & Samuel, A. D. Thermotaxis in Caenorhabditis elegans analyzed by measuring responses to defined Thermal stimuli. *J Neurosci* **22,** 5727–33 (2002).

105. Saeki, S., Yamamoto, M. & Iino, Y. Plasticity of chemotaxis revealed by paired presentation of a chemoattractant and starvation in the nematode Caenorhabditis elegans. eng. *J Exp Biol* **204,** 1757–1764. ISSN: 0022-0949 (05/2001).

106. Safaai, H., Onken, A., Harvey, C. D. & Panzeri, S. Information estimation using nonparametric copulas. *Phys. Rev. E* **98,** 053302 (5 11/2018).

107. Salvador, L. C. M., Bartumeus, F., Levin, S. A. & Ryu, W. S. Mechanistic analysis of the search behaviour of Caenorhabditis elegans. *J R Soc Int* **11,** 20131092. ISSN: 1742-5689 (2014).

108. Saravanan, V., Berman, G. J. & Sober, S. J. Application of the hierarchical bootstrap to multi-level data in neuroscience. *Neurons, behavior, data analysis and theory* **3** (2020).

109. Schultz, W. Multiple reward signals in the brain. *Nature Reviews Neuroscience* **1,** 199–207 (2000).

110. Schultz, W. Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology* **80,** 1–27 (1998).

111. Schwarz, G. Estimating the Dimension of a Model. *The Annals of Statistics* **6,** 461–464, 4 (1978).

112. Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal* **27,** 379–423 (1948).

113. Srivastava, K. H. *et al.* Motor control by precisely timed spike patterns. *Proceedings of the National Academy of Sciences* **114,** 1171–1176 (2017).

114. Stein, G. M. & Murphy, C. T. The Intersection of Aging, Longevity Pathways, and Learning and Memory in C. elegans. eng. *Frontiers in Genetics* **3,** 259 (2012).

115. Stiernagle, T. Maintenance of C. elegans. *WormBook* **11** (02/2006).

116. Strong, S. P., Koberle, R., van Steveninck, R. R. d. R. & Bialek, W. Entropy and information in neural spike trains. *Physical Review Letters* **80,** 197 (1998).

117. Sun, T., Guo, H. & Grant, M. Dynamics of driven interfaces with a conservation law. *Physical Review A* **40,** 6763–6766 (1989).

118. Sutton, R. & Barto, A. *Reinforcement learning: An introduction* (MIT Press, Cambridge, MA, 1998).

119. Takens, F. *Detecting strange attractors in turbulence* in *Dynamical Systems and Turbulence, Warwick 1980* (eds Rand, D. & Young, L.-S.) (Springer Berlin Heidelberg), 366–381. ISBN: 978-3-540-38945-3.

120. Tishby, N., Pereira, F., Bialek, W., Hajek, B. & Sreenivas, R. Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing (1999).

121. Tomioka, M. *et al.* The insulin/PI 3-kinase pathway regulates salt chemotaxis learning in Caenorhabditis elegans. *Neuron* **51,** 613–25 (2006).

122. Torayama, I., Ishihara, T. & Katsura, I. Caenorhabditis elegans integrates the signals of butanone and food to enhance chemotaxis to butanone. eng. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **27,** 741–750 (01/2007).

123. Victor, J. D. Binless strategies for estimation of information from neural data. *Phys. Rev. E* **66,** 051903 (5 11/2002).

124. Villain, J. Continuum models of crystal growth from atomic beams with and without desorption. *J. Phys. I France* **1,** 19–42 (1991).

125. Vinck, M., Battaglia, F. P., Balakirsky, V. B., Vinck, A. J. H. & Pennartz, C. M. A. Estimation of the entropy based on its polynomial representation. *Phys. Rev. E* **85,** 051139 (5 05/2012).

126. Vogel, E. H., Castro, M. E. & Saavedra, M. A. Quantitative models of Pavlovian conditioning. *Brain Res Bull* **63,** 173–202 (2004).

127. Wang, C. J., Bergmann, A., Lin, B., Kim, K. & Levchenko, A. Diverse Sensitivity Thresholds in Dynamic Signaling Responses by Social Amoebae. *Science Signaling* **5,** ra17–ra17 (2012).

128. Ward, S. Chemotaxis by the nematode Caenorhabditis elegans: identification of attractants and analysis of the response by use of mutants. *Proc Natl Acad Sci (USA)* **70,** 817–821 (03/1973).

129. Wendel, J. Zero-free intervals of semi-stable Markov processes. *Mathematica Scandinavica* **14,** 21–34 (1964).

130. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode Caenorhabditis elegans. eng. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **314,** 1–340 (11/1986).

131. Witten, T. & Sander, L. Diffusion-limited aggregation. *Phys Rev B* **27,** 5686 (1983).

132. Wolf, D. E. & Villain, J. Growth with Surface Diffusion. *Europhysics Letters (EPL)* **13,** 389–394 (1990).

133. Wolpert, D. H. & DeDeo, S. Estimating functions of distributions defined over spaces of unknown size. *Entropy* **15,** 4668–4699 (2013).

134. Wolpert, D. H. & Wolf, D. R. Estimating functions of probability distributions from a finite set of samples. *Physical Review E* **52,** 6841 (1995).

135. Xiao, R. *et al.* A genetic program promotes C. elegans longevity at cold temperatures via a thermosensitive TRP channel. *Cell* **152,** 806–817 (02/2013).

136. Xiong, W., Faes, L. & Ivanov, P. C. Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. *Phys. Rev. E* **95,** 062114 (6 06/2017).

137. Yamada, Y. & Ohshima, Y. Distribution and movement of Caenorhabditis elegans on a thermal gradient. *J Exp Biol* **206,** 2581–93 (2003).

138. Zhang, Y., Lu, H. & Bargmann, C. I. Pathogenic bacteria induce aversive olfactory learning in Caenorhabditis elegans. en. *Nature* **438,** 179–184. ISSN: 1476-4687 (11/2005).

139. Zhao, C. *et al.* Predictive olfactory learning in Drosophila. *Sci Rep* **11,** 6795 (2021).