

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:

Baohua Zhou

Date

Explorations of Sensorimotor Learning and Individuality
in Mathematical Models of Behavioral Data

By

Baohua Zhou
Doctor of Philosophy
Physics

Ilya Nemenman, Ph.D
Advisor

Dr. Gordon Berman, Ph.D
Committee Member

Dr. Minsu Kim, Ph.D
Committee Member

Dr. Stefan Boettcher, Ph.D
Committee Member

Dr. Samuel J. Sober, Ph.D
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D
Dean of the James T. Laney School of Graduate Studies

Date

Explorations of Sensorimotor Learning and Individuality
in Mathematical Models of Behavioral Data

By

Baohua Zhou
B.S., Qingdao University, 2008
M.S., Beijing University of Technology, 2011

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
2019

Abstract

Explorations of Sensorimotor Learning and Individuality in Mathematical Models of Behavioral Data

By Baohua Zhou

One of the most obvious properties of a living organism is its highly coordinated behavioral patterns. These patterns reflect the orderly activities of the underlying nervous system that controls the behavior, and also may contain information about this organism's individuality. This thesis is devoted to the studies of behavior in diverse animal systems, and includes two distinct themes: computational rules of sensorimotor control and species classification. First, we try to build a full Bayesian theory of sensorimotor learning by investigating the singing behavior of songbirds and its response to imposed perturbations. Unlike previous models, our normative Bayesian filter model with multiple temporal scales and non-Gaussian sensory feedback can describe the dynamics of the entire probability distribution of the pitch of the song produced by the animals. We will show that our model not only accounts for salient features of the adaptation dynamics, such as the non-monotonic dependence of the behavioral compensation on the perturbation size, but also makes predictions about the behavioral patterns in the classical three-phase perturbation experiments, about the effects of pharmacological silencing of certain nuclei in the bird brain, and about the neural dynamics that is required to implement the model. After this, we move from a single behavior to the whole behavioral repertoire, which is automatically extracted from naturally moving flies. We can show that this behavioral repertoire contains information about the individuality of young flies, and that it can be used to classify fly species in high accuracy. This result suggests that animal behavior patterns might be used to study evolution without the need of any genetic information.

Explorations of Sensorimotor Learning and Individuality
in Mathematical Models of Behavioral Data

By

Baohua Zhou
B.S., Qingdao University, 2008
M.S., Beijing University of Technology, 2011

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
2019

Acknowledgement

It takes me almost seven years to arrive at this thesis, but the whole experience along the way is well worth the time. At the end of this long journey, I owe my thanks to my advisor Ilya Nemenman. He is knowledgeable and creative, and has the special ability to connect seemingly unrelated ideas together. Without his guidance, this thesis can never happen. Besides, it is simply enjoyable to listen to him discussing science, and I feel lucky that I have the chance to sit in his group for so many years and have him in front explaining his ideas.

I also need to thank Sam Sober for his excellent data set on songbird vocal learning and for the great discussions when we develop the theory. I thank Gordon Berman for his mentoring on the fly project, and Stefan Boettcher and Minsu Kim for their teaching and mentoring at the early stage of my graduate life at Emory.

I feel grateful for having met so many wonderful people in Nemenman's Lab: Catalina Rivera, Ahmed Roman, David Hofmann, Michael Martini, Itai Pinkovezky, Damian Hernandez, Joe Natale, George Leung, Xinxian Shao, Vijay Singh, Caroline Holmes, and many others. Thanks for your friendship and encouragement, and for making N117 so active and fun.

Special thanks goes to the champion team Physics FC: Emrah Simsek, Michael Thees, Ryan Freeman, Xiang Cheng, and other team members. Winning the soccer tournament in 2017 with you guys is one of the most memorable moments during my seven years' stay at Emory.

My parents have devoted their entire lives to me and to my two little brothers, and I can never pay off what I owe them. This thesis is dedicated to them for their selfless love, and I just want them to know that I am so proud to be their son. My two brothers, Tong and Kanghua, are always supportive and encouraging, and they are the best gifts that my parents can ever give to me.

Most importantly, I want to thank my life partner, Chenchen, who is also my best friend ever. I understand her complaint and criticism, but also appreciate her support and care. I need to become a much better person in order not to feel indebted to her.

At last, thanks goes to many others who has helped and supported me in my life.

Table of Contents

1	Introduction	1
2	Sensorimotor learning as a 1d Bayesian filter	5
2.1	Introduction	5
2.2	Experimental data	8
2.3	Mathematical model	10
2.3.1	1d Bayesian filter	10
2.3.2	Non-Gaussian statistics	14
2.3.3	Fits and predictions	15
2.4	Discussion	18
2.5	Materials and methods	21
2.5.1	Experiments	21
2.5.2	Stable distributions	22
2.5.3	Fitting	23
2.5.4	Choice of the shape of distributions	25
2.5.5	Linear dependence on pitch shift in a Kalman filter with multiple time scales	26
3	Sensorimotor learning as a 2d Bayesian filter	30
3.1	Introduction	30
3.2	Mathematical model	33
3.2.1	2d Bayesian filter	33
3.2.2	Fits and predictions	39
3.2.3	Biological interpretations	44
3.3	Discussion	48
3.4	Materials and methods	51
3.4.1	Experimental data	51
3.4.2	Data bootstrapping	53
3.4.3	Simplification of representations	53
3.4.4	Powerlaw-like distributions	54
3.4.5	Model fitting and parameters	54
3.5	Additional figures	55
4	Behavioral map of a fly predicts species identity	68
4.1	Introduction	68
4.2	Behavior maps of fly	68
4.3	Embedding by t-SNE	70
4.4	Classification by logistic regression	71
4.5	Discussion	72

5 Conclusion	73
Bibliography	78

List of Figures

2.1	1d dynamical Bayesian filter	7
2.2	Fits to the baseline distribution and average learning curves	9
2.3	Predictions about the dynamics of variance	17
2.4	Objective function landscapes	19
2.5	Fits and predictions with the power law distributions	26
3.1	Physical analogy of the 2d dynamical Bayesian filter	34
3.2	2d dynamical Bayesian filter	35
3.3	Fits to all the pitch distributions	40
3.4	Fits quality shown as Jensen-Shannon divergence	41
3.5	Dynamics of medians of the pitch distributions	42
3.6	Model predictions for two three-phase experiments	43
3.7	Biological interpretation of the model	44
3.8	Model-predicted LMAN contribution to learning	47
3.9	Relations between variability and learning	49
3.10	Predicted joint distributions	50
3.11	Cascade model	52
3.12	The landscapes of the objective function	53
3.13	Dynamics of medians when fitting the one-timescale model to all the distributions	55
3.14	All distributions for the 0.5-semitone perturbation	56
3.15	All distributions in log scale for the 0.5-semitone perturbation	57
3.16	All distributions for the 1.0-semitone perturbation	58
3.17	All distributions in log scale for the 1.0-semitone perturbation	59
3.18	All distributions for the 1.5-semitone perturbation	60
3.19	All distributions in log scale for the 1.5-semitone perturbation	61
3.20	All distributions for the 3.0-semitone perturbation	62
3.21	All distributions in log scale for the 3.0-semitone perturbation	63
3.22	All distributions for the staircase perturbation	64
3.23	All distributions in log scale for the staircase perturbation	65
3.24	Fits of the model with a correlated kernel	66
3.25	Fits of the cascade model	67
4.1	Phylogenetic tree and behavioral repertoire	70
4.2	Classification of fly species	71

Chapter 1 Introduction

An animal uses its behaviors to strategically respond to the challenges and opportunities imposed by the surrounding environment, and the adaptation of these strategies is crucial for its survival. Through billions of years of evolution, animals are able to perform diverse and complex behaviors. Just to name a few of them, a moth needs to properly track a fluctuating flower in order to get food [82]; the nocturnal predator barn owl can reliably turn its head to the location of a sound source with the help of an internal mapping between the sound signal and the encoded spatial information [51]; Tiger Woods has to rely on his arm swing patterns to perform well in a golf contest, and in fact, he has changed these patterns three times throughout his career to achieve better performance [23]. On the one hand, behaviors are direct consequences of the upstream motor commands generated by the nervous system, and the features of the observed behaviors must be embedded in some way in the activities of this nervous system; on the other hand, behavioral patterns may also carry information about the organism's individual features since different organisms behave differently. In this thesis, two distinct topics are discussed: (1) behavioral dynamics can be used to infer general computational rules of the sensorimotor learning systems; (2) a whole behavioral repertoire can be used to distinguish individual animals.

One of the biggest problems in neuroscience is to understand how the nervous system controls behavior. According to Marr [64], there are three levels in this problem: behavior, algorithm and neural implementation. As technology advances, neural recordings can yield much more high-fidelity data than before, and thus, many people are applying various techniques and models to analyze data, hoping that algorithms used by the nervous systems to control the behavior can emerge from the analysis

of the neural activities. However, it has been argued that this bottom-up approach might not be very effective without careful behavioral studies. Unlike the neural activities that are well characterized by the spiking patterns, behaviors are usually hard to define. Part of the reason is that many behaviors are analog and have too many degrees of freedom, and thus it is not easy to determine what to measure. Despite this difficulty, many experimental paradigms and computational models have been developed to study behaviors, many of which have yielded fruitful results. In general, in order to build mathematical models of behavioral data, one needs to define macroscopic variables and their associated dynamics that are related to the behavior under study. If a model agrees well with behavioral observations, it might capture something that is truly embedded in the nervous system—the algorithm. These experimentally tested mathematical models (or algorithms) can then be used as guidance to study neural recordings.

The nervous system controls the behavior by sending motor commands to instruct the activities of relevant muscles, and the consequences of the motor commands are sent back to the nervous system in the forms of sensory feedback [8, 11, 56, 104]. The sensory feedback is then processed by the nervous system to generate an instructive teaching/error signal, which is used for adaptation of the motor commands [91, 93]. This sensorimotor learning loop has been extensively studied across different behaviors and animal species. A typical approach is to add perturbations to the sensory consequences of the motor command and to observe how the behavior responds to these perturbations. Many computational models have been proposed to explain the observed phenomena. One of the most famous ones is the two-rate model [92], which is formulated to explain the observations in the arm-reaching experiments. This model assumes that two states or memories, ϕ_t and ψ_t , are present in the motor system, and each state learns from the same error with different learning rates. Specifically, it is

formulated as

$$\phi_{t+1} = A_\phi \phi_t + B_\phi \delta_t \quad (1.1)$$

$$\psi_{t+1} = A_\psi \psi_t + B_\psi \delta_t \quad (1.2)$$

$$\delta_t = r - (\phi_t + \psi_t) \quad (1.3)$$

$$B_\phi > B_\psi, A_\phi < A_\psi, \quad (1.4)$$

where A_ϕ and A_ψ are constant retention factors, B_ϕ and B_ψ are constant learning rates, and r is the preset target output value, which the system is trying to learn [92]. This model can explain the phenomenon of savings where a spontaneous recovery of previous learned behavior is observed after an adaptation-extinction training episode [92]. The two-rate model suggests a specific algorithm that the nervous system might use to control the behavior, and how this algorithm is implemented in the level of neurons and neural circuits needs further investigation.

While this two-rate model and many other sensorimotor learning models have successfully explained a large amount of experimental observations, they only focus on the dynamics of average behaviors. However, many experiments have shown that motor output is highly variable, and that a large portion of this variability is not random biological noise, but can actually be controlled by the animal itself [19, 44, 45]. For example, the song of a male Bengalese finch is usually very variable if there is no female finch nearby (undirected song); however, the male finch can produce highly accurate songs in front of a female finch (directed song) [45]. This controlled variability provides the animal opportunities to explore more of the motor command space in order to respond more quickly and adapt better to sudden environmental changes. In fact, it has been shown in some experiments that the initial learning rate is positively correlated with the task-relevant motor variability [105]. In addition, none of the previous sensorimotor learning models can account for the fact that animals adapt to small perturbations faster and better than to large perturbations. This type

of behaviors has been observed in many animals, such as barn owls adapting to prism-induced visual perturbations [50], Bengalese finches adapting to headphone-induced auditory shifts [94], and human beings adapting to visual disturbances [102].

In Chapter 2 and Chapter 3, we propose a new sensorimotor learning model. The model has the classical form of a Bayesian filter, which has long been suggested to account for learning and information integration in the brain in face of fluctuating internal and external environments [3, 24, 25, 49, 54, 79, 101]. We test our model on the singing behavior data of Bengalese finches. The distributions of the recorded song pitches show non-Gaussian tails, and thus, we incorporated various non-Gaussian terms in our Bayesian theory. Surprisingly, this turns out to account for the negative correlation between the perturbation sizes and adaptations. In addition, with the hypothesis that there are more than one interacting dynamical processes in the motor system [92], we show that our model can describe the dynamics of the whole distributions of the singing behaviors.

After discussing sensorimotor learning models of a single behavior, we switch the topic and focus on a group of behaviors. In Chapter 4, we demonstrate that automatically defined behavioral map of a fly has information about its individuality, and that the map can be used to classify fly species with high accuracy.

Chapter 2 Sensorimotor learning as a 1d Bayesian filter

2.1 Introduction

¹ Learned behaviors — reaching for an object, talking, and hundreds of others — allow the organism to interact with the ever-changing surrounding world. To learn and execute skilled behaviors, it is vital for such behaviors to fluctuate from iteration to iteration. Such variability is not limited to inevitable biological noise [65, 91], but rather a significant part of it is controlled by animals themselves and is used for exploration during learning [45, 71]. Furthermore, learned behaviors rely heavily on sensory feedback. The feedback is needed, first, to guide the initial acquisition of the behaviors, and then to maintain the needed motor output in the face of changes in the motor periphery and fluctuations in the environment. Within such sensorimotor feedback loops, the brain computes how to use the inherently noisy sensory signals to change patterns of activation of inherently noisy muscles to produce the desired behavior. This transformation from sensory feedback to motor output is both robust and flexible, as demonstrated in many species in which systematic perturbations of the feedback dramatically reshape behaviors [51, 60, 91–93].

Since many complex behaviors are characterized by both tightly controlled motor variability and by robust sensorimotor learning, we propose that, during learning, the brain controls the *distribution of behaviors*. In contrast, most prior theories of animal learning have assumed that there is a single optimal motor command that the animal tries to produce, and that, after learning, deviations from the optimal behavior result from the unavoidable (Gaussian) downstream motor noise. Such prior models include

¹This chapter is adapted from the published paper: B Zhou, D Hofmann, I Pinkoviezky, SJ Sober, I Nemenman Proceedings of the National Academy of Sciences 115 (36), E8538-E8546

the classic Rescorla-Wagner (RW) model [80], as well as more modern approaches belonging to the family of reinforcement learning [42, 59, 96], Kalman filters [53, 103], or dynamical Bayesian filter models [32, 34]. In many of these theories, the variability of the behavior is intrinsic to motor exploration and deliberately controlled, but the distribution of this variability is not itself shaped by the animal’s experience [21, 26, 27]. Such theories have addressed many important experimental questions, such as evaluating the optimality of the learning process [2, 20, 53, 99, 102], accounting for multiple temporal scales in learning [48, 53, 92, 101], identifying the complexity of behaviors that can be learned [33], and pointing out how the necessary computations could be performed using networks of spiking neurons [18, 29, 59, 72, 87, 89].

However, despite these successes, most prior models that assume that the brain aims to achieve a single optimal output have been unable to explain some commonly observed experimental results. For example, since such theories assume that errors between the target and the realized behavior drive changes in future motor commands, they typically predict large behavioral changes in response to large errors. In contrast, experiments in multiple species report a decrease in both the speed and the magnitude of learning with an increase in the experienced sensory error [51, 81, 94, 102]. One can rescue traditional theories by allowing the animal to *reject* large errors as “irrelevant”—unlikely to have come from its own actions [37, 102]. However, such rejection models have not yet explained why the same animals that cannot compensate for large errors can correct for even larger ones, as long as their magnitude grows gradually with time [51, 94].

Here we present a theory (Fig. 2.1) of a classic model system for sensorimotor learning — vocal adaptation in a songbird — in which the brain controls a probability distribution of motor commands and updates this distribution by a recursive Bayesian inference procedure. The distribution of the song pitch is empirically heavy tailed (Fig. 2.2C), and the pitch variability is much smaller in a song directed at a female vs

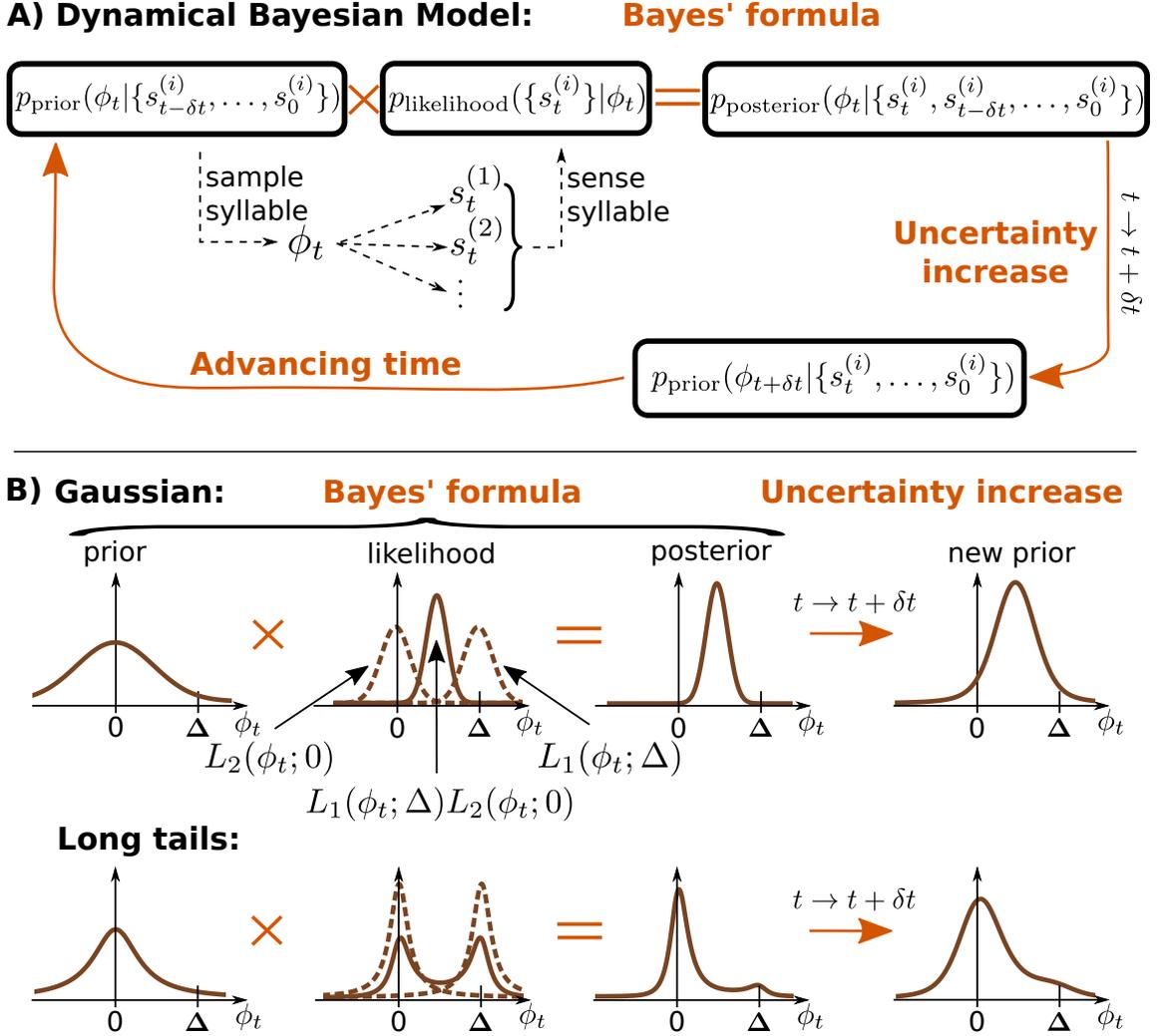


Figure 2.1: The dynamical Bayesian model (Bayesian filter). (A) A Bayesian filter consists of the recursive application of two general steps [43]: i) an *observation update*, which corresponds to novel sensory input and updates the underlying probability distribution of plausible motor commands using the Bayes' formula; ii) a *time evolution update*, which denotes the temporal propagation and corresponds to uncertainty increasing with time (see main text); here the probability distribution is updated by convolution with a propagator. These two steps are repeated for each new sensory data in a recursive loop. (B) Example distributions for the entire procedure in two scenarios: Gaussian (top) and heavy-tailed (bottom) distributions. The x-axis, ϕ_t , represents the motor command which results in a specific pitch sung by the bird. The outcome of this motor command is then measured by two different sensory modalities, represented by $\{s_t^{(i)}\}_{i=1,2}$, with corresponding likelihood functions $L_1(\phi_t; \Delta)$ and $L_2(\phi_t; 0)$, respectively. The $-\Delta$ shift for modality 1 is induced by the experimentalist, which results in animal compensating its pitch towards $+\Delta$. Dashed brown lines represent the individual likelihood functions from the individual modalities, and the solid lines represent their product, which signals how likely it is that the correct motor command corresponds to ϕ_t . Heavy-tailed distributions can produce a bimodal likelihood, which, multiplied by the prior, suppresses large-error signals. In contrast, Gaussian likelihoods are unimodal and result in greater compensatory changes in behavior.

the undirected song, suggesting that the variability and hence its non-Gaussian tails are deliberately controlled. Thus our model does not make the customary Gaussian assumptions. The focus on learning and controlling (non-Gaussian) distributions of behavior allows us to capture successfully all of the above-described nonlinearities in learning dynamics and, furthermore, to account for previously unnoticed learning-dependent changes in the *shape* of the distribution of the behavior.

2.2 Experimental data

Vocal control in songbirds is a powerful model system for examining sensorimotor learning of complex tasks [10]. The phenomenology we are trying to explain arises from experimental approaches to inducing song plasticity [94]. Songbirds sing spontaneously and prolifically and use auditory feedback to shape their songs towards a “template” learned from an adult bird tutor during development. When sensory feedback is perturbed (see below) using headphones to shift the pitch (fundamental frequency) of auditory feedback [94], birds compensate by changing the pitch of their songs so that the pitch they hear is closer to the unperturbed one. As shown in Fig. 2.2A, the speed of the compensation and its maximum value, which is measured as a fraction of the pitch shift and referred to as the magnitude of learning hereafter, decrease with the increasing shift magnitude, so that a shift of 3 semitones results in near-zero fractional compensation. Crucially, the small compensation for large perturbation does not reflect the limited plasticity of the adult brain since imposing the perturbation gradually, rather than instantaneously, results in a large compensation (Fig. 2.2B).

We use experimental data collected in our previous work [93, 94] to develop our mathematical model of learning. As detailed in Ref. [58], we used a virtual auditory feedback system [39, 93] to evoke sensorimotor learning in adult songbirds. For this, miniature headphones were custom-fitted to each bird and used to provide online au-

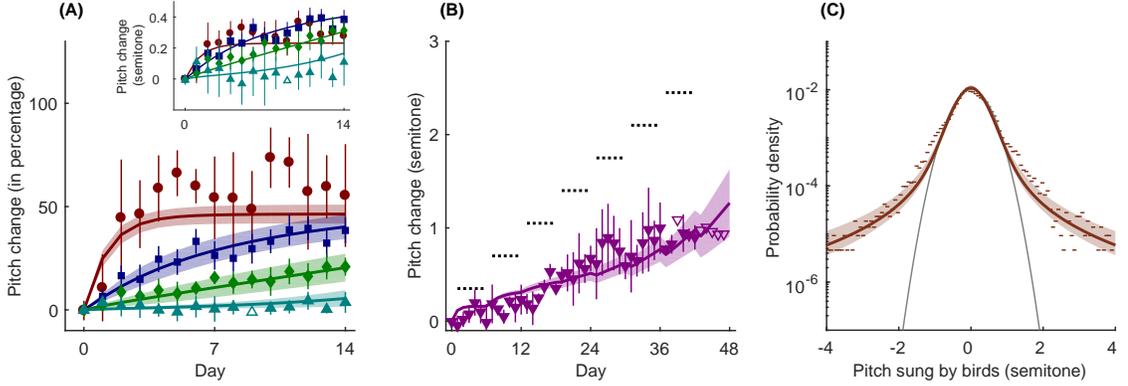


Figure 2.2: Experimental data and model fitting. The same six parameters of the model are used to *simultaneously* fit all data. (A) The dots with error bars are four groups of experimental data, with different colors and symbols indicating different shift sizes (red-brown circle, 0.5 semitone shift; blue square, 1 semitone shift; green diamond, 1.5 semitones shift; cyan upper triangle, 3 semitones shift). The error bars indicate the standard error of the group mean, accounting for variances across individual birds and within one bird, see *Materials and Methods*). For each group, the data are combined from three to eight different birds, and the sign of the experimental perturbation (lowering or raising pitch) is always defined so that adaptive (i. e., error-correcting) vocal changes are positive. Data points without error bars only had a single bird, and they are not used for the fitting, which we denote by hollow symbols. The mean pitch sung on day 0 by each bird is defined as the zero semitone compensation ($\phi = 0$). The solid lines with one standard deviation bands (see *Materials and Methods*) are results of the model fits, with the same color convention as in experimental data. The inset shows learning curves in absolute units, without rescaling by the shift size, and without model error bands. (B) The lower triangles with error bars show the data from staircase-shift experiment, with the same plotting conventions as in (A). The data are combined from three birds. During the experiment, every six days, the shift size is increased by 0.35 semitone, as shown by the dotted horizontal line segments. On the last day of the experiment, the experienced pitch shift is 2.8 semitones. The magenta solid line with one standard deviation band is the model fit. The combined combined quality of fit for the five curves collectively (4 step perturbations and a staircase perturbation) is χ^2 per degree of freedom ≈ 1.47 (compared to χ^2 per degree of freedom of ≈ 28.4 for the null model of the non-adaptive, zero pitch compensation line). Note, however, that such Gaussian statistics of the fit quality should be taken with a grain of salt for non-normally distributed data. (C) Dots represent the distribution of pitch on day 0, before the pitch shift perturbation (the baseline distribution), where the data are from 23 different experiments (all pitch shifts combined). The gray parabola is a Gaussian fit to the data within the ± 1 semitone range. The empirical distribution has long, non-exponential tails. The brown solid line with one standard deviation band is the model fit. Deviance of the model fit relative to the perfect fit (the latter estimated as the NSB entropy of the data [69, 70]) is 0.057 per sample point.

ditory feedback in which the pitch (fundamental frequency) of the bird’s vocalizations could be manipulated in real time, with a loop delay of roughly 10 ms. In addition to providing pitch-shifted feedback, the headphones largely blocked the airborne transmission of the bird’s song from reaching the ear canals, thereby effectively replacing the bird’s natural airborne auditory feedback with the manipulated version. Pitch shifts were introduced after a baseline period of at least 3 days in which birds sang while wearing headphones but without pitch shifts. All pitch shifts were implemented relative to the bird’s current vocal pitch and were therefore “correctable” in the sense that if the bird changed its vocal pitch to fully compensate for the imposed pitch shift, the pitch of auditory feedback heard through the headphones would be equal to its baseline value. All data were collected during undirected singing (i. e., no female bird was present).

2.3 Mathematical model

2.3.1 1d Bayesian filter

To describe the data, we introduce a dynamical Bayesian filter model, Fig. 2.1A. We focus on just one variable learned by the animal during repeated singing — the pitch of the song syllables. Even though the animal learns the motor command and not the pitch directly, we do not distinguish between the produced pitch ϕ and the motor command leading to it because the latter is not known in behavioral experiments. We set the mean “baseline” pitch sung by the animal as $\phi = 0$, representing the “template” of the tutor’s song, or the scalar target memorized during development, and nonzero values of ϕ denote deviations of the sung pitch from the target.

However, while an instantaneous output of the motor circuit in our model is a scalar value of the pitch, the state of the motor learning system at each time step is a probability distribution over motor commands that the animal expects can lead to the target motor behavior. This is in contrast to the more common assumption that

the state of the learning system is a scalar, usually the mean behavior, which is then corrupted by the downstream noise [37]. Thus at time t , the animal has access to the prior distribution over plausible motor commands, $p_{\text{prior}}(\phi_t)$. We remain deliberately vague about how this distribution is stored and updated in the animal memory (e. g., as a set of moments, or values, or samples, or yet something else), and focus instead not on *how* the neural computation is performed, but on modeling *which* computation is performed by the animal. We assume that the bird randomly selects and produces the pitch from this distribution of plausibly correct motor commands. In other words, we suggest that the experimentally observed variability of sung pitches is dominated by the deliberate exploration of plausible motor commands, rather than by noise in the motor system. This is supported by the experimental finding that the variance of pitch during singing directed at a female (performance) is significantly smaller than the variance during undirected singing (practice) [45, 77].

After producing a vocalization, the bird then senses the pitch of the produced song syllable through various sensory pathways. Besides the normal airborne auditory feedback reaching the ears, which we can pitch-shift, information about the sung pitch may be available through other, unmanipulated pathways. For example, efference copy may form an internal short term memory of the produced specific motor command [74]. Additionally, proprioceptive sensing presumably also provides unshifted information [95]. Finally, unshifted acoustic vibrations might be transmitted through body tissue in addition to the air, as is thought to be the case in studies that use pitch shifts to perturb human vocal production [61, 86].

We denote all feedback signals as $s_t^{(i)}$ where the index i denotes different sensory modalities. Because sensing is noisy, feedback is not absolutely accurate. We posit that the animal interprets it using the Bayes' formula. That is, the posterior probability of which motor commands would lead to the target with no error is changed by the observed sensory signals, $p_{\text{post}}(\phi_t) \propto p_{\text{likelihood}}(\{s_t^{(i)}\}|\phi_t)p_{\text{prior}}(\phi_t)$, where $p_{\text{likelihood}}$

represents the probability of observing a certain sensory feedback value given the produced motor command ϕ_t was the correct one. In its turn, the motor command is chosen from the prior distribution, p_{prior} , which represents the *a priori* probability of the command to result in no sensory error. In other words, if the sensory feedback indicates that the pitch was likely too high, then the posterior is shifted towards motor commands that have a higher probability of producing a lower pitch, and hence no sensory error – similar to how an error would be corrected in a control-theoretic approach to the same problem. We discuss this in more detail below.

Finally, the animal expects that the motor command needed to produce the target pitch with no error may change with time because of slow random changes in the motor plant. In other words, in the absence of new sensory information, the animal must increase its uncertainty about which command to produce with time (this is a direct analogue of increase in uncertainty of the Kalman filter without new measurements). Such increase in uncertainty is given by $p_{\text{prop}}(\phi_{t+\delta t}|\phi_t)$, the propagator of statistical field theories [107]. Overall, this results in the distribution of motor outputs after one cycle of the model

$$p_{\text{prior}}(\phi_{t+\delta t}) = \frac{1}{Z} \int p_{\text{prop}}(\phi_{t+\delta t}|\phi_t) \times p_{\text{likelihood}}(\{s_t^{(i)}\}|\phi_t) p_{\text{prior}}(\phi_t) d\phi_t, \quad (2.1)$$

where Z is the normalization constant.

We choose δt to be one day in our implementation of the model and lump all vocalizations (which we record) and all sensory feedback (which are unknown) in one time period together. That is, we look at timescales of changes across days, rather than faster fluctuations on timescales of minutes or hours. This matches the temporal dynamics of the learning curves (Fig. 2.2A, B). Since the bird sings hundreds of song bouts daily, we now use the law of large numbers and replace the unknown sensory feedback for individual vocalizations by its expectation value $s_t^{(i)} \rightarrow \bar{s}_t^{(i)}$. For simplicity, we focus on just two sensory modalities, the first affected by the

headphones, and the second one not, and we remain agnostic about the exact nature of this second modality among the possibilities noted above. Thus the expectation values of the feedbacks are the shifted and the unshifted versions of the expected value of the sung pitch, $\bar{s}_t^{(1)} = \bar{\phi}_t - \Delta$ and $\bar{s}_t^{(2)} = \bar{\phi}_t$, where $-\Delta$ is the experimentally induced shift (see more on the $-$ sign below). Note that since ϕ_t is the motor command that the animal expects to produce the target pitch, the term $p_{\text{likelihood}}(s_t^{(i)}|\phi_t)$ should be viewed as the probability of generating the feedback $s_t^{(i)}$ given that ϕ_t was the correct motor command, or as a likelihood of ϕ_t being the correct command given the observed $s_t^{(i)}$. This introduces a negative sign, the compensation, into the analysis — for a positive $s_t^{(i)}$, the most likely ϕ_t to lead to the target is negative, and vice versa. While potentially confusing, this is the same convention that is used in all filtering application — a positive sensory signal mean the need to compensate and to lower the motor command, and the negative signal leads to the opposite. In other words, the bird uses the sensory feedback to determine what it should have sung, and not only what it sang. With that, we refer to the conditional probability distributions $p_{\text{likelihood}}(s_t^{(i)}|\phi_t)$ for each sensory modality i as the likelihood functions $L_i(\phi_t)$ for a certain motor command being the target given the observed sensory feedback. Thus assuming that both sensory inputs are independent measurements of the motor output, we rewrite (2.1) as

$$\begin{aligned}
 p_{\text{prior}}(\phi_{t+\delta t}) &= \frac{1}{Z} \int p_{\text{prop}}(\phi_{t+\delta t}|\phi_t) \\
 &\quad \times L_1(\phi_t; \Delta)L_2(\phi_t; 0)p_{\text{prior}}(\phi_t)d\phi_t,
 \end{aligned}
 \tag{2.2}$$

where 0 and Δ represent the centers of the likelihoods (or the maximum likelihoods). This explains our choice of denoting the experimental shift $-\Delta$, so that the compensation by the animal is instead $+\Delta$, and L is centered on $+\Delta$ as well. Notice that the likelihoods for the shifted and unshifted modalities are centered around Δ and 0, respectively, and bias the learning of what should be sung towards these centers irrespectively of the current value of ϕ_t . We emphasize again that, in this formalism,

we do not distinguish the motor noise and the sensory noise, and assume that both are smaller than the deliberate exploratory variance (which is supported by the substantial variance reduction in directed vs. undirected song). This is consistent with not distinguishing individual vocalizations and focusing on time steps of one day in the Bayesian update equation above.

2.3.2 Non-Gaussian statistics

As illustrated in Fig. 2.1B, such Bayesian filtering behaves differently for Gaussian and heavy-tailed likelihoods and propagators. Indeed, if the two likelihoods are Gaussians, their product is also a Gaussian centered between them. In this case, the learning speed of an animal is linear in the error Δ , no matter how large this error is, which conflicts with the experimental results in songbirds and other species [10, 60, 93, 102]. Similarly, if the two likelihoods have long tails, then when the error is small, their product is also a single-peaked distribution as in the Gaussian case. However, when the error size Δ is large, the product of such long-tailed likelihoods is bimodal, with evidence peaks at the shifted and the unshifted values, with a valley in the middle. Since the prior expectations of the animal are developed before the sensory perturbation is turned on, they peak near the unshifted value. Multiplying the prior by the likelihood then leads to suppression of the shifted peak and hence of large error signals in animal learning.

In (2.2), there are three distributions to be defined: $L_1(\phi_t; \Delta)$, $L_2(\phi_t; 0)$, and $p_{\text{prop}}(\phi_{t+\delta t}|\phi_t)$, corresponding to the evidence term from the shifted channel, the evidence term from the unshifted channel, and the time propagation kernel, respectively. The prior at the start of the experiment $t = 0$, $p_{\text{prior}}(\phi_0)$, is not an independent degree of freedom: it is the steady state of the recurrent application of (2.2) with no perturbation, $\Delta = 0$. We have verified numerically that a wide variety of shapes of L_1 , L_2 and p_{prop} result in learning dynamics that can approximate the experimental

data (see *Materials and Methods*). To constrain the selection of specific functional forms of the distributions, we point out that the error in sensory feedback obtained by the animal is a combination of many noisy processes, including both sensing itself and the neural computation that extracts the pitch from the auditory input and then compares it to the target pitch. By the well-known generalized central limit theorem, the sum of these processes is expected to converge to what are known as Lévy alpha-stable distributions, often simply called stable distribution [75] (see *Materials and Methods*). If the individual noise sources have finite variances, the stable distribution will be a Gaussian. However, if the individual sources have heavy tails and infinite variances, then their stable distribution will be heavy-tailed as well (Cauchy distribution is one example). Most stable distributions cannot be expressed in a closed form, but they can be evaluated numerically (see *Materials and Methods*). Here we assume symmetric stable distributions, truncated at ± 8 semitones (see *Materials and Methods*). Such distributions are characterized by three parameters: the stability parameter α (measuring the proportion of the probability in the tails), the scale or width parameter γ , and the location or the center parameter μ (the latter can be predetermined to be 0, Δ , or the previous time step value in our case). For three distributions $L_1(\phi_t; \Delta)$, $L_2(\phi_t; 0)$, and p_{prop} , this results in the total of six unknown parameters.

2.3.3 Fits and predictions

We fit the set of six parameters of our model simultaneously to *all* the data shown in Fig. 2.2. Our dataset consists of twenty-three individual experiments across five experimental conditions: four constant pitch shift learning curves and one gradual, staircase shift learning curve (see *Material and Methods* for details). As mentioned previously, birds learn the best (larger and faster compensation) for smaller perturbations, here 0.5-semitone, panel (A). In contrast, for a large 3-semitone perturbation,

the birds do not compensate at all within the 14 days of the experiment. However, the birds are able to learn and compensate large perturbations when the perturbation increases gradually, as in the staircase experiment in panel (B). Importantly, the baseline distribution (panel C) has a robust non-Gaussian tail, supporting our model. We note that our six-parameter model fits are able to simultaneously describe *all* of these data with a surprising precision, including their most salient features: dependence of the speed and the magnitude of the compensation on the perturbation size for the constant and the staircase experiments, as well as the heavy tails in the baseline distribution.

Mathematical models are useful to the extent that they can predict experimental results not used to fit them. Quantitative predictions of *qualitatively* new results are particularly important for arguing that the model captures the system’s behavior. To test the predictive power of our model, we used it to predict the dynamics of higher-order statistics of pitches during learning, rather than using it to simply predict the mean behavior. We first use the model to predict time-dependent measures of the variability (standard deviation in this case) of the pitch. As shown in Figure 2.3A-E, our model correctly predicted time-dependent behaviors in the standard deviation in both single-shift (Figure 2.3A-D) and staircase-shift experiments (Figure 2.3E) with surprising accuracy. We stress again that no new parameter fits were done for these curves. Potentially even more interestingly, Fig. 2.3F shows that our model is capable of predicting unexpected features of the probability distribution of pitches, such as the asymmetric and bimodal structure of the pitch distribution at the end of the staircase-shift experiment. This bimodal structure is predicted by our theory, since the theory posits that the (bimodal) likelihood distribution (Fig. 2.1B, bottom) will iteratively propagate into the observable pitch distribution (the prior). The existence of the bimodal pitch distribution in the data therefore provides strong evidence in support of our theory. Importantly, this phenomenon can never be reproduced by models

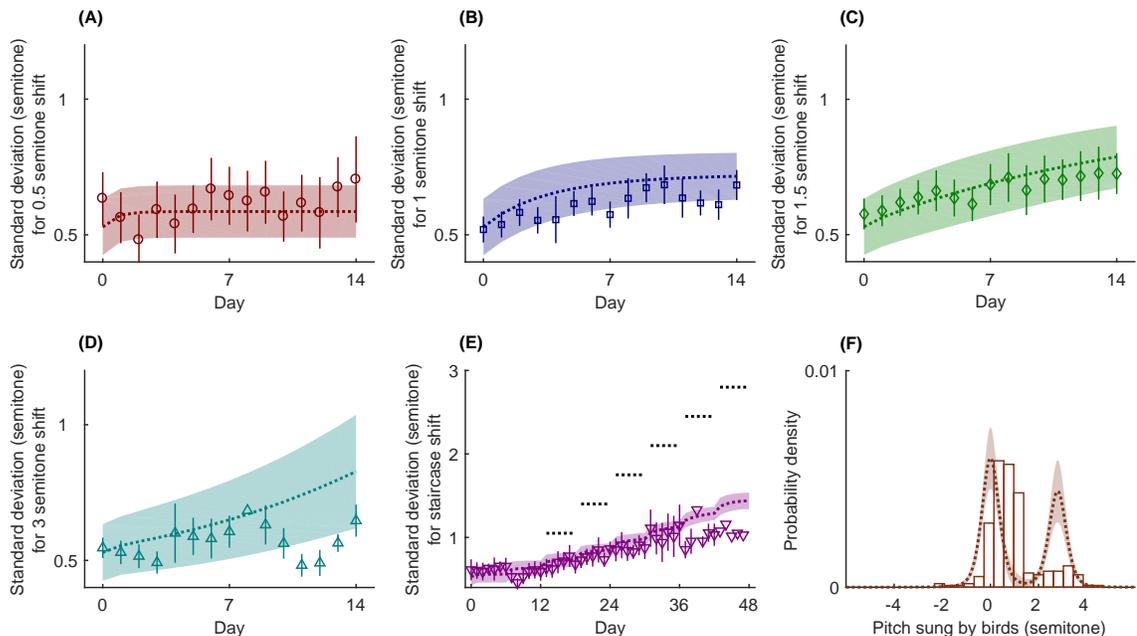


Figure 2.3: Predictions of our model using the parameter values obtained from fitting the data shown in Fig. 2.2. The dots with error bars (A-E) and the histogram (F) represent experimental data with colors, symbols, error bars (from bootstrapping), and other plotting conventions as in Fig. 2.2. The dotted lines with one standard deviation bands represent model predictions. Our model correctly predicts the behaviors of the standard deviations of the pitch distributions. Specifically, the best-fit model lines predict increases in the standard deviation in panels (B,C,E), which correspond to 1 semitone, 1.5 semitones, and the staircase shift, respectively. At the same time, the data says that the standard deviation increases for these panels (p-value for a positive dependence of the standard deviation when regressed on time is 4×10^{-4} , 5×10^{-5} and $< 10^{-6}$ for these panels). (F) Our model predicts that, at the end of the staircase experiment (mean and standard deviation shown in Figs. 2.2B and 2.3E, respectively), the pitch distribution should be bimodal, while it is unimodal initially (cf. Fig. 2.2C). This is also supported by the data. Specifically, a fit with a mixture of two Gaussian peaks has the AIC score higher than a fit to a single Gaussian by 50 (in decimal log units), which is highly statistically significant (the data here is from day 47 from the single bird who was exposed to the staircase shift for the longest time, and the amount of data is insufficient to fit more complex distributions). Further, the two peaks are centered far from each other (0.59 ± 0.04 of a semitone and 2.17 ± 0.62 semitones, with error bars obtained by bootstrapping), illustrating the true bimodality. Neither the data nor the models show unambiguous bimodality in other learning cases.

based on animals learning a single motor command with Gaussian noise around it, rather than a heavy-tailed distribution of motor commands.

2.4 Discussion

We introduced a novel mathematical framework within the class of observation-evolution models [43] for understanding sensorimotor learning: a dynamical Bayesian filter with non-Gaussian (heavy-tailed) distributions. Our model describes the dynamics of the *whole probability distribution* of the motor commands, rather than just its mean value. We posit that this distribution controls the animal’s deliberate exploration of plausible motor commands. The model reproduces the learning curves observed in a range of songbird vocal adaptation experiments, which classical behavioral theories have not been able to do to date. Further, also unlike the previous models, our approach predicts learning-dependent changes in the width and shape of the distribution of the produced behaviors.

To further increase the confidence in our model, we show analytically (see *Materials and Methods*) that traditional linear models with Gaussian statistics [53] cannot explain the different levels of compensation for different perturbation sizes. While we cannot exclude that birds would continue adapting if exposed to perturbations for longer time periods and would ultimately saturate at the same level of adaptation magnitude, the Gaussian models are also argued against by the shape of the pitch distribution, which shows heavy tails (Fig. 2.2C and 2.3F) and by our ability to predict not just the mean pitch, but the whole pitch distribution dynamics during learning.

An important aspect of our dynamical model is its ability to reproduce multiple different time scales of adaptation (Fig. 2.2A, B) using a nonlinear dynamical equation with just a single time scale of the update (1 day). As with other key aspects of the model, this phenomenon results from the non-Gaussianity of the distributions employed, and specifically the many time scales built into a single long-tailed kernel

of uncertainty increase over time. This is in contrast to other multiscale models that require explicit incorporation of many time scales [53, 92]. While multiple time scales could be needed to account for other features of the adaptation, our model clearly avoids this for the present data. In the future, we hope that an extension of our model to include multiple explicit time scales will account for individual differences across animals, for the dynamics of acquisition of the song during development, and for the slight shift of the peak of the empirical distribution in Fig. 2.3F from $\phi = 0$.

Previous analyses of the speed and magnitude of learning in the Bengalese finch have noted that both depend on the overlap of the distribution of the natural variability at the baseline and at the shifted means [48, 58]: small overlaps result in slower and smaller learning, so that different overlaps lead to different time scales. However, these prior studies have not provided a computational mechanism, or a learning-theoretic explanation of why or how such overlap might determine the dynamics of learning. Our dynamical inference model provides such a computational mechanism.

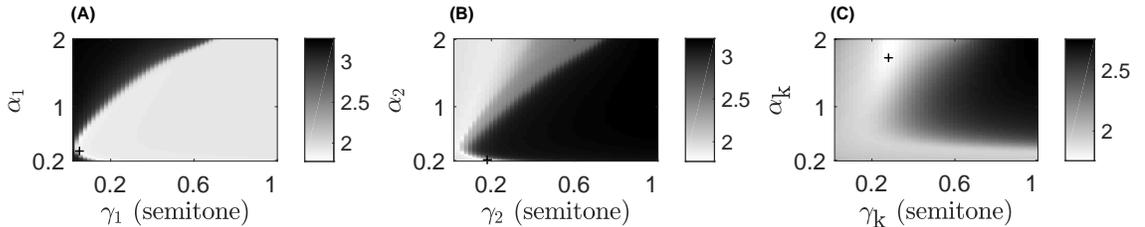


Figure 2.4: Objective function as a function of the two parameters (stability and scale) for (A) the first (shifted) likelihood, (B) the second (unshifted) likelihood, and (C) the propagation kernel, while the respective other four parameters are held fixed. The gray shades represent the decimal logarithm of the objective function (effectively, logarithms of the negative log-likelihood), and lighter colors mean a better fit. Because of the logarithmic scaling, small changes in the color represent large changes in the quality of the fit. The black crosses show the parameter values for the deepest local minimum in this range of parameters. Note that, even though the minimum in (C) is close to the Gaussian kernel ($\alpha_k = 2$), a Gaussian kernel cannot fit the data well. Specifically, it cannot reproduce a non-Gaussian distribution of the baseline pitch, instead essentially matching the parabola in Fig. 2.2(C).

We have chosen the family of so-called Lévy alpha-stable distributions to pro-

vide the central ingredient of our model: the heavy tails of the involved probability distributions. In general, a symmetric alpha-stable distribution has a relatively narrow peak in the center and two long fat tails, and this might provide some valuable qualitative insights into how the nervous system processes sensory inputs. For example, a narrow peak in the middle of the likelihood function suggests that the brain puts a high belief in the sensory feedback. However, the heavy tails say that it also puts certain weight (nearly constant) on the probability of very large errors outside of the narrow central region. We have verified that the actual choice of the stable distributions is not crucial for our modeling. For example, one could instead take each likelihood as a power law distribution, or as a sum of two Gaussians with equal means, but different variances. The latter might correspond to mixture of high (narrow Gaussian) and low (wide Gaussian) levels of certainty about sensory feedback, potentially arising from variations in environmental or sensory noise or from variations in attention. As shown in *Materials and Methods*, different choices of the underlying distributions result in essentially the same fits and predictions. This suggests that the heavy tails themselves, rather than their detailed shape, are crucial for the model.

Another extension of this work would be to use this framework to account for interindividual differences in behavior and neural activity [36, 66], which are not easily addressable given current experimental limitations. Like in many other behavioral modeling studies [37, 53], the present version of our model can fit only an average animal (because of our need to aggregate large datasets to accurately estimate behavioral distributions), making our results semiquantitative with respect to the statistics of any particular individual. Nevertheless, our framework represents a class of model that can explain a broader range of qualitative results than previous efforts, including, in particular, the shape of the distribution of exploratory behaviors.

Finally, while we used Bengalese finches as the subject of this study, nothing in the model relies on the specifics of the songbird system. Our approach might, therefore,

also be applied to studies of sensorimotor learning in other model systems, and we predict that any animal with heavy-tailed distribution of motor outputs should exhibit similar phenomenology in its sensorimotor learning. Exploring whether the model allows for such cross-species generalizations is an important topic for future research, as are questions of how networks of neurons might implement such computations [14, 30, 46, 78].

2.5 Materials and methods

2.5.1 Experiments

The data used are taken from the experiments in Ref. [94] and is described in detail there. Briefly, subjects were nine male adult Bengalese finches (females do not produce song) aged over 190 days. Lightweight headphones and microphones were used to shift the perceived pitches of birds’ own songs by different amounts, and the pitch of the produced song was recorded. For each day, only data from 10 am to 12 pm is used. The same birds were used in multiple (but not all) pitch shift experiments separated by at least 32 days. Changes in vocal pitch were measured in semitones, which is a relative unit of the fundamental frequency (pitch) of each song syllable:

$$\text{pitch in semitone} \approx 1.2 \log_2 \frac{\text{syllable frequency}}{\text{mean of baseline syllable frequency}}.$$

The error bars reported for the group means in Fig. 2.2 indicate the error of the mean that accounts for variances both across individual birds and within one bird. Specifically, if y_i^μ represents the pitch of the i th vocalization from the μ th bird on a specific day ($i = 1, \dots, n$, $\mu = 1, \dots, m$), then the mean pitch for the day for each bird is $\bar{y}^\mu = \frac{1}{n} \sum_i y_i^\mu$, and the global mean pitch is $\bar{y} = \frac{1}{m} \sum_\mu \bar{y}^\mu$. With these, we define the error of the mean used in Fig. 2.2 as

$$\delta y = \sqrt{\frac{\sum_\mu (\bar{y}^\mu - \bar{y})^2}{(m-1)^2} + \frac{1}{m} \sum_\mu \left[\frac{\sum_i (y_i^\mu - \bar{y}^\mu)^2}{(n-1)^2} \right]}, \quad (2.3)$$

where the first term in the square root represents the variance across birds, and the second term is the variance within one bird, averaged over the birds.

2.5.2 Stable distributions

A probability distribution is said to be stable if a linear combination of two variables distributed according to the distribution has the same distribution up to location and scale [75]. By the generalized central limit theorem, the probability distribution of sums of a large number of i. i. d. random variables with infinite variances tend to be stable distributions [75]. A general stable distribution does not have a closed form expression, except for three special cases: Lévy, Cauchy and Gaussian. A symmetric stable variable x can be written in the form $x = \gamma y + \mu$, where y is called the standardized symmetric stable variable and follows the following distribution [75]:

$$f(y; \alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} du e^{-|u|^\alpha} \cos(yu). \quad (2.4)$$

Thus any symmetric stable distribution is characterized by three parameters: the type, or the tail weight, parameter α ; the scale parameter γ ; and the center μ . α takes the range $(0, 2]$ [75]. If $\alpha = 2$, the corresponding distribution is the Gaussian, and if $\alpha = 1$, it is the Cauchy distribution. γ can be any positive real number, and μ can be any real number. The above integral is difficult to compute numerically. However, due to the common occurrence of stable distributions in various fields, such as finance [67], communication systems [73], and brain imaging [84], there are many algorithms to compute it approximately. We used the method of Ref. [4]. In this method, the central and tail parts of the distribution are calculated using different algorithms: the central part is approximated by 96-points Laguerre quadrature and the tail part is approximated by Bergstrom expansion [5].

Note that even though we take the propagator and the likelihood distributions as stable distributions in our model, their iterative application (effectively, a *product* of many likelihood distributions iterated with a *convolution* with the kernel), as well

as truncation, result in finite variance predictions, allowing us to compare predicted variances of the behavior with experimentally measured ones.

2.5.3 Fitting

Our model consists of three *truncated* stable distributions, one for each of the two likelihood functions resembling the feedback modalities and a third for the propagation kernel. We use truncation to ensure biological plausibility: neither extremely large errors nor extremely large pitch changes are physiologically possible. We truncate the distributions to the range $[-8, 8]$ semitones — much larger than imposed pitch shifts and slightly larger than the largest observed pitch fluctuations in our data, 7 semitones. This leaves us with 9 parameters of which we need to fit 6 from data, namely the type parameters α and the scale parameters γ , while the center parameters μ are predetermined: the two likelihoods are at 0 and Δ respectively, while the propagation kernel is centered around the previous time step value (see (2.1)). The prior (and accordingly the posterior) is a discrete distribution with a resolution of 1600 bins covering equidistantly the entire support of the distribution ($[-8, 8]$ semitones). The value within each bin is computed by the Bayes formalism given in (2.2). As described earlier in the *Mathematical Model* subsection, the initial prior is given as the steady state distribution after repeatedly applying (2.2) to a uniform distribution. In other words, the prior is *not* an independent variable in the model, but it is determined self-consistently by the likelihoods and the kernel. The resulting prior distribution is compared with the empirical distribution to compute the quality of fit according to the objective functions described below. Furthermore we point out that it is essential to have a non-Gaussian kernel: with a Gaussian kernel the baseline ($t = 0$ prior distribution) is, essentially, Gaussian and would thus not match the data in Fig. 2.2(D). Similarly, the likelihood functions must be long-tailed otherwise the combined likelihood would not be bimodal as found to be empirically in Fig. 2.3(F),

and large perturbations would not be rejected - even if the prior was long-tailed (for example due to a long-tailed kernel).

We construct an objective function that is a sum of terms representing the quality of fit for each of the three data sets to our disposition: the χ^2 for four adaptations of the means to the respective constant shifts (Fig. 2.2A), the χ^2 for the adaptation of the mean to the staircase shifts (Fig. 2.2B), and the log-likelihood of the observed baseline pitch probability distribution (Fig. 2.2C). Using (2.3) to define the error of the mean, we calculate χ^2 as

$$\chi^2 = \frac{1}{T} \sum_{t=1}^T \frac{(\bar{\phi}_t - \bar{y}_t)^2}{\delta y(t)^2}, \quad (2.5)$$

where $t \in [1, T]$ represents the days for a specific experiment with total duration of T days, while $\bar{\phi}_t$, \bar{y}_t and $\delta y(t)$ represent the theoretical result, the mean, and standard error of the experimental data on day t , respectively. To make sure that all three terms contribute on about the same scale to the objective function, we multiply the baseline fit term by 10. We use this objective function because the data we fit is heterogeneous: for learning with a sensory perturbation, we use only the means and the error bars of the produced pitch curves, while for the baseline distribution, we use the whole distribution. We choose to do it this way because it is computationally intensive to calculate likelihoods for distributions of vocalizations for every day, and to do it repeatedly for parameter sweeps. Given that we are fitting only a handful of parameters, while the data sets are very constraining, we do not think that we loose in accuracy by fitting the summary statistics instead of performing a full maximum likelihood estimation.

The objective function landscape is not trivial in this case, and there is not a single best set of parameters. Figure 2.4 illustrates this by showing the quality of fit as a function of each pair of (α, γ) , while keeping the other four parameters fixed. There is a large subspace (a plateau or a long nonlinear valley, depending on the projection used) that provides similar fit values. In other words, the effective number

of important parameters is less than six. Thus choosing the maximum of the objective function and characterizing the error ellipsoid, or linear sensitivity to the parameters, to get the best-fit parameter values and their uncertainties is not appropriate. As suggested in the literature on Sloppy Models [35, 98], where such nontrivial likelihood landscapes are discussed, instead we focus on values and uncertainties of the fits and predictions themselves. For this, we sweep through the entire parameter space and, for each set of parameters $\vec{\theta} = \{\alpha_1, \gamma_1, \alpha_2, \gamma_2, \alpha_k, \gamma_k\}$, we calculate the value of the objective function $\mathcal{L}(\vec{\theta})$ and the corresponding fitted or predicted curve $f(\vec{\theta})$. Then for the mean fits/predictions (lines in Fig. 2.2, Fig. 2.3, and Fig. 2.5), we have

$$\langle f(\vec{\theta}) \rangle = \frac{\sum_{\vec{\theta}} e^{-\mathcal{L}(\vec{\theta})} f(\vec{\theta})}{\sum_{\vec{\theta}} e^{-\mathcal{L}(\vec{\theta})}}. \quad (2.6)$$

For the SDs, denoted by shaded regions in Fig. 2.2, Fig. 2.3, and Fig. 2.5, we have

$$\sigma_f^2 = \langle f(\vec{\theta})^2 \rangle - \langle f(\vec{\theta}) \rangle^2. \quad (2.7)$$

There are many ways of doing the sweep over the parameters. Here we choose first to find a local minimum (however shallow it is). Then for each parameter, we choose six data points on each side of the minimum, distributed uniformly in the log space between the local minimum and the extremal parameter values ($(0.2, 1.9]$ for each α and $[0.01, 8]$ for each γ). The extremal values avoid $\alpha = 0, 2$ and $\gamma = 0$, which are singular and dramatically slow down computations. Thus there are total of 13 grid points for each parameter, and the total of $13^6 \approx 4.8 \cdot 10^6$ total parameter samples.

2.5.4 Choice of the shape of distributions

For Fig. 2.2 and Fig. 2.3 in the main text, we have chosen stable distributions for L_1 , L_2 and p_{prop} . To investigate effects of this choice, we repeated the fitting and the predictions for different distribution choices. We consider a family of power law distributions $\propto 1/(1 + (\phi/\gamma)^{2\alpha})$ and a family of mixtures of Gaussians of different width $\rho N(0, \gamma^2) + (1 - \rho)N(0, \delta^2)$. Distributions in either family produce very similar fits

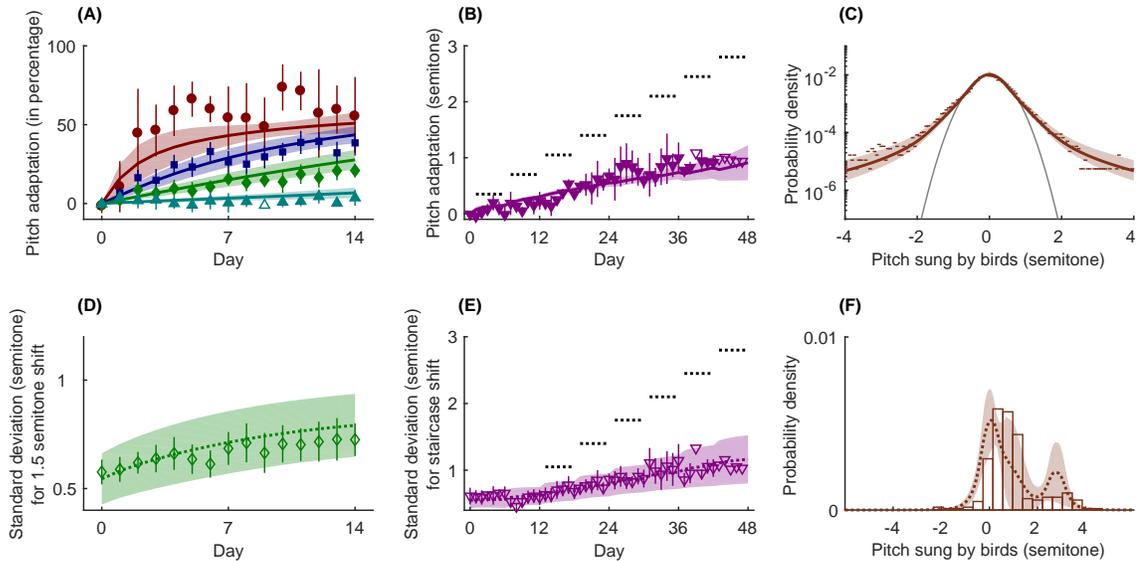


Figure 2.5: Fits and predictions with the power law family of heavy tailed distributions instead of the stable distribution family. (A-C) Equivalent to the panels in Figs. 2.2 A-C. (D-F) Equivalent to the panels in 2.3 C, E and F. The shaded areas around the theoretical curves represent confidence intervals for one standard deviation. The quality of all of the five fitted mean compensation curves combined is χ^2 per degree of freedom ≈ 1.56 , so that the truncated stable distributions used in the main text provide for (slightly) better fits. At the same time, the deviance of the fitted baseline distribution in (C) relative to the perfect fit, estimated as the NSB entropy of the data [69, 70], is 0.022 per sample point, slightly better than for the truncated stable distribution model, Fig. 2.2.

to the stable distribution model. For example, Fig. 2.5 shows the fits and predictions for the power law distribution model, and the power law family fits the means of the pitch compensation data slightly worse, but the baseline pitch distribution slightly better than the truncated stable distribution model, Fig. 2.2. The detailed shape of the distributions seems less important than the existence of the heavy tails.

2.5.5 Linear dependence on pitch shift in a Kalman filter with multiple time scales

We emphasized that traditional learning models cannot account for the nonlinear dependence of the speed and the magnitude of learning on the error signal. Here we show this for one such common model, originally proposed by Körding et al. [53]. This Kalman filter model belongs to the family of Bayes filters, which are dynamical

models describing the temporal evolution of the probability distribution of a hidden state variable (can be a vector or a scalar) and its update using the Bayes formula for integrating information provided by observations, which are conditionally dependent on the current state of the hidden variable. The specific attributes of a Kalman filter within the general class of Bayes filters [43] are the linearity of the temporal evolution of the hidden state (the pitch ϕ for the birds, but referred to as *disturbances* d in Ref. [53] and hereon), the linear relation between the measurements (observations) and the hidden variable, and the Gaussian form of the measurement noise and the distribution of disturbances.

One can argue that Kalman filter models with multiple time scales may be able to account for the diversity of learning speeds in our pitch shift experiments. We explore this in the context of an experimentally induced constant shift Δ to one disturbance d in the Kalman filter model with multiple time scales from Ref. [53]. If there is a constant shift Δ , Eq. (3) in Ref. [53] takes the form

$$O_t = \Delta + H \cdot d_t + W_t. \quad (2.8)$$

The first step in the Kalman filter dynamics is the prediction:

$$\langle d \rangle_{t+1|t} = A \langle d \rangle_{t|t}, \quad (2.9)$$

where $\langle d \rangle_{s|t}$ is the mean disturbance vector at time s given measurements up to time t and $A = \text{diag}(1 - \tau_i^{-1})$ with τ_i being the relaxation timescale of d_i . We assume that the shift occurs when the disturbances have relaxed to the steady state: $\langle d \rangle = 0$. Therefore, we approximate the standard Kalman filter equation describing the observation update of the expectation value of the disturbance after a measurement at time $t + 1$ as (see [43] for a detailed formal description)

$$\langle d \rangle_{t+1|t+1} = \langle d \rangle_{t+1|t} + \frac{\Sigma_{t+1|t} H^T}{H \Sigma_{t+1|t} H^T + R} (\Delta - H \cdot \langle d \rangle_{t+1|t}), \quad (2.10)$$

where R is the covariance matrix of the measurement noise, and Σ is the covariance matrix of the hidden variables. Σ does not depend on the measurement and is thus

not affected by the shift Δ . Thus the steady state prediction variance Σ_s is given by a solution to the equation

$$\Sigma_s = A \left(\Sigma_s - \frac{\Sigma_s H^T H \Sigma_s}{H \Sigma_s H^T + R} \right) A^T + Q, \quad (2.11)$$

where A is the matrix determining the temporal evolution of the mean disturbances, (2.9), and Q is the covariance matrix of the intrinsic (temporal evolution) noise.

From (2.11) we see that Σ_s is constant if the perturbation occurs when the system was at the steady state. We now wish to find the new steady-state given the constant perturbation Δ . Consider, for simplicity, two disturbances, each with its own temporal scale $n = 2$. The components of the steady state covariance are

$$\Sigma_s = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}, \quad (2.12)$$

and we define

$$\begin{aligned} f_1 &= \frac{\Sigma_{11} + \Sigma_{12}}{\Sigma_{11} + 2\Sigma_{12} + \Sigma_{22}}, \\ f_2 &= \frac{\Sigma_{12} + \Sigma_{22}}{\Sigma_{11} + 2\Sigma_{12} + \Sigma_{22}}. \end{aligned} \quad (2.13)$$

Substituting (2.9) in (2.10) we get

$$\begin{aligned} \begin{bmatrix} \langle d_1 \rangle_{t+1|t+1} \\ \langle d_2 \rangle_{t+1|t+1} \end{bmatrix} &= \begin{bmatrix} 1 - \tau_1^{-1} & 0 \\ 0 & 1 - \tau_2^{-1} \end{bmatrix} \begin{bmatrix} \langle d_1 \rangle_{t|t} \\ \langle d_2 \rangle_{t|t} \end{bmatrix} \\ &+ \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} (\Delta - (1 - \tau_1^{-1})\langle d_1 \rangle_{t|t} - (1 - \tau_2^{-1})\langle d_2 \rangle_{t|t}). \end{aligned} \quad (2.14)$$

In the steady state, $\langle d \rangle_{t+1|t+1} = \langle d \rangle_{t|t} = d_s$, we get

$$d_s^{(1)} = \Delta \frac{f_1 \tau_1}{1 + f_1(\tau_1 - 1) + f_2(\tau_2 - 1)} \quad (2.15)$$

$$d_s^{(2)} = \Delta \frac{f_2 \tau_2}{1 + f_1(\tau_1 - 1) + f_2(\tau_2 - 1)} \quad (2.16)$$

Thus we find that the sum of the disturbances is *proportional* to Δ *independent* of the size of Δ even for systems with multiple time scales.

Generalizing the result to n disturbances with different time scales, we get the following equations at steady state:

$$\tau_i^{-1}d_s^i = f_i\Delta - f_i \sum_{j=1}^n (1 - \tau_j^{-1})d_s^j \quad (2.17)$$

These equations are solved by

$$d_s^i = \Delta \frac{f_i\tau_i}{1 + \sum_{j=1}^n f_j(\tau_j - 1)} \quad (2.18)$$

which generalizes the linear dependence of learning on Δ for arbitrary n . Thus this (and similar) Kalman filter based model cannot explain the experimental results studied here.

Chapter 3 Sensorimotor learning as a 2d Bayesian filter

3.1 Introduction

An animal continuously interacts with the ever-changing environment, and incorporates its past experiences into its current beliefs about the world [22]. For motor control, these beliefs encapsulate expectations about which motor commands are likely to result in the desired motor output [104, 106], and the beliefs are updated based on the sensory feedback following execution of motor behaviors. Understanding the computation performed by the animal to update its beliefs and the neurophysiological mechanisms involved in this computation are the central questions in the field of sensorimotor learning.

Experimental studies of sensorimotor learning usually proceed by imposing perturbations on the sensory feedback or on the motor plant. The ensuing adaptation to the perturbation then reveals information about the function and the structure of the sensorimotor learning system. However, in essentially all analyses of sensorimotor learning in different animal systems, the adaptation has been monitored – and modeled – as a mean change of some behavioral characteristic, averaged over multiple renditions of the behavior in a single animal, and over multiple animals [53, 92, 106]. Where variance of the behavior is analyzed, it is usually not a dynamical variable. And yet, recent work has shown that the shape of the entire distribution of behavior – the mean, the variance, and maybe even higher moments – is crucial for animal performance and dynamically changes over time [19, 37, 48, 106]. Our goal in this paper is to present a theory that is capable of explaining the dynamics of the entire probability distribution of animal behavior during learning, focusing on pitch adap-

tation in the Bengalese finch as a model system. We build the theory based on the Bayesian filter model with non-Gaussian likelihood terms that are recently introduced [106]. The previously-introduced model has been able to account for many salient features of the songbird sensorimotor adaptation, including: (i) the non-monotonic dependence of the speed and the magnitude of adaptation on the perturbation size (small perturbations are compensated proportionally, while large ones are neglected), (ii) the ability of the animal to compensate even very large perturbations provided they were introduced gradually, and (iii) the observed dynamics of the behavioral variance during the adaptation (for some perturbations, the variance increased, and for some it stayed constant). Crucially, the model has not been powerful enough to model the entire probability distribution of the behavior, and hence only the mean adaptive behavioral dynamics were fitted by the model. Since the non-Gaussian tails are so important to the model success, the ability to account for the dynamics of the entire probability distribution is an essential test of the validity of the approach.

The situation is, in fact, even more dramatic: currently, there are no known models – whether developed by us or by other groups – that are able to track and predict complex, dynamic probability distributions of animal behaviors, while at the same time many models rely on the probabilistic nature of the behavior as a mechanism for learning. The goal of our contribution is to rectify this problem.

To endow the Bayesian filter theory with the ability to account for the dynamics of distributions, we note that experiments on different model systems, and within different behavioral paradigms, have observed that sensorimotor adaptation happens on multiple time scales. For example, relearning of a perturbation is usually faster than the initial learning [16, 52, 92]. Similar effects have been observed on even longer time scales, showing that animals exposed to relevant learning experiences as juveniles tend to learn better in the adulthood than adults without the experience [51]. Additionally, many studies focusing on adaptation back to the baseline following

removal of a perturbation consistently report that this so called *washout* involves at least two distinct time scales: a fast immediate decrease of the adapted response is usually followed by a slow, prolonged tail [40, 93, 100].

A simple one-dimensional dynamical system with a single timescale cannot generate these phenomena, arguing for the need of more complex, multiscale models of sensorimotor learning. In fact, there is evidence that even individual neurons and neural networks operate on diverse timescales [7, 13, 68, 83]. Correspondingly, many computational models have been proposed to study how multiple timescales can give rise to the above experimental observations. These models have been successful at explaining phenomena such as the faster relearning and the long decay tails [53, 92], but they usually assume a linear relation between the learning and error size, where a smaller (larger) error induces less (more) learning.

However, as has been shown by many experiments, biology does exactly the opposite: a smaller (larger) error generally induces more (less) learning [81, 94, 102]. Moreover, it has been suggested that the motor variability contains meaningful biological information that is crucial for learning, conflicting with the common thought that the observed variance in motor outputs is due to biological noises [38]. For example, the motor variability has been found to be positively correlated with animal's ability to learn in a new environment [105]. This feature echoes the exploration-exploitation trade-off in the reinforcement learning algorithms, where proper amount of uncertainty in action selection helps the convergence of the value function to the global ground state [97]. Another example is that the variance of the motor outputs can be voluntarily decreased or increased by the animal itself, which means that the motor variability is a well-controlled biological feature [45, 71]. However, most previous models only focus on the dynamics of the mean behavior without considering how the motor variability changes [53, 92]. Modeling the dynamics of higher moments or even the whole distributions can provide more perspective to understand the sensorimotor

system and its neural-behavior mappings.

Here, we introduce a full Bayesian model of sensorimotor learning with multiple timescales and test it on the vocal learning system of a songbird. The model assumes that there are many interacting hidden components that learn from the observed error, and that their linear weighted sum is the final output motor command. The model also assumes that the animal controls and updates the whole (non-Gaussian) distribution of motor commands [106]. Our model can explain the temporal evolution of the song distributions observed in experiments including both pitch shift data and washout data. In addition, the model makes various predictions that can be verified by future experiments.

3.2 Mathematical model

3.2.1 2d Bayesian filter

Let's first look at a physical system that shows the main idea of our model (Fig. 3.2A). There are two boxes with different masses on two frictionless surfaces. The box ψ has a larger mass, and the box ϕ has a smaller mass. We assume that the whole system does not have inertia, and the boxes only move when there are net forces acting on them. The two boxes are connected by a massless asymmetric nonlinear string (orange), which means that the string exerts different forces on ψ and ϕ , represented by $s^{(\phi)}$ and $s^{(\psi)}$, respectively. Each of the boxes has another massless nonlinear string (pink) attached to it, and a group of common external forces $s^{(i)}$, where $i = 1, 2, \dots$, are exerted on both of them. The over-damped surface guarantees that the movements of the two boxes have to be driven by the forces from the strings. Since the mass of the ϕ is smaller, the external forces $s^{(i)}$ will move it much faster than ψ . However, at the same time, the orange strings will try to pull the two boxes together.

In our model, at any time point t , we assume that the plausible motor commands are represented as a joint distribution of two interacting dynamical variables ψ_t and

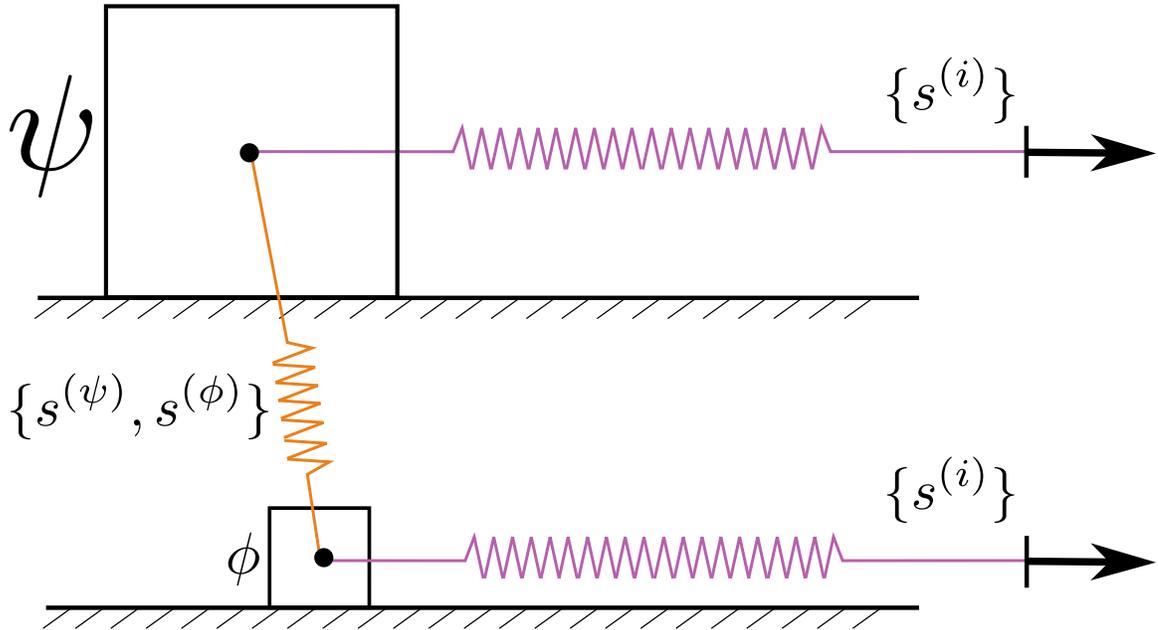


Figure 3.1: A physical analogy of the model. Suppose there are two boxes on two frictionless surfaces: one box with a large mass and denoted by ψ , and the other box with a small mass and denoted by ϕ . Imagine that the whole system is in liquid with low Reynolds number, and so there is no inertia. The movements of the boxes must be generated by forces. The external forces $\{s^{(i)}\}$, where $i = 1, 2, \dots$, act on both boxes through massless nonlinear springs (pink), and the lighter box ϕ will move faster than the heavier box ψ . The two boxes interact with each other through an asymmetric massless nonlinear spring (orange), where the force on ψ is generated by $s^{(\phi)}$ and the force on ϕ is generated by $s^{(\psi)}$.

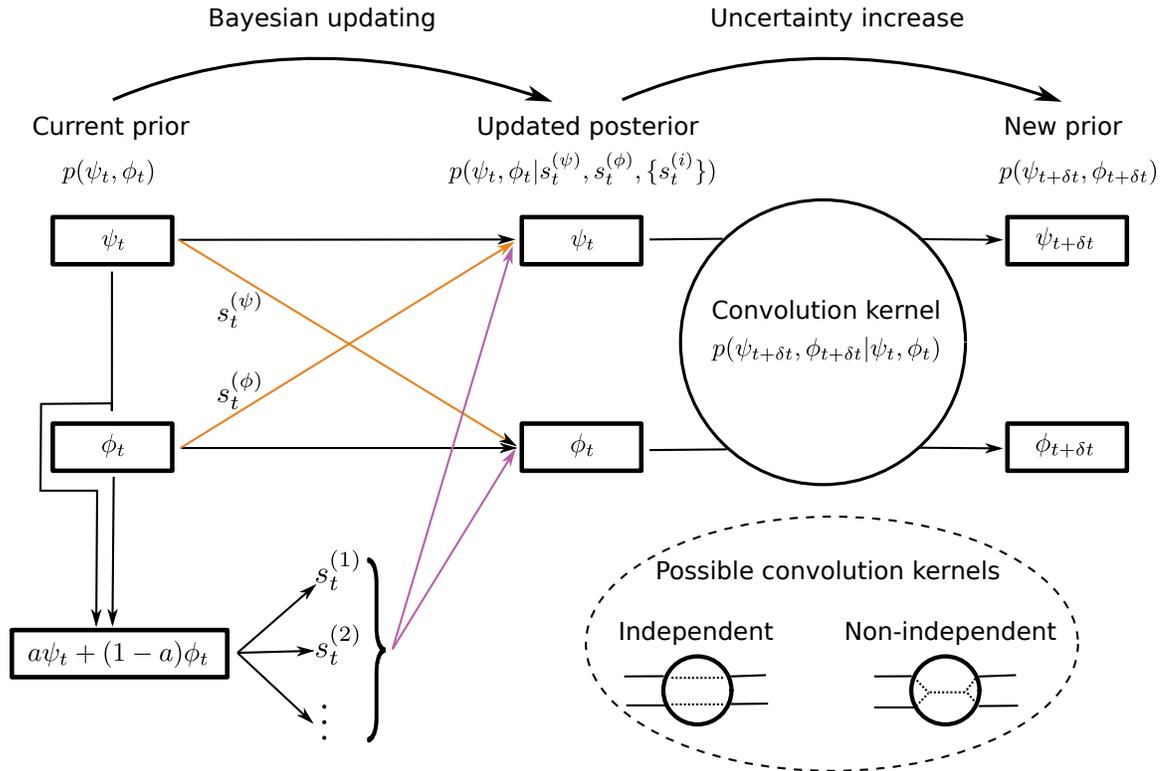


Figure 3.2: The dynamical Bayesian filter model. At each time step, a prior $p(\psi_t, \phi_t)$ is kept in memory by an animal. For each instance of motor performance, a sample pair (ψ_t, ϕ_t) is extracted from the prior $p(\psi_t, \phi_t)$, and their weighted sum $a\psi_t + (1-a)\phi_t$ is the motor output. This motor output is then measured by various sensory or non-sensory modalities, which form the feedback signals $s_t^{(i)}$, $i = 1, 2, \dots$ that are used in the Bayesian updating. At the same time, the two variables also send signals directly to each other: $s_t^{(\psi)}$ and $s_t^{(\phi)}$. All these signals are integrated with the prior using Bayes' Theorem, and we have the posterior $p(\psi_t, \phi_t | s_t^{(\psi)}, s_t^{(\phi)}, \{s_t^{(i)}\})$. This posterior is convolved with a kernel $p(\psi_{t+\delta t}, \phi_{t+\delta t} | \psi_t, \phi_t)$ to get the new prior $p(\psi_{t+\delta t}, \phi_{t+\delta t})$ at the next time step.

ϕ_t , each with its own timescale. The weighted sum of the two variables gives the final output of a motor command $a\psi_t + (1 - a)\phi_t$, where the parameter $a \in [0, 1]$ controls how much each variable contributes to the motor output. The joint distribution is denoted by $p(\psi_t, \phi_t)$, and serves as the animal's prior belief about the probability for each motor command to generate the desired sensory consequences. For applicable reasons, we assume that the two variables talk to each other by sending signals $s_t^{(\psi)}$ (from ψ_t to ϕ_t) and $s_t^{(\phi)}$ (from ϕ_t to ψ_t). This assumption is not justified yet in experiments, but it is very important for our model to work, since it serves as some regularization to constrain the values of ψ_t and ϕ_t . The weight a is assumed to be a constant here for simplicity, but more complicated models can have a weight that is updated every time step. The linear additivity of the two variables has been suggested by several experimental and computational studies of neuronal activities [47, 62]. A more complicated model with nonlinear summations is also possible, and we leave that for future investigations. As in our previous work [106], we take the motor command to be equivalent to the song pitches recorded in experiments, which means we don't distinguish between the motor command and its resulted behavior.

As shown in Fig. (3.2), at each time step, the motor system will sample a pair (ψ_t, ϕ_t) from the prior distribution, and send the motor command $a\psi_t + (1 - a)\phi_t$ to initiate a singing event (recorded as a song pitch). Then, this motor command or the song pitch can be measured by various sensory/non-sensory modalities [106], which generate a list of feedback signals $s_t^{(i)}$ ($i = 1, 2, \dots$). These feedback signals are sent back to the motor system to instruct the adaptation of the prior $p(\psi_t, \phi_t)$ in the form of likelihoods $p(\{s_t^{(i)}\} | a\psi_t + (1 - a)\phi_t)$. Besides, the interacting signals $s_t^{(\psi)}$ and $s_t^{(\phi)}$ are also received by the two variables, respectively, and can be incorporated into the prior as two likelihood terms $p(s_t^{(\phi)} | \psi_t)$ and $p(s_t^{(\psi)} | \phi_t)$. Thus, combining these

likelihoods with the prior using Bayes' Theorem, we have the posterior:

$$p(\psi_t, \phi_t | s_t^{(\phi)}, s_t^{(\psi)}, \{s_t^{(i)}\}) \propto p(s_t^{(\phi)} | \psi_t) p(s_t^{(\psi)} | \phi_t) p(\{s_t^{(i)}\} | a\psi_t + (1-a)\phi_t) p(\psi_t, \phi_t). \quad (3.1)$$

which is the updated memory of the two hidden variables based on the feedback data obtained from the current trial of the motor output. The posterior $p(\psi_t, \phi_t | s_t^{(\phi)}, s_t^{(\psi)}, \{s_t^{(i)}\})$ can be seen at the same time step as the current prior $p(\psi_t, \phi_t)$, which means that the time gap between the initiation of the song production and the Bayesian integration of the feedback is assumed to be much smaller than the time gap between two consecutive song productions. When there are no feedback inputs, an animal is assumed to become less certain about its belief of the world, and thus, the updated posterior should become wider and wider until the next sing event [106]. This increase of uncertainty can be modeled by a convolution kernel $p(\psi_{t+\delta t}, \phi_{t+\delta t} | \psi_t, \phi_t)$, which propagates the posterior to the new prior in the next time step. Thus, one complete iteration of the Bayesian filter is as follows:

$$p(\psi_{t+\delta t}, \phi_{t+\delta t}) = \frac{1}{Z} \iint p(\psi_{t+\delta t}, \phi_{t+\delta t} | \psi_t, \phi_t) p(s_t^{(\phi)} | \psi_t) p(s_t^{(\psi)} | \phi_t) \times p(\{s_t^{(i)}\} | a\psi_t + (1-a)\phi_t) p(\psi_t, \phi_t) d\psi_t d\phi_t, \quad (3.2)$$

where Z is the normalization constant.

The convolution kernel can be constructed in many different forms depending on specific questions one is asking. As shown in Fig. 3.2, the two variables ψ_t and ϕ_t can evolve either independently or non-independently. We will talk about the non-independent case later in the discussion section, and for now let's look at the independent case first, which means $p(\psi_{t+\delta t}, \phi_{t+\delta t} | \psi_t, \phi_t) = p(\psi_{t+\delta t} | \psi_t) p(\phi_{t+\delta t} | \phi_t)$. With this, the Eq. (3.2) becomes:

$$p(\psi_{t+\delta t}, \phi_{t+\delta t}) = \frac{1}{Z} \iint p(\psi_{t+\delta t} | \psi_t) p(\phi_{t+\delta t} | \phi_t) p(s_t^{(\phi)} | \psi_t) \times p(s_t^{(\psi)} | \phi_t) p(\{s_t^{(i)}\} | a\psi_t + (1-a)\phi_t) p(\psi_t, \phi_t) d\psi_t d\phi_t \quad (3.3)$$

If we take δt to be one day, and since during one day, the bird will sing thousands of song bouts, we need some averaging operation on those data received during a day [106]. Thus, we only consider averaged sensory signals: $s_t^{(\phi)} \rightarrow \bar{s}_t^{(\phi)}$, $s_t^{(\psi)} \rightarrow \bar{s}_t^{(\psi)}$, and $s_t^{(i)} \rightarrow \bar{s}_t^{(i)}$. The signals that the two variables send to each other are nothing but themselves, and thus, the average signals are actually the average values of the two variables: $\bar{s}_t^{(\phi)} = \bar{\phi}_t$ and $\bar{s}_t^{(\psi)} = \bar{\psi}_t$. For simplicity, we assume that there are only two feedback channels, $\bar{s}_t^{(1)}$ (channel 1, affected by headphones) and $\bar{s}_t^{(2)}$ (channel 2, not affected by headphones) [106]. Thus, these two averaged feedback terms can be written as $\bar{s}_t^{(1)} = a\bar{\psi}_t + (1-a)\bar{\phi}_t - \Delta$ and $\bar{s}_t^{(2)} = a\bar{\psi}_t + (1-a)\bar{\phi}_t$, respectively, where $-\Delta$ is the perturbation imposed through the headphones. The bird is learning to control its own song production, and the feedback signals reflect the performance of the last motor command produced by itself. Thus, for a proper control model, a minus sign is implied any feedback term; that is, if the feedback $\bar{s}^{(i)}$ is above (below) zero, then the corresponding likelihood term should drive the next plausible motor command down (up) [106]. We rewrite the four likelihood terms, $p(\bar{s}_t^{(\phi)}|\psi_t)$, $p(\bar{s}_t^{(\psi)}|\phi_t)$, $p(\{\bar{s}_t^{(1)}\}|a\psi_t + (1-a)\phi_t)$, and $p(\{\bar{s}_t^{(2)}\}|a\psi_t + (1-a)\phi_t)$ as $L(\psi_t; \bar{\phi}_t)$, $L(\phi_t; \bar{\psi}_t)$, $L(a\psi_t + (1-a)\phi_t; \Delta)$, and $L(a\psi_t + (1-a)\phi_t; 0)$, and the model shown in Eq. 3.3 can be rewritten as

$$p(\psi_{t+1}, \phi_{t+1}) = \frac{1}{Z} \iint p(\psi_{t+1}|\psi_t)p(\phi_{t+1}|\phi_t)L(\psi_t; \bar{\phi}_t)L(\phi_t; \bar{\psi}_t) \\ \times L(a\psi_t + (1-a)\phi_t; \Delta)L(a\psi_t + (1-a)\phi_t; 0)p(\psi_t, \phi_t)d\psi_t d\phi_t, \quad (3.4)$$

where $\bar{\phi}_t$, $\bar{\psi}_t$, Δ , and 0 represent the centers of the likelihoods. The likelihood $L(\psi_t; \bar{\phi}_t)$ means that ψ_t should get closer to $\bar{\phi}_t$; the likelihood $L(\phi_t; \bar{\psi}_t)$ means that ϕ_t should get closer to $\bar{\psi}_t$; the likelihood $L(a\psi_t + (1-a)\phi_t; \Delta)$ means that the motor output $a\psi_t + (1-a)\phi_t$ should be at Δ since there is a $-\Delta$ perturbation imposed on the feedback channel 1; the likelihood $L(a\psi_t + (1-a)\phi_t; 0)$ means that the motor output $a\psi_t + (1-a)\phi_t$ should be at 0 since the feedback channel 2 can genuinely measure the song pitches.

3.2.2 Fits and predictions

In the model described by Eq. 3.4, there are six distributions to be determined, the two kernels and the four likelihoods. The prior distribution does not contribute new degrees of freedom, which is simply the steady state of the Eq. 3.4 when the shift size Δ is 0. As we have pointed out previously [106], many distribution families work similarly well as long as they are heavy-tailed. Here, we use powerlaw-like distributions for the sake of convenience (Methods). For each of these distributions, there are three parameters, the exponent α , the scale γ and the mean μ . All the means for the six distributions have been fixed by the model itself, since for the kernels, the means are the variables in the previous step, and for the likelihoods, the means are $\bar{\phi}_t$, $\bar{\psi}_t$, Δ , and 0. Thus, including the parameter a that is used to control how much each variable contribute to the motor output and the cut-off point L of each distribution in the tail, we have 14 free parameters in total. The cut-off point L is fixed at 10 semitones, which includes all the song pitches in the current data set.

We fit the rest of the 14 free parameters in our model to more than 150 distributions (one distribution for each day) simultaneously using maximum likelihood, and the results are shown in Fig. 3.3. The pitch perturbation experiments end on day 14, and day 15 is the start of the washout period during which the perturbation is lifted. When the perturbation is on, the distributions of the song pitch (thin lines) on each day deviates gradually from the baseline data (gray shade), and after the pitch shift is lifted (washout period), the distribution gradually comes back to align with the baseline data (Fig. 3.3A). The staircase shift experiments last longer (in total 47 days), and the deviations from the baseline data are even more obvious. Our model (thick lines) can describe the dynamics of the distributions surprisingly well, and the quality of the fits can be justified by calculating the Jensen-Shannon divergence (JSD) between the data and the theoretical curves as shown in Fig. 3.4. The JSDs between the data and the theoretical curves are very close to zero (Fig.

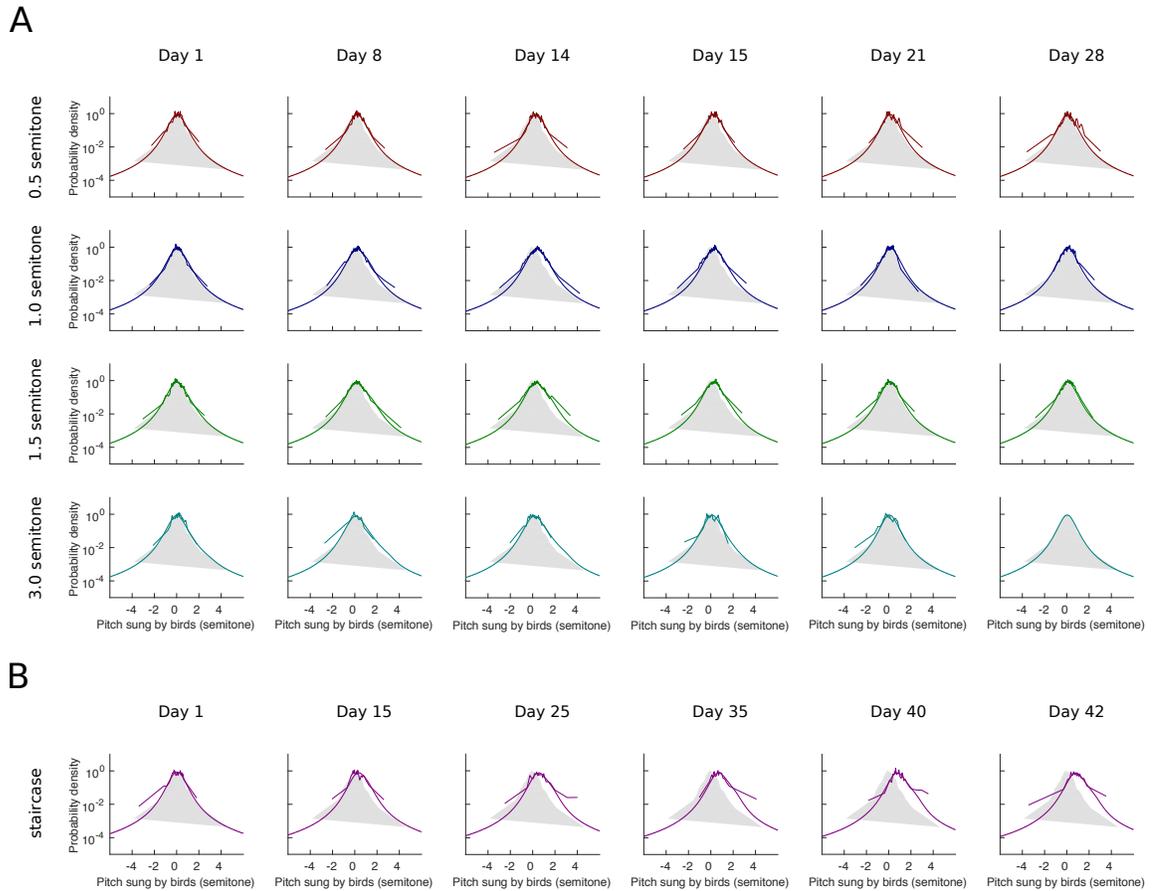


Figure 3.3: Experimental data and model fits. For each small panel, the gray-shaded region is the baseline data distribution obtained by collecting all the baseline data from all the birds. The thin line shows the data on a specific day, and the thick line shows the results of the model on the same day. The distributions from data are obtained by adaptive binning. (A) From top to bottom, the perturbation sizes are 0.5 semitone, 1.0 semitone, 1.5 semitone and 3.0 semitone. The first three columns show the results of perturbation periods, which end on Day 14. The last three columns show the results of washout periods (perturbation lifted), which starts on Day 15. There are no data from the 3.0-semitone group on Day 28. (B) In staircase experiments, there is no washout period, and the perturbation size is gradually increased by 0.35 semitone for every 6 days. Note that the following color code is used throughout the paper: red, 0.5-semitone constant perturbation; blue, 1.0-semitone constant perturbation; green, 1.5-semitone constant perturbation; cyan, 3.0-semitone constant perturbation; magenta, staircase perturbation.

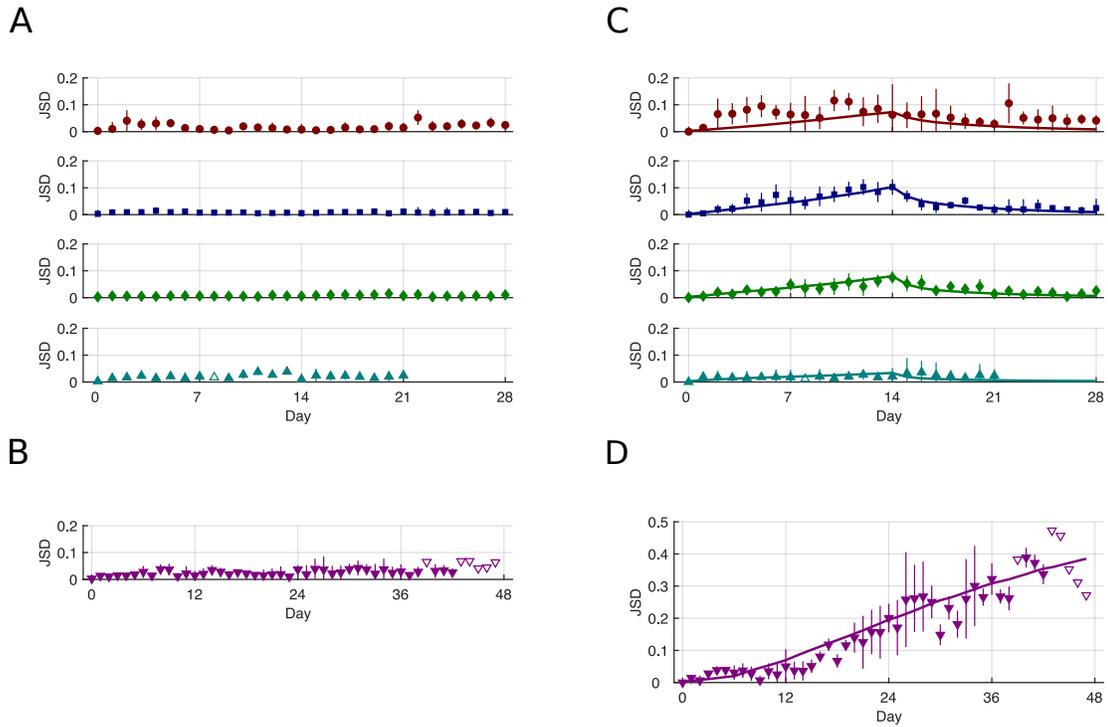


Figure 3.4: Jensen-Shannon divergence (JSD). (A) & (B) The dots represent the JSDs between the data and the model on the same day, which indicate the quality of the fits in Fig. 3.3. The error bars represent the standard deviation obtained from the bootstrapped sampling of the data, including both cross-bird and within-bird variability. Note that the symbols used here apply to the whole paper: circles, 0.5-semitone constant perturbation; squares, 1.0-semitone constant perturbation; diamonds, 1.5-semitone constant perturbation; up triangles, 3.0-semitone constant perturbation; down triangles, staircase perturbation; hollow symbols, the data on that day are not used for the fitting in Fig. 3.3. The mean of all dots here is 0.0161 ± 0.0155 . (C) & (D) The dots represent the JSDs between the baseline data and the data on a specific day, and the error bars represent the standard deviation obtained from the bootstrapped sampling of the data, including both cross-bird and within-bird variability. The solid lines represent the JSDs between the baseline data and the model on a specific day.

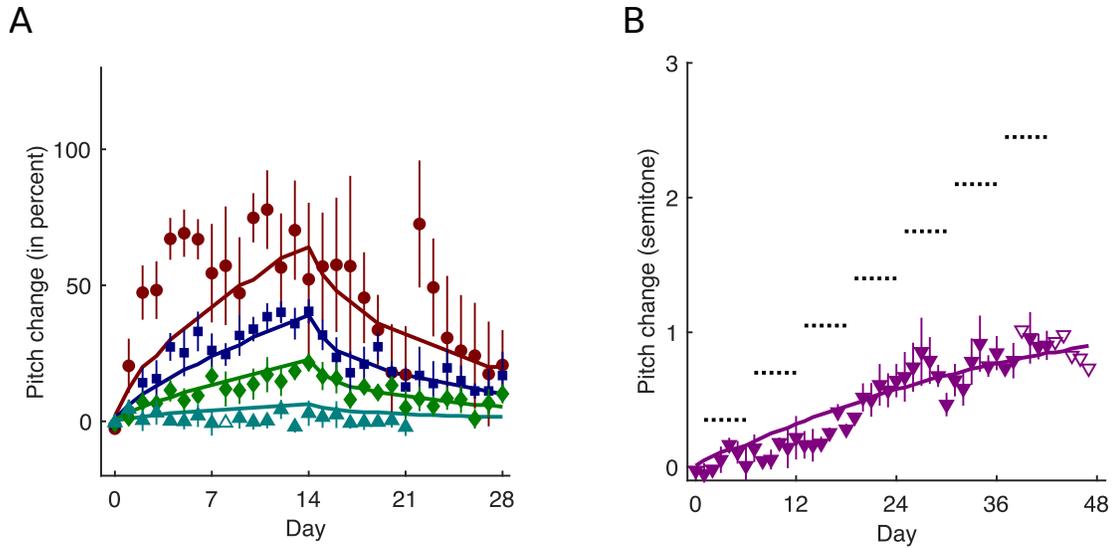


Figure 3.5: Medians. (A) The dots show the normalized medians of the distributions in Fig. 3.3A, and the error bars represent the standard deviation obtained from the bootstrapped sampling of the data, including both cross-bird and within-bird variability. The normalized medians are calculated by dividing the medians by the corresponding perturbation size and times 100, which indicate the percentage of the perturbation been learned. The solid lines represent the normalized medians from the model. (B) The dots show the medians of the distributions in Fig. 3.3B, and the error bars have the same meaning as in (A). The solid lines represent the medians from the model. The dotted lines in show the gradually increased perturbation size.

3.4A and 3.4B), indicating that our model has a very good description of the data. Fig. 3.4C and 3.4D shows the JSDs between the data and the baseline (dots with errorbars) and the JSDs between the model and the baseline (solid lines) for each day of the experiments. It can be seen that at the end of the shift experiments (day 14), both data and model deviates from the baseline significantly where the JSDs reach around 0.1 for both 1-semitone and 1.5-semitone experiments. This effect is even more obvious in the staircase experiment, where the JSD can reach above 0.4.

In Fig. 3.5, we plot the medians of the data and the model. After washout period starts (on day 15), there are prolonged decaying tails for 0.5-semitone, 1.0-semitone and 1.5-semitone shift sizes. One can argue that an exponential model with a very long timescale can also have similar tails, and thus, an one-timescale model should be enough. However, we fit our previous one-timescale model to all the distributions,

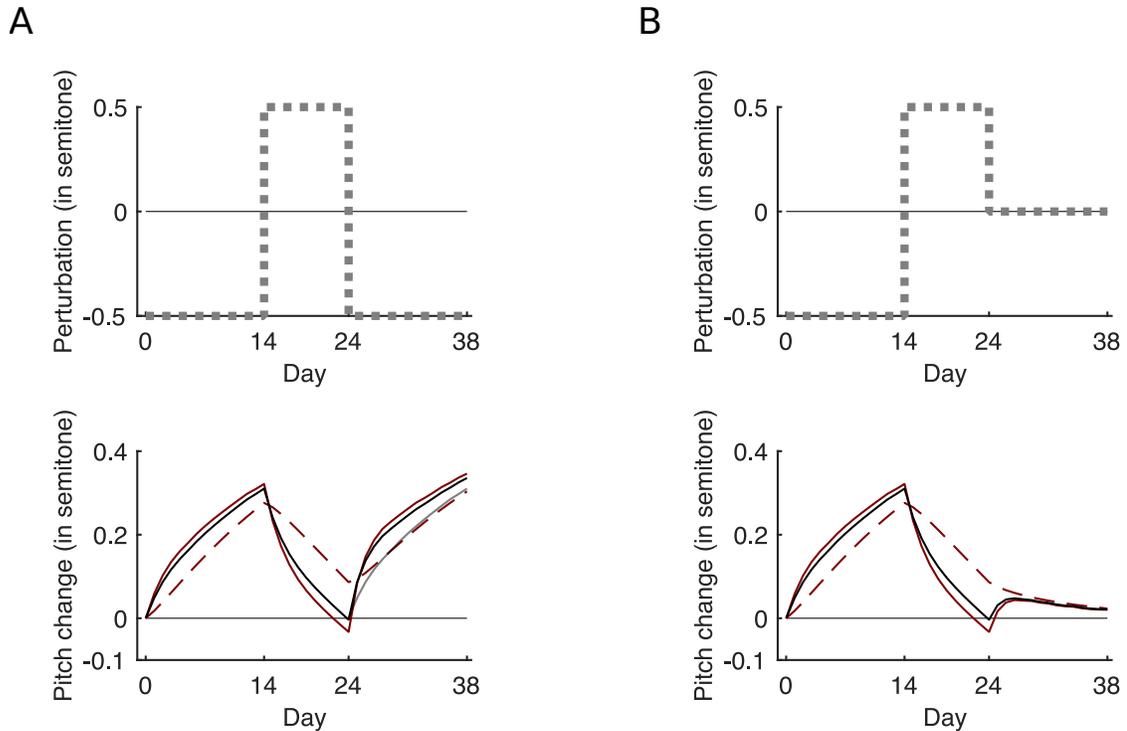


Figure 3.6: Model predictions for two three-phase experiments [92]. The parameters used in this figure are the same ones used in Fig. 3.3. (A) As shown by the dotted grey line in the upper panel, the perturbation is imposed as follows: -0.5 semitone for the first 14 days (phase 1), 0.5 semitone for the next 10 days (phase 2), and -0.5 semitone again for another 14 days (phase 3). The behavior, or the song pitch, is predicted by the black solid curve in the lower panel. The solid orange curve represents ϕ_t , and the dashed orange curve represents ψ_t . The solid grey curve between day 24 and day 38 is the horizontal translation of the solid black curve between day 0 and day 14. (B) The perturbation (upper panel) has a similar structure as in (A), and the only difference is that in the phase 3, the perturbation is 0 semitone. Line types and colors are the same as in (A).

and it turns out that it just cannot fit the fine structures of the tails nicely (3.13). Besides, it cannot fit the entire data set very well. For a comparison, the Bayesian information criteria (BIC) for the fit of the two-timescale model shown in Fig. 3.3 is around 4063463, and the BIC for the best fit of the one-timescale model [106] is around 4069086, which is 5622 higher.

With the fitted parameters, we are interested to see what type of predictions our model can make for new types of perturbation paradigms. A classical experimental design is the three-phase perturbations as shown in the upper two panels in the Fig.

More comments are needed here.

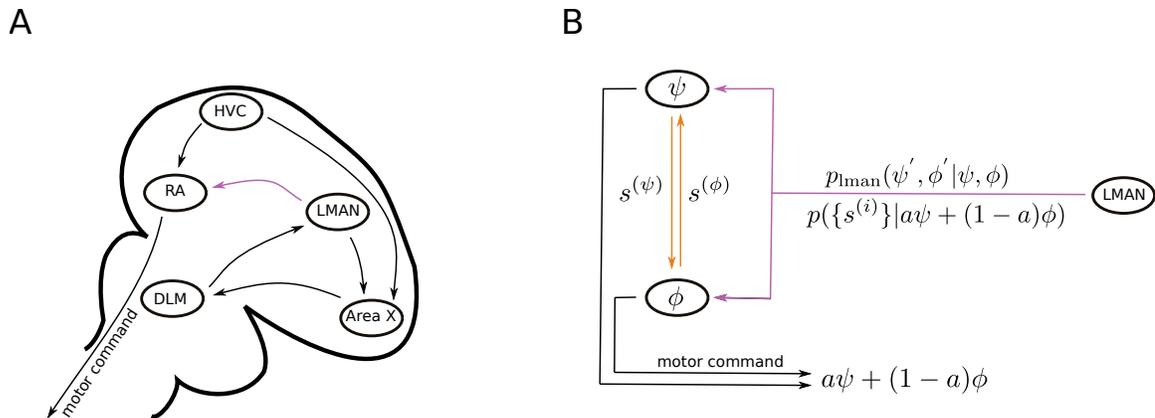


Figure 3.7: One possible biological interpretation of the model. (A) Schematic diagram of a songbird’s vocal system. The pink arrow shows the projection from LMAN to RA, and it has been suggested that this projection carries two effect of LMAN on RA: motor command bias and motor variability (B) The diagram shows one of the ways in which our model can be interpreted in the context of the songbird vocal system. Rather than a single motor command generator RA, we have two interacting components, ψ and ϕ , and their linear sum is the final motor output. The orange arrows represent the interactions between the two variables by exchanging information $s^{(\psi)}$ and $s^{(\phi)}$, which results in the synchronization of their activities. The pink arrow carries the bias $p(\{s^{(i)}\} | a\psi + (1-a)\phi)$ and variability $p_{\text{lman}}(\psi', \phi' | \psi, \phi)$ from LMAN, and projects them to both of the two motor components.

3.6. It can be seen in Fig. 3.6A that the relearning rate of the same perturbation size is obviously faster than the original learning. This is due to the fact that although at the end of the phase 2 the overall adaptation is already around 0 (solid black curve), the slower variable ψ_t is still way above 0, which will accelerate the relearning process. The non-zero slow process at the end of phase 2 is also the reason of the spontaneous rebound if in the phase 3 the perturbation is set to 0. This spontaneous rebound has been observed in arm-reaching experiments [92]. Here, we expect that the faster relearning and spontaneous rebound can also be observed in songbird’s pitch-perturbation experiments.

3.2.3 Biological interpretations

It should be noted that we want to keep our model as general as possible, and thus the model should not rely on any specific neural circuit structures. However, it

is still interesting to have a discussion on the possible interpretations of our model in the case of songbird vocal system. Let's first look at the main pathways that are involved in the songbird vocal executing and learning (Fig. 3.7A). HVC controls the temporal and sequential patterns of the song, i.e., the start and end points of each song syllable, and initiates singing behaviors [28]. At any singing event, HVC sends signals carrying temporal information of the syllables to robust nucleus of the arcopallium (RA). Then, RA fills in specific properties for each syllable, such as the pitch, and send them to downstream organs. RA also gets inputs from LMAN (the pink arrow), which is suggested to contribute to motor variability and learning [9, 12, 77, 85]. There are two phenomena observed if the projection from LMAN is blocked: (1) the distribution of the song bouts becomes very narrow [45, 77] and (2) the bird stops learning and does not adapt to new situations [9, 12]. Thus, it seems that LMAN contributes two effects: variability and a teaching signal (bias).

In our model (Fig. 3.7B), we assume that in the motor system, there are two components, ψ and ϕ , that can store past experience and contribute to the final motor command, and they tend to synchronize by sending signals ($s^{(\psi)}$ and $s^{(\phi)}$) to each other. The linear sum of the two components generates the motor command. However, the specific neural basis of this two components is not clear, and they might correspond to two separate subregions somewhere in the motor system, such as the robust nucleus of the arcopallium (RA). These subregions should have axon projections to each other to exchange information, which results in the synchronization between the behavior-relevant activities of the two subregions. We hope that future experiments could be done to justify these assumptions, and to reveal how they are implemented in the neuron and circuit levels [55].

Both of the two components receive signals from LMAN, and to compare our

model with experimental results, we rewrite the model in Eq. 3.2 as follows:

$$p(\psi_{t+\delta t}, \phi_{t+\delta t}) = \frac{1}{Z} \iint \left[\iint p(\psi_{t+\delta t}, \phi_{t+\delta t} | \psi'_t, \phi'_t) p_{\text{lman}}(\psi'_t, \phi'_t | \psi_t, \phi_t) d\psi'_t d\phi'_t \right] \\ \times p(s_t^{(\phi)} | \psi_t) p(s_t^{(\psi)} | \phi_t) p(\{s_t^{(i)}\} | a\psi_t + (1-a)\phi_t) p(\psi_t, \phi_t) d\psi_t d\phi_t. \quad (3.5)$$

Thus, as suggested by experiments, we assume that there are two types of signals: teaching signals indicated by the likelihood terms $p(\{s_t^{(i)}\} | a\psi_t + (1-a)\phi_t)$, which represent the processed sensory feedback (bias), and part of the convolution kernel $p_{\text{lman}}(\psi'_t, \phi'_t | \psi_t, \phi_t)$, which provides the motor variability that is associated with the exploration. It should be noted that the kernel models more than the variability from the LMAN, and it also contains other variance sources, such as the intrinsic noises in the motor pathway. Whether these two types of signals are sent together or separately is not clear [28], but some deafening studies show that LMAN still affects the song production even without auditory feedback [76]. This might suggest that the signals from LMAN to RA could be rather diverse, and different information could be sent separately. However, more experiments should be done in order to clarify the true nature of the LMAN-RA projections [28].

In Fig. 3.8, we show what our model predicts if the LMAN-RA projections are blocked, which means in Eq. 3.5 that the term $p(\{s_t^{(i)}\} | a\psi_t + (1-a)\phi_t)$ is deleted and the term $p_{\text{lman}}(\psi'_t, \phi'_t | \psi_t, \phi_t)$ is set to be a delta function, $\delta(\psi'_t - \psi_t, \phi'_t - \phi_t)$. For any specific day, we change our model in this manner for the previous day, and see its effect on learning, as shown by the red dots in Fig. 3.8A. Our model predicts a drop from the one in the normal condition (grey dots connected by grey lines). This effect is qualitatively comparable to the experimental results shown in Fig. 3.8B, where they tested how much the LMAN contributes to learning on a daily basis in a white noise experiment [1].

In our Bayesian framework, the two dynamical variables, ψ_t and ϕ_t , represent certain neural states with their own specific timescales. The biological interpretation of our model shown in Fig. 3.8 assumes that both of the two variables are related

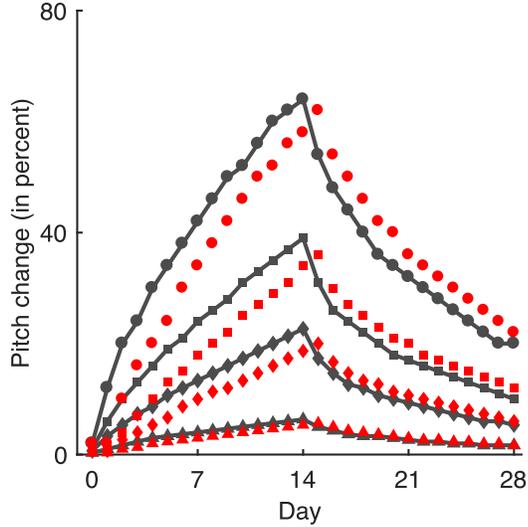


Figure 3.8: Model-predicted LMAN contribution to learning in 1 day. Grey dots and lines are the same as in Fig. 3.5 when the LMAN is intact, and red dots represent what will happen if in Eq. 3.5 p_{lman} is set to be a delta function (no variability from LMAN) and $p(\{s_t^{(i)}\} | a\psi_t + (1-a)\phi_t)$ is deleted (no teaching signals) in the previous day.

to RA or other relevant regions involved in the motor pathway. Thus, the activity of LMAN is not explicitly represented by any dynamical variable, and the LMAN’s contributions are modeled as part of the convolutional kernel, which is static, and as the sensory-related likelihood terms, which have the timescale as the feedback data. This means that our model does not have the intrinsic timescale of the neural activities in LMAN. However, an alternative interpretation is to assume that the activity of LMAN is explicitly represented by the fast variable ϕ_t . This variable receives feedback information directly from sensory modalities (likelihood terms). However, the problem is how the slow variable ψ_t learns. If we assume that the motor pathway or RA only learns through LMAN, then the slow variable ψ_t should be only affected by the fast variable, but not by the sensory likelihoods. In this case, the motor output should not explicitly depend on ψ_t , since the feedback information of the motor output is only integrated into the LMAN activity ϕ_t . Thus, it seems that our model is hard to be interpreted in the alternative way, and a proper LMAN-

specific Bayesian model is needed.

3.3 Discussion

Learning from the past experience is essential for animals to shape and improve their skills to deal with the complicated environments that are changing dynamically. Thus, how the nervous system represents previously accumulated data and how to integrate new ones are important and interesting questions to answer in order to understand why our brain can perform its functions successfully. Behavioral studies have their advantage in that the quantities measured are usually the ones that interact with the environment directly, and the generative models developed based on them may capture the relevant features of the true neural algorithms. Here, we use song-bird's song pitch data collected from pitch perturbation experiments to investigate the general computational rules that underlie motor representation and learning. To capture the rich dynamics shown by the data, we introduced a mathematical model under the general Bayesian framework. With various hypotheses made about the motor system and only 14 parameters, our model can simultaneously describe in high precision the dynamics of more than 150 distributions of the song pitches.

In our model, we assume that multiple components in the motor system are learning from the experience at different speeds. At the same time, they also send signals to each other, and these interactions resulted in a synchronization of their activities. The weighted linear sum of these variables produces a motor command that is sent to instruct muscle movements. We also assume that the representation of these variables is not a list of scalars, but rather a whole joint probability distribution of them is stored in the motor system. The probability shows the animal's confidence about how likely these variables can produce the desired sensory feedback. Thus, our model is able to trace the dynamics of the whole distribution of the motor output. This provides a powerful tool to study how motor variability is gradually shaped by changes

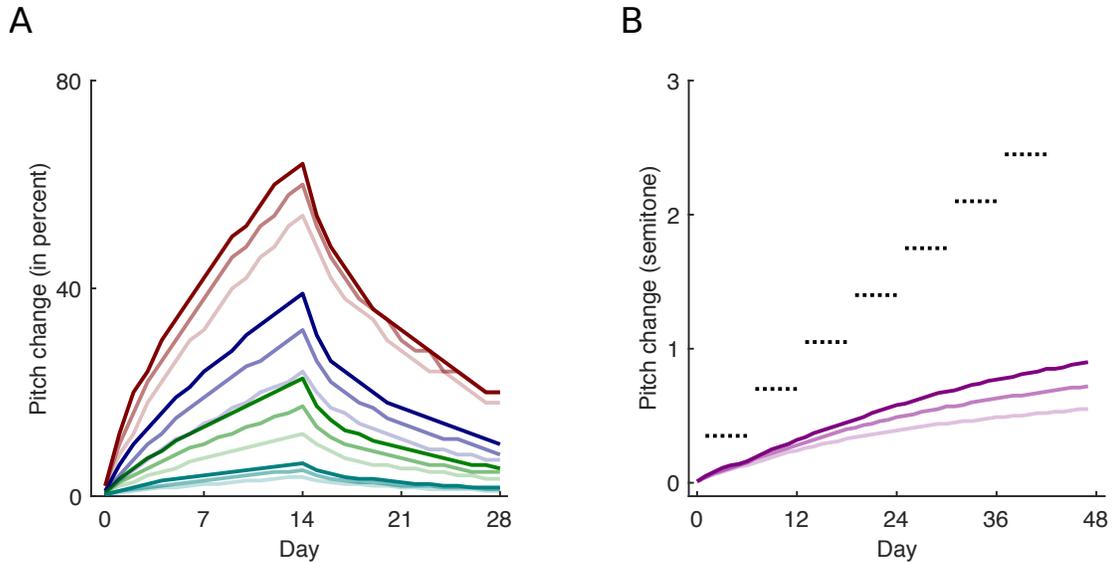


Figure 3.9: Relations between variability and learning. (A) & (B) The solid lines are the same as the solid lines in Fig. 3.5, and the colored dotted lines are obtained by turning off the LMAN on that day. (C) & (D) The transparency of the lines indicate how much the variance is reduced, which is done by multiplying the scale parameters in the two kernels in Eq. 3.4 by a fractional number r . For each color, the normal curve has $r = 1$; the more transparent curve, $r = 0.8$; the most transparent curve, $r = 0.6$.

in the environment. This is interesting, since it has been suggested that motor variability is an essential part of the exploration-exploitation trade off of any learning agent. Indeed, many experimental results have already shown that variability in the motor outputs is not useless noise but rather a well controlled biological signal [45, 71]. What’s more, a positive correlation between learning and variance has already been shown in experiments [105]. In fact, in agreement with experiments, our model predicts that learning is slower for smaller variance, if the variability is reduced (Fig. 3.9).

From data, we have shown that the distributions of song pitches illustrate a clear non-Gaussian structure. Surprisingly, it is this non-Gaussian statistics that explains the negative correlation between the error size and the learning speed/magnitude [106]. It will be interesting to see whether this heavy-tailed distributions also exist in other motor systems and animals.

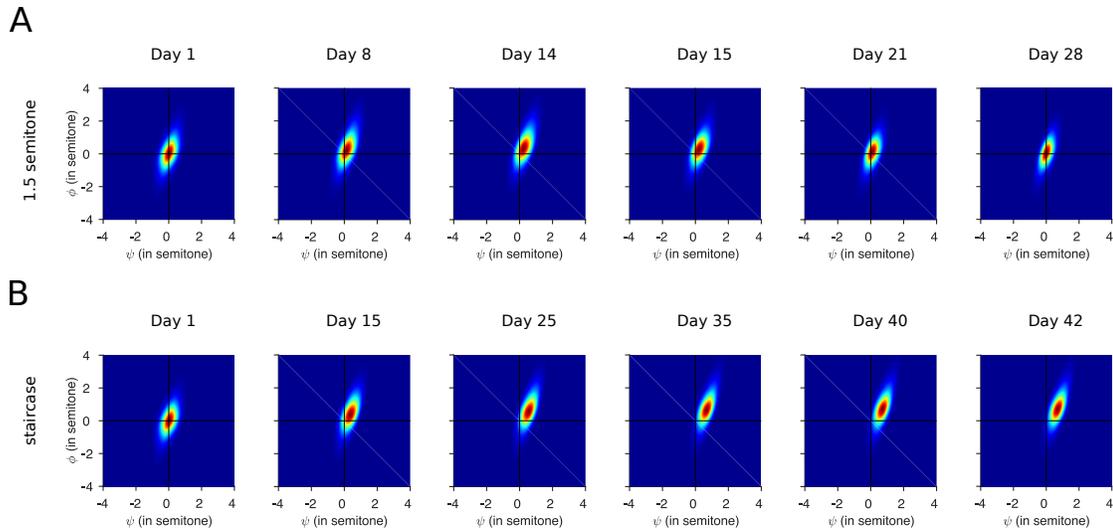


Figure 3.10: Predicted joint distributions from the model where the two variables, ψ_t and ϕ_t , evolve non-independently with a kernel shown in Eq. 3.9. (A) The shift size is 1.5 semitone. As in Fig. 3.3A, Day 1 is the first day of the shift experiment, Day 14 is the last day of the shift experiment, and Day 15 is the first day of the washout period. (B) As in Fig. 3.3B, the shift size is increased by 0.35 semitone for every 6 days.

The nature of the two neural components assumed in our model is an interesting topic for future experimental investigations. In the vocal system of songbird, whether they are both part of the RA region or there exist new regions in the brain participating motor learning that has not been identified yet. Do other motor systems or other animals have similar structures in the brain? Answering those questions might give us more insights on the motor learning and further our understanding of brain functions.

The construction of the convolution kernel is variable (Fig. 3.2), and besides the independent kernel we have discussed, one alternative of the model is a kernel that evolve the two variables non-independently, which has the form shown in Eq. 3.9. The matrix Σ has three free parameters, which can be used to describe the covariance between the two components. This model can fit the data equally well (Fig. 3.24). The dynamics of the joint distributions is shown in Fig. 3.10. Behavioral observations such as the song pitch are the final outputs of the motor system, and to obtain the

information of the two hidden units, neuronal recordings might be used. Thus, what the joint distributions reflect may be the neuronal activities in the motor pathway or its latent representations [15].

Another alternative of the model is when the weight $a = 0$, which transforms Eq. (3.3) into the following form

$$\begin{aligned} p(\psi_{t+\delta t}) &= \frac{1}{Z_\psi} \int p(\psi_{t+\delta t}|\psi_t)p(s_t^{(\phi)}|\psi_t)p(\psi_t)d\psi_t \\ p(\phi_{t+\delta t}) &= \frac{1}{Z_\phi} \int p(\phi_{t+\delta t}|\phi_t)p(s_t^{(\psi)}|\phi_t)p(\{s_t^{(i)}\}|\phi_t)p(\phi_t)d\phi_t \end{aligned} \quad (3.6)$$

This introduces a whole new model, where only one of the variables ϕ_t contributes to the final motor output and learns from data. The other variable ψ_t can only get access to data indirectly through ϕ_t . The dynamics of the two variables are shown in Fig. 3.11. This architecture might be useful to model the memory consolidation. during the initial learning of a new motor skill, two processes happens: the quick acquisition of this motor skill in one brain region and the gradual consolidation or transfer of the learned skill to other brain regions [31, 57, 90]. This also resembles some recently developed AI algorithms [41].

In summary, the presented model provides a powerful framework that can be potentially used to tackle a wide range of learning problems, not just in the case of animal learning, but might also be useful for designing new AI algorithms.

3.4 Materials and methods

3.4.1 Experimental data

There are in total nine birds used across the five different groups of experiments, and some birds have been used multiple times. Here, if a bird has been used in two experiments, we count it as two effective birds, and in this way, there are in total 23 birds. They are distributed as follows: the 0.5-semitone group has three birds, the 1.0-semitone group has eight birds, the 1.5-semitone group has six birds, the

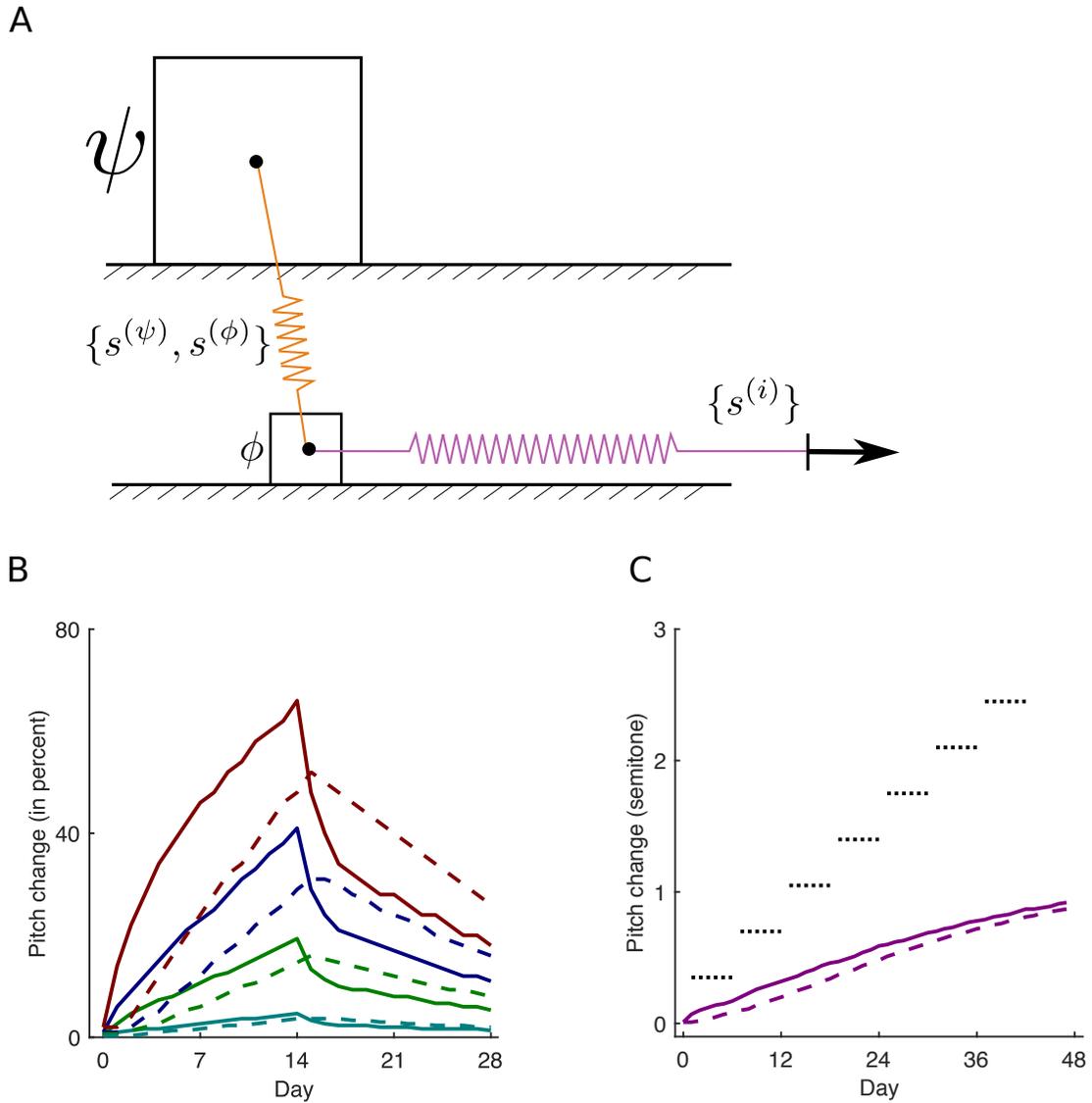


Figure 3.11: (Cascade model. (A) When $a = 0$, the feedback forces $\{s^{(i)}\}$ only act on the variable ϕ . (B) & (C) The medians of the two variables are shown, by simulating the cascade model Eq. 3.6. The solid lines are the medians of ϕ_t and the dashed lines are the medians of ψ_t .

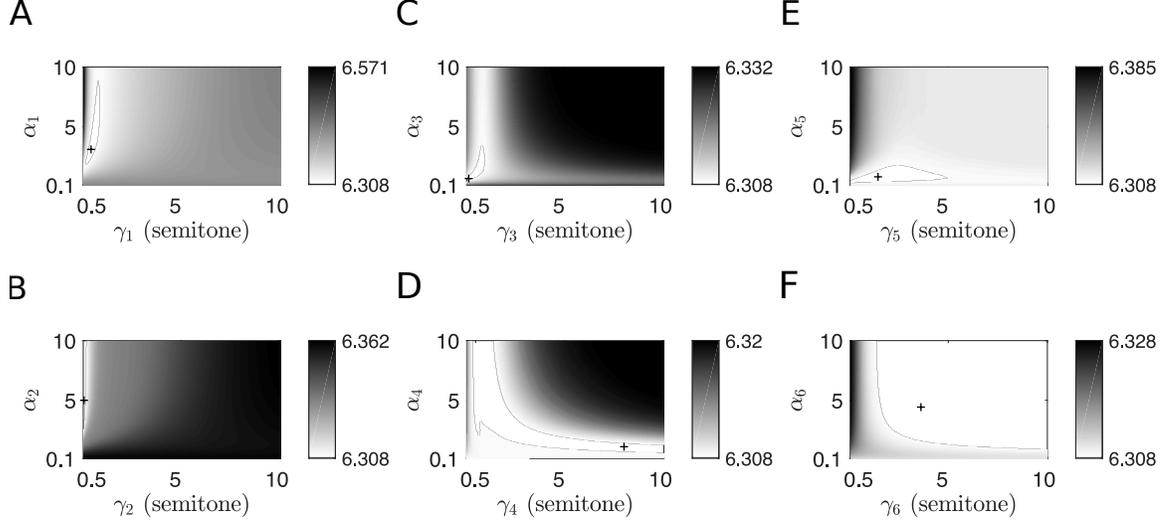


Figure 3.12: The landscapes of the objective function for the model shown in Eq. 3.4. (A-F) Each panel shows how the objective (in log 10 scale) varies with respect to a certain pair of parameters, while all the other parameters are fixed. Each pair of parameters shown as the two axes for each panel represents a distribution used in the model: α_1 and γ_1 represent $p(\psi_{t+\delta t}|\psi_t)$, α_2 and γ_2 represent $p(\phi_{t+\delta t}|\phi_t)$, α_3 and γ_3 represent $L(\psi_t; \bar{\phi}_t)$, α_4 and γ_4 represent $L(\phi_t; \bar{\psi}_t)$, α_5 and γ_5 represent $L(a\psi_t + (1-a)\phi_t; \Delta)$, α_6 and γ_6 represent $L(a\psi_t + (1-a)\phi_t; 0)$.

3.0-semitone group has three birds, and the staircase group has three birds.

3.4.2 Data bootstrapping

We do bootstrapping in two steps: first, for each group, we bootstrap the birds; second, for each bird, we bootstrap the recorded song pitches. In this way, we take into account both across-bird and within-bird variations.

3.4.3 Simplification of representations

The complete forms of the prior and the posterior should be written as

$$\text{prior} : p(\psi_t, \phi_t | s_{t-\delta t}^{(\phi)}, \dots, s_0^{(\phi)}; s_{t-\delta t}^{(\psi)}, \dots, s_0^{(\psi)}; \{s_{t-\delta t}^{(i)}\}, \dots, \{s_0^{(i)}\})$$

$$\text{posterior} : p(\psi_t, \phi_t | s_t^{(\phi)}, s_{t-\delta t}^{(\phi)}, \dots, s_0^{(\phi)}; s_t^{(\psi)}, s_{t-\delta t}^{(\psi)}, \dots, s_0^{(\psi)}; \{s_t^{(i)}\}, \{s_{t-\delta t}^{(i)}\}, \dots, \{s_0^{(i)}\}),$$

respectively. But here, for simplicity, we get rid of the data before the current step, and only write the prior and the posterior in the simplified version:

$$\text{prior : } p(\psi_t, \phi_t)$$

$$\text{posterior : } p(\psi_t, \phi_t | s_t^{(\phi)}, s_t^{(\psi)}, \{s_t^{(i)}\}).$$

3.4.4 Powerlaw-like distributions

The powerlaw-like distribution used in this paper has the following form:

$$f(y; \alpha, \gamma, \mu) \propto \frac{1}{1 + |\frac{y-\mu}{\gamma}|^\alpha}, \quad (3.7)$$

where α , γ , and μ are parameters. It should be noted that, when the exponent α is smaller than 1, this distribution cannot be normalized. In this case, a truncation point should be introduced. In fact, throughout the paper, we confine all of our distributions to the interval [-10 semitones, 10 semitones]. This also applies to the 2d case as shown below. The 2d kernel that allows the two variables evolve in a correlated way is defined as:

$$\mathbf{X}_t = \begin{bmatrix} \psi_t \\ \phi_t \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \quad (3.8)$$

$$p(\mathbf{X}_{t+\delta t} | \mathbf{X}_t) \propto \frac{1}{1 + [(\mathbf{X}_{t+\delta t} - \mathbf{X}_t)^\top \Sigma^{-1} (\mathbf{X}_{t+\delta t} - \mathbf{X}_t)]^{\alpha/2}}, \quad (3.9)$$

where Σ represents some correlation measure between the two variables, and $\gamma_{12} = \gamma_{21}$.

3.4.5 Model fitting and parameters

For each day t , we have a joint distribution $p(\psi_t, \phi_t)$, and the distribution of the final motor output $O_t = a\psi_t + (1 - a)\phi_t$ is

$$p_{\text{output}}(O_t) = \iint p(\psi_t, \phi_t) \delta(O_t - a\psi_t - (1 - a)\phi_t) d\psi_t d\phi_t. \quad (3.10)$$

In fact, this distribution of motor output is characterized by the model parameters Θ , which means we can rewrite it as $p_{\text{output}}(O_t; \Theta)$. Here, the parameter Θ contains

all the 14 parameters in our model. We calculate the negative log likelihood of the parameters Θ , and summing over all experimental days to generate the objective function H that is used for the fitting:

$$H = - \sum_t \log p_{\text{output}}(\mathbf{D}_t; \Theta), \quad (3.11)$$

where \mathbf{D}_t represents all the song pitches recorded on day t . The optimization problem is solved in MatLab using the function `fmincon`. All the theoretical results shown in main text are generated by a single group of parameters: 12 of them are indicated by the 6 plus signs '+' in Fig. 3.12, and the weight a is 0.754. It can be seen that besides the sets of parameters we choose, there are a large number of other combinations can give reasonably good fits. In Fig. 3.25, we show the fitting results when a is 0. This degeneracy shows that our model does not rely on fine tuning.

3.5 Additional figures

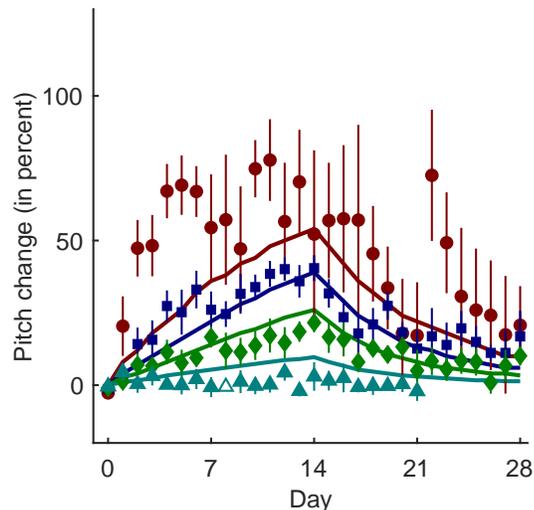


Figure 3.13: This shows the dynamics of medians if we fit the one-timescale model to all the distributions. It can be seen that the model shows very different average behaviors compared with the ones in Fig. 2.2A, and it cannot capture the structures of the decaying tails in washout period. All plot conventions are the same as in Fig. 3.5A

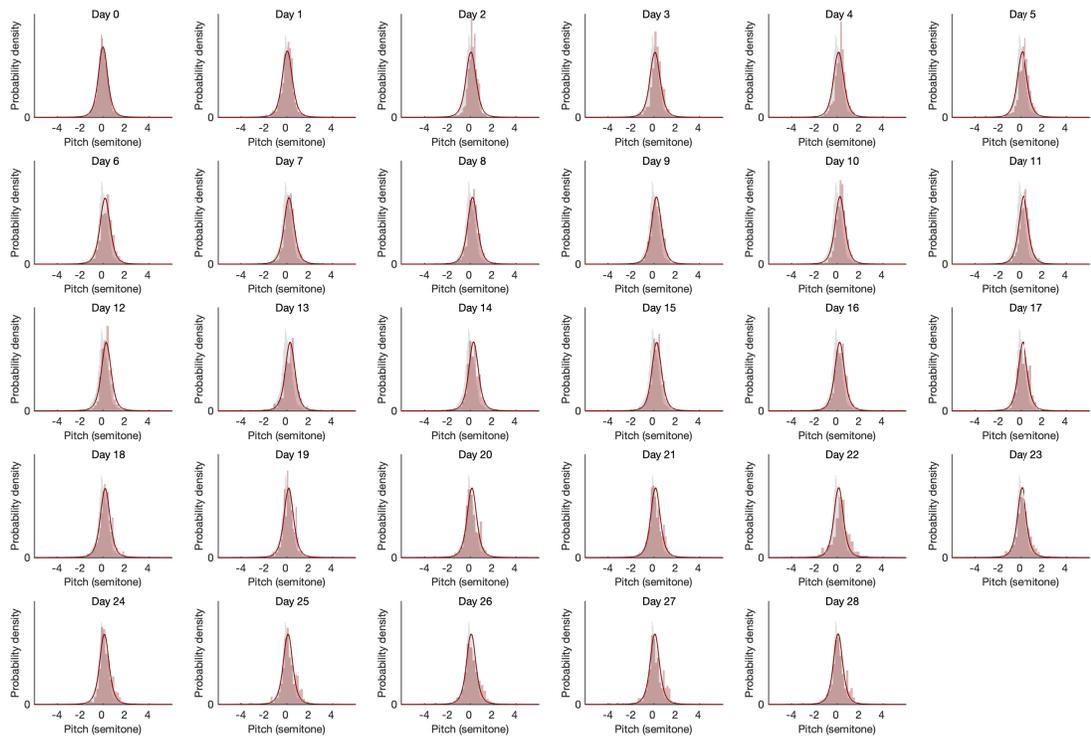


Figure 3.14: This figure shows all distributions for the 0.5-semitone perturbation experiments, which corresponds to the red curve in Fig. 3.5. The grey histograms represent the baseline distribution of the song pitches; the colored histograms represents the distributions of the song pitches on the specific day; the colored curves represent the results of the model.

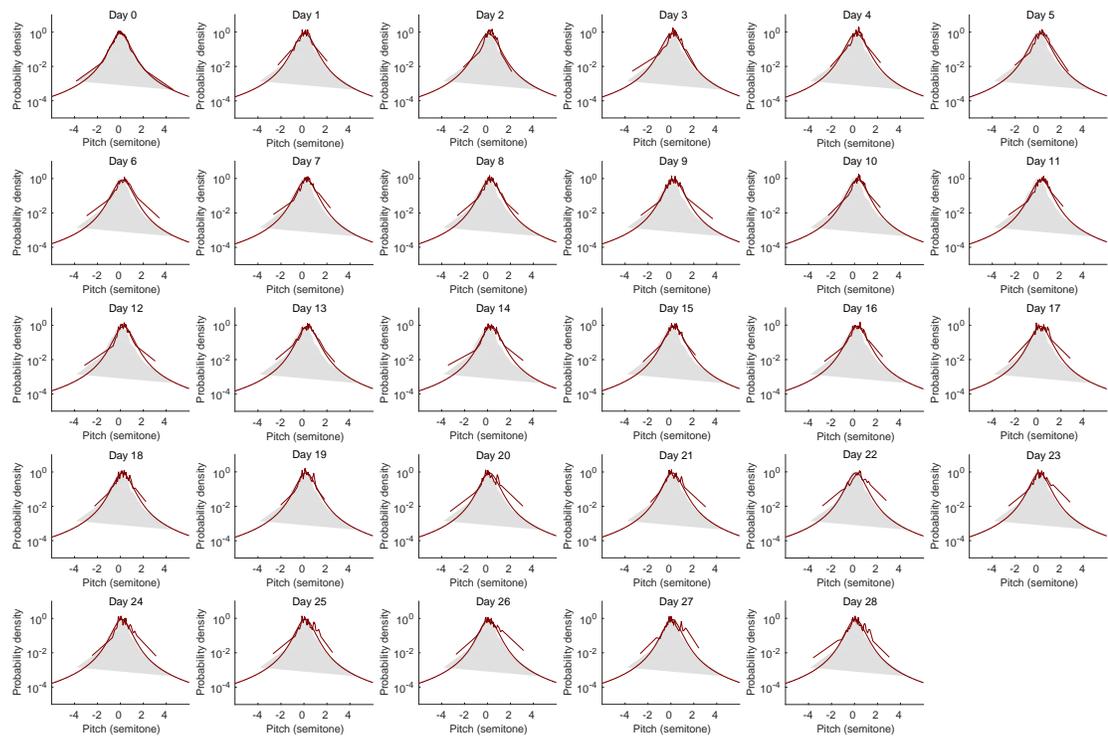


Figure 3.15: This figure shows all distributions in log scale for the 0.5-semitone perturbation experiments, which corresponds to the red curve in Fig. 3.5. The grey shades represent the baseline distribution of the song pitches; the thin colored lines represents the distributions of the song pitches on the specific day; the thick colored curves represent the results of the model.

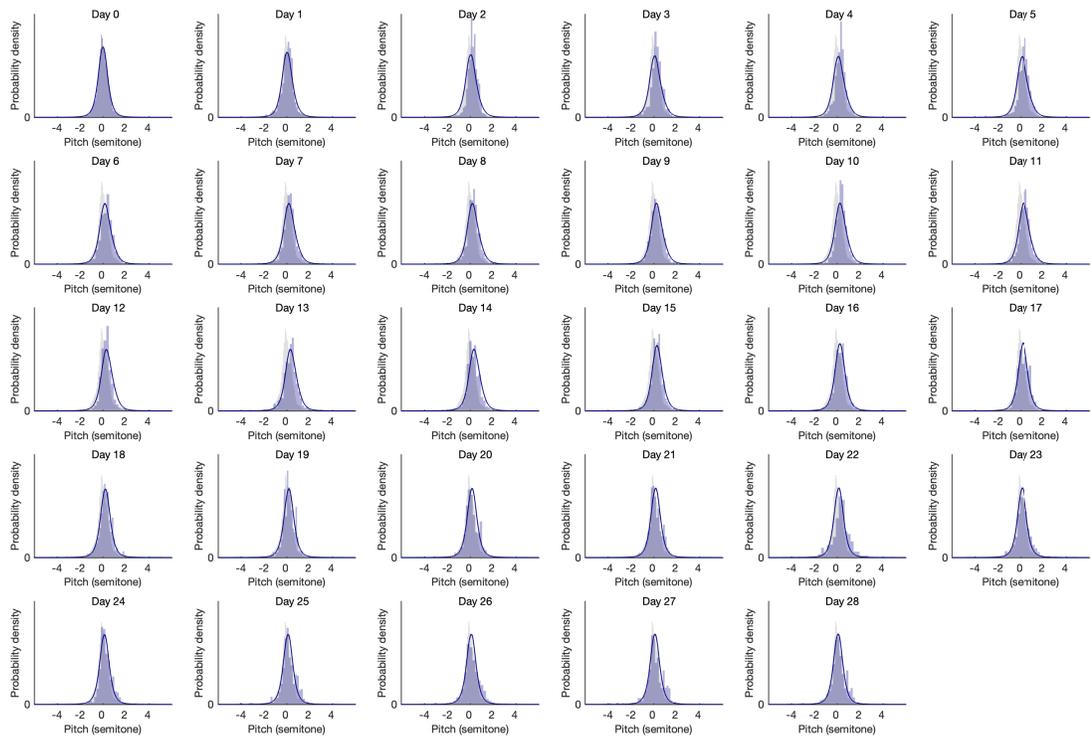


Figure 3.16: This figure shows all distributions for the 1.0-semitone perturbation experiments, which corresponds to the blue curve in Fig. 3.5. The grey histograms represent the baseline distribution of the song pitches; the colored histograms represents the distributions of the song pitches on the specific day; the colored curves represent the results of the model.

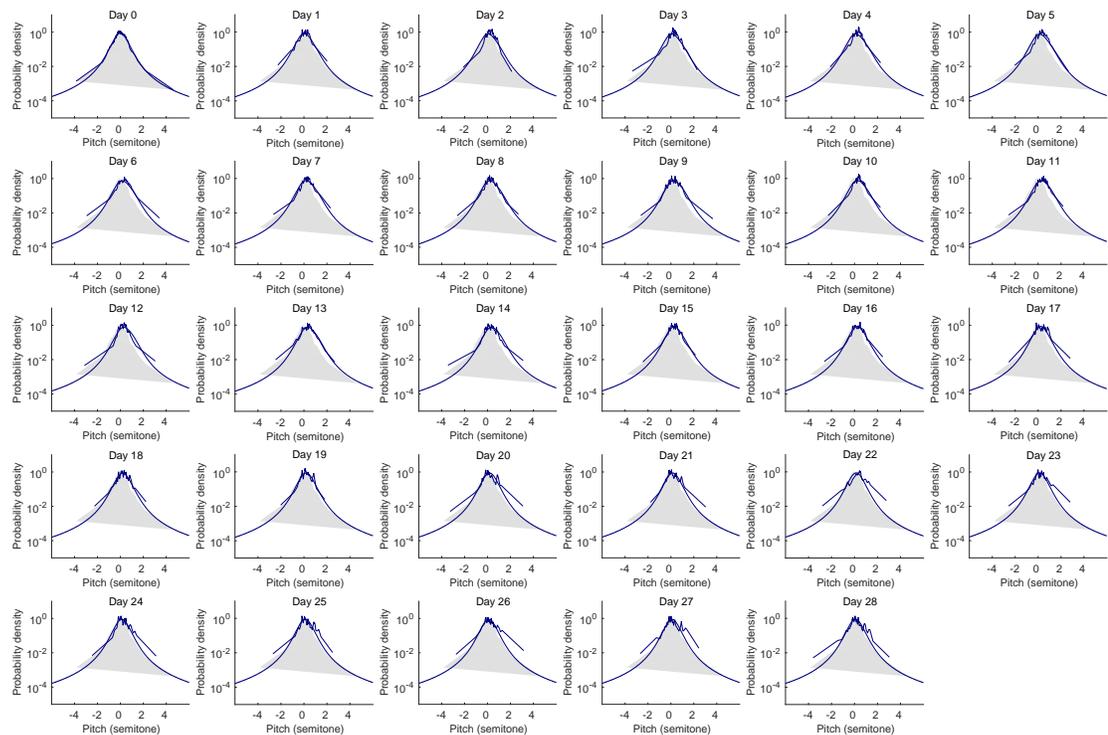


Figure 3.17: This figure shows all distributions in log scale for the 1.0-semitone perturbation experiments, which corresponds to the blue curve in Fig. 3.5. The grey shades represent the baseline distribution of the song pitches; the thin colored lines represents the distributions of the song pitches on the specific day; the thick colored curves represent the results of the model.

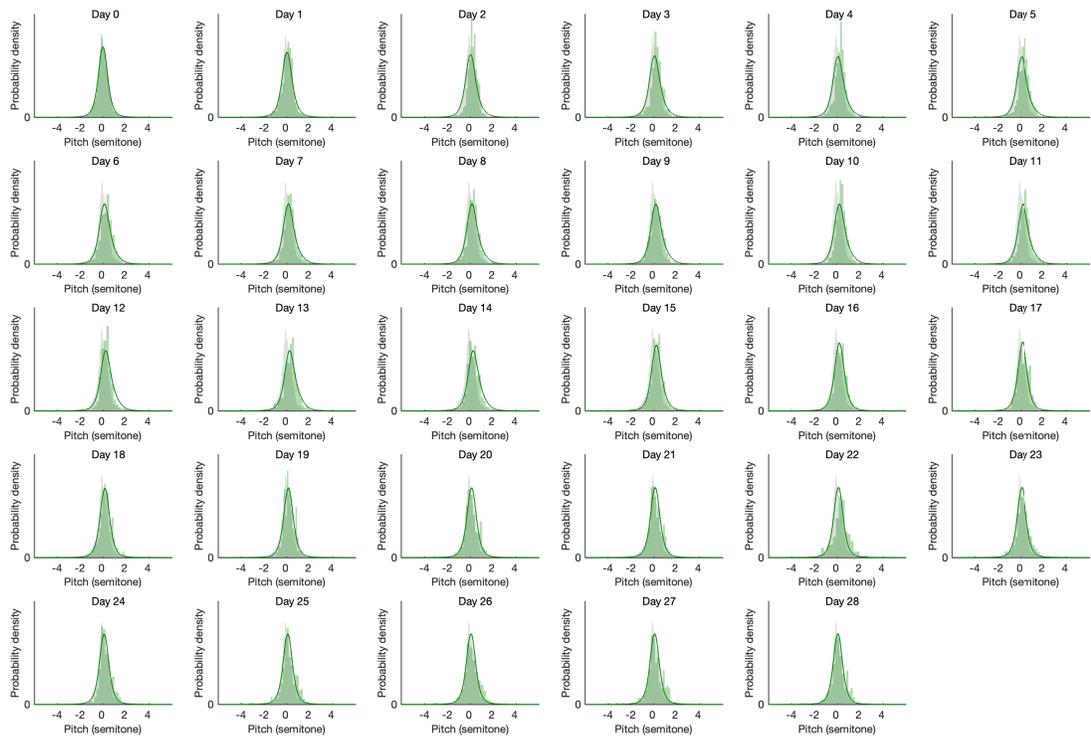


Figure 3.18: This figure shows all distributions for the 1.5-semitone perturbation experiments, which corresponds to the green curve in Fig. 3.5. The grey histograms represent the baseline distribution of the song pitches; the colored histograms represents the distributions of the song pitches on the specific day; the colored curves represent the results of the model.

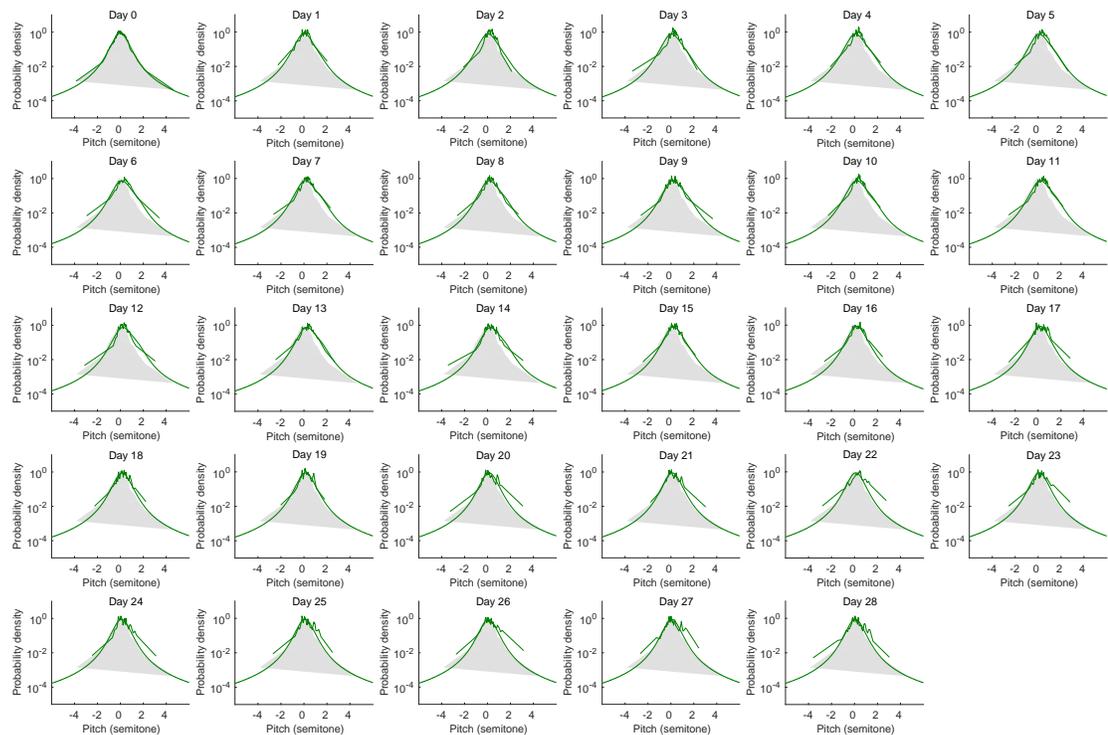


Figure 3.19: This figure shows all distributions in log scale for the 1.5-semitone perturbation experiments, which corresponds to the green curve in Fig. 3.5. The grey shades represent the baseline distribution of the song pitches; the thin colored lines represents the distributions of the song pitches on the specific day; the thick colored curves represent the results of the model.

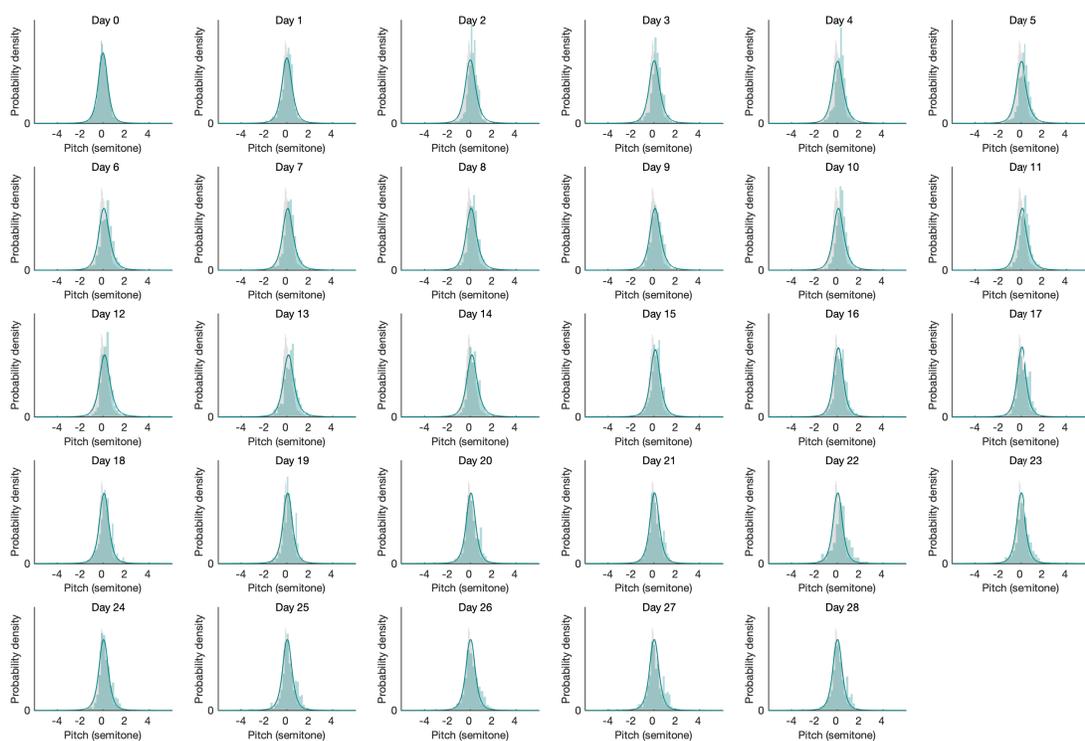


Figure 3.20: This figure shows all distributions for the 3.0-semitone perturbation experiments, which corresponds to the cyan curve in Fig. 3.5. The grey histograms represent the baseline distribution of the song pitches; the colored histograms represents the distributions of the song pitches on the specific day; the colored curves represent the results of the model.

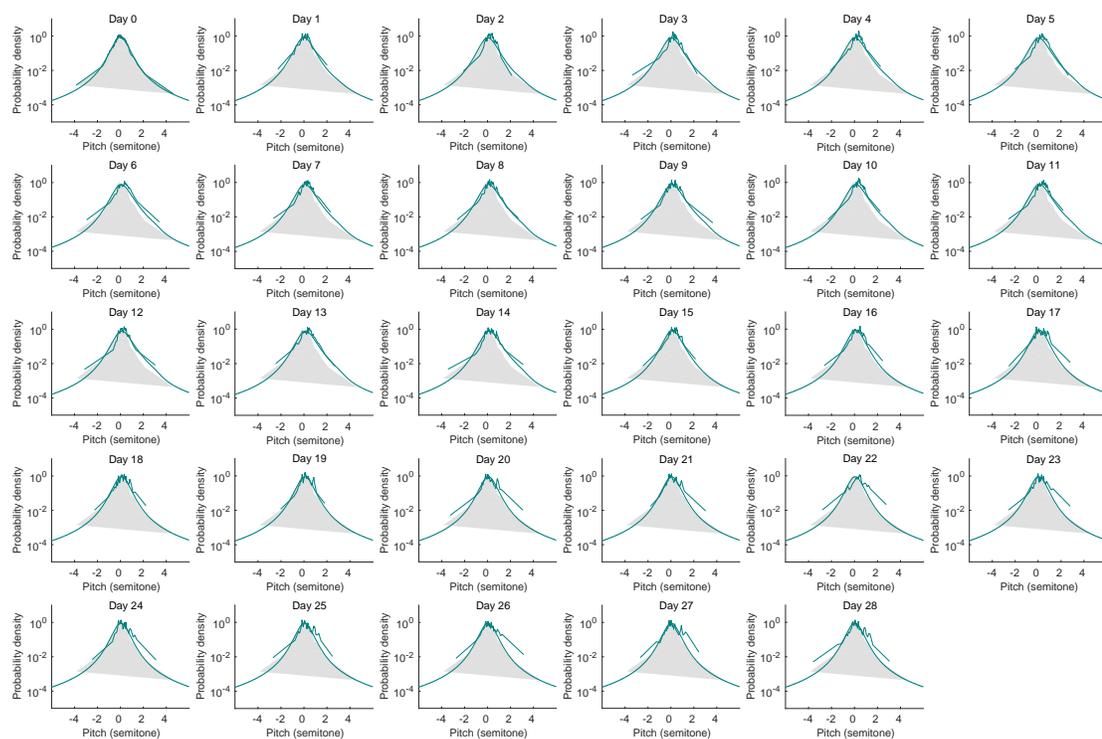


Figure 3.21: This figure shows all distributions in log scale for the 3.0-semitone perturbation experiments, which corresponds to the cyan curve in Fig. 3.5. The grey shades represent the baseline distribution of the song pitches; the thin colored lines represents the distributions of the song pitches on the specific day; the thick colored curves represent the results of the model.

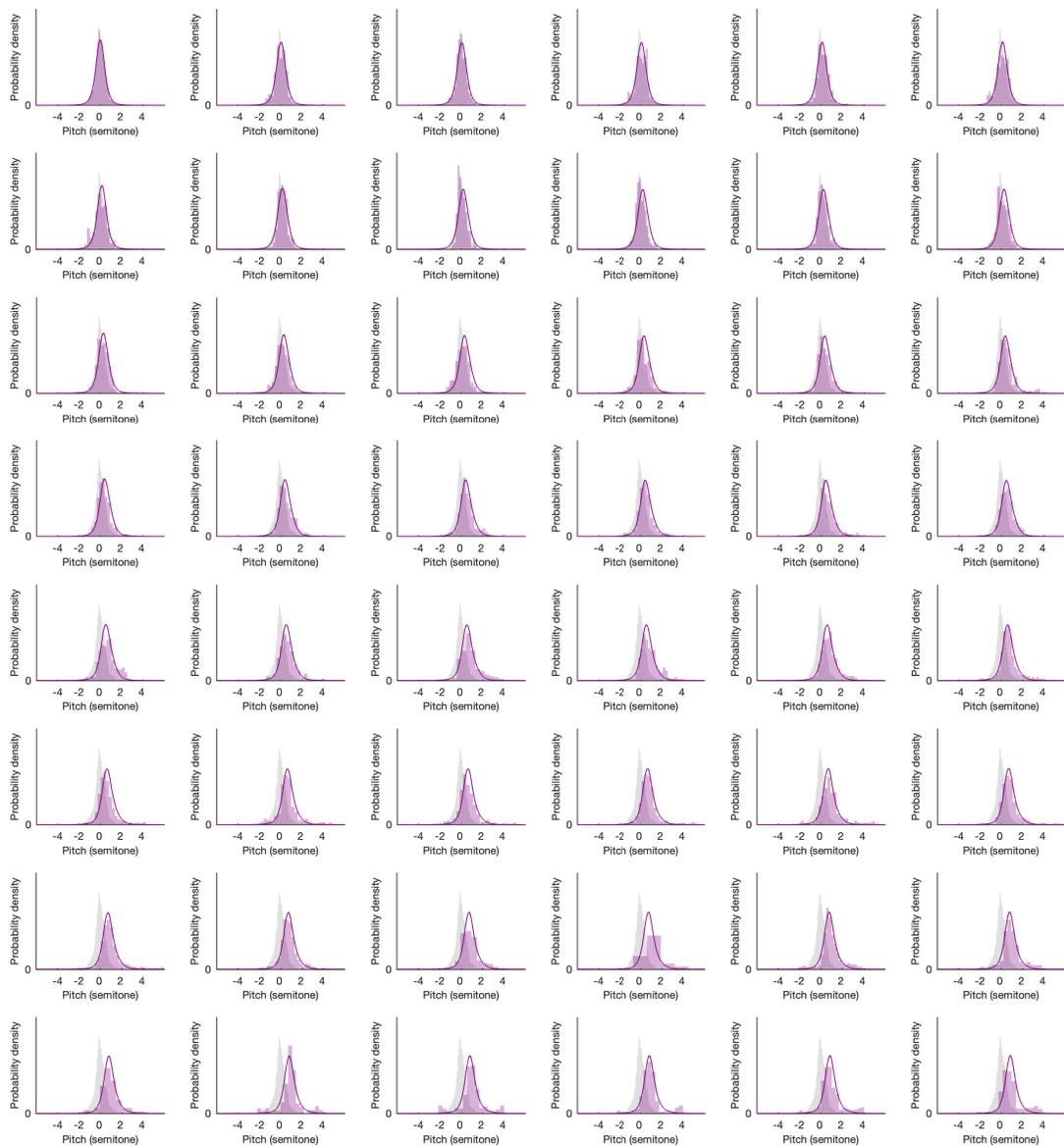


Figure 3.22: This figure shows all distributions for the staircase perturbation experiments, which corresponds to the magenta curve in Fig. 3.5. The grey histograms represent the baseline distribution of the song pitches; the colored histograms represents the distributions of the song pitches on the specific day; the colored curves represent the results of the model.

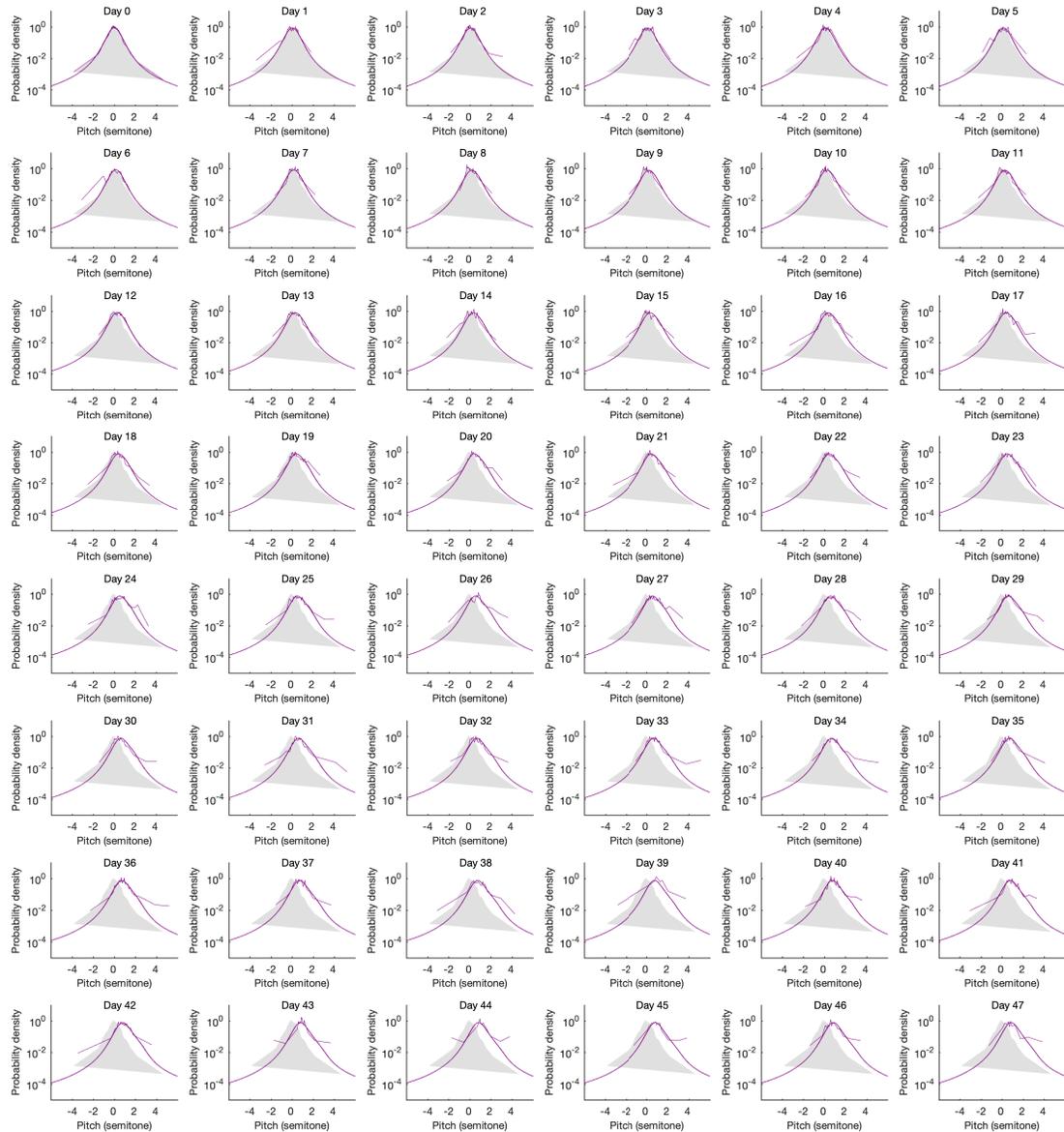


Figure 3.23: This figure shows all distributions in log scale for the staircase perturbation experiments, which corresponds to the magenta curve in Fig. 3.5. The grey shades represent the baseline distribution of the song pitches; the thin colored lines represents the distributions of the song pitches on the specific day; the thick colored curves represent the results of the model.

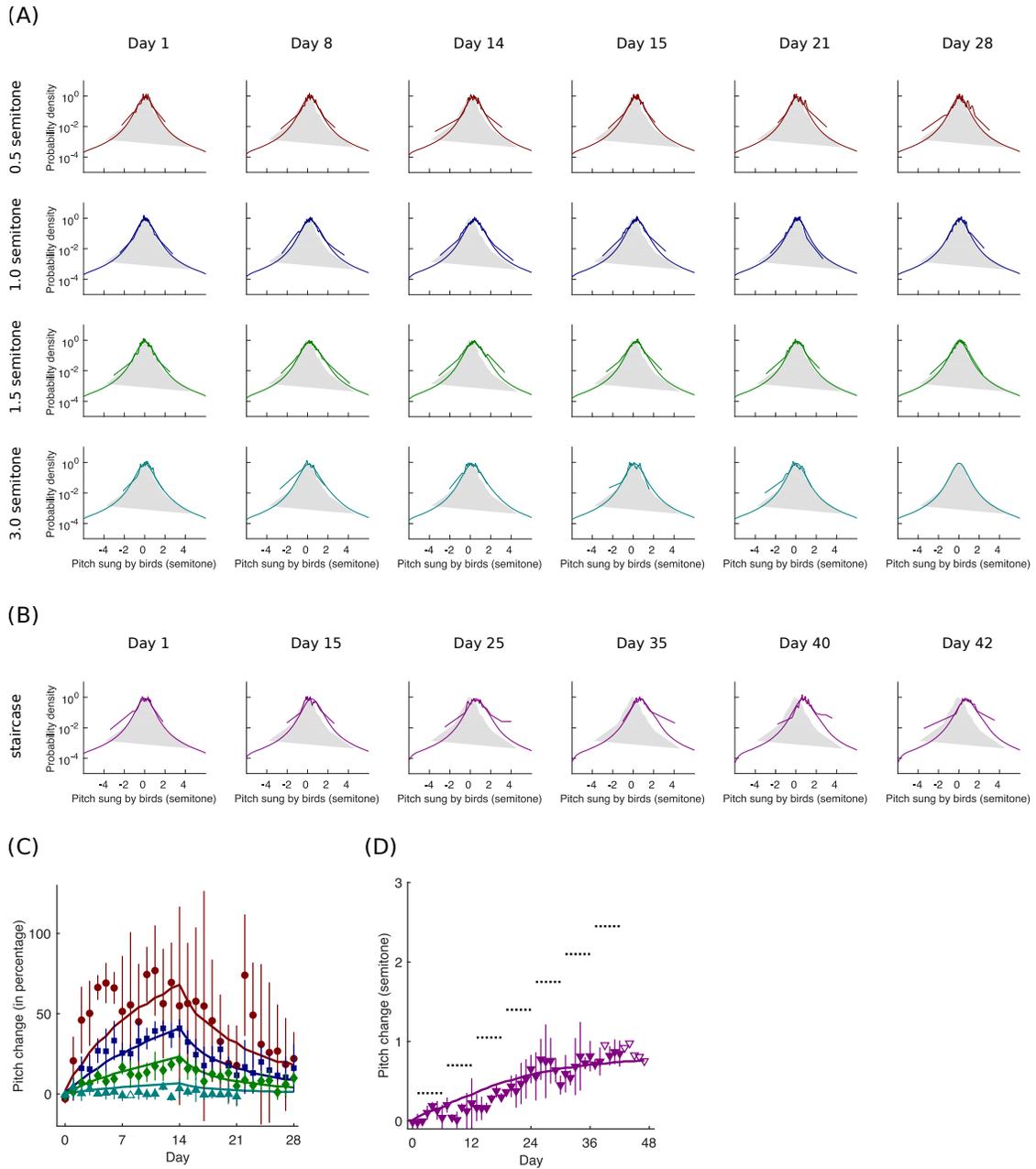


Figure 3.24: Fits of the model with a kernel as in Eq. 3.9. All symbols and colors are the same as in Fig. 3.3 and Fig. 3.5.

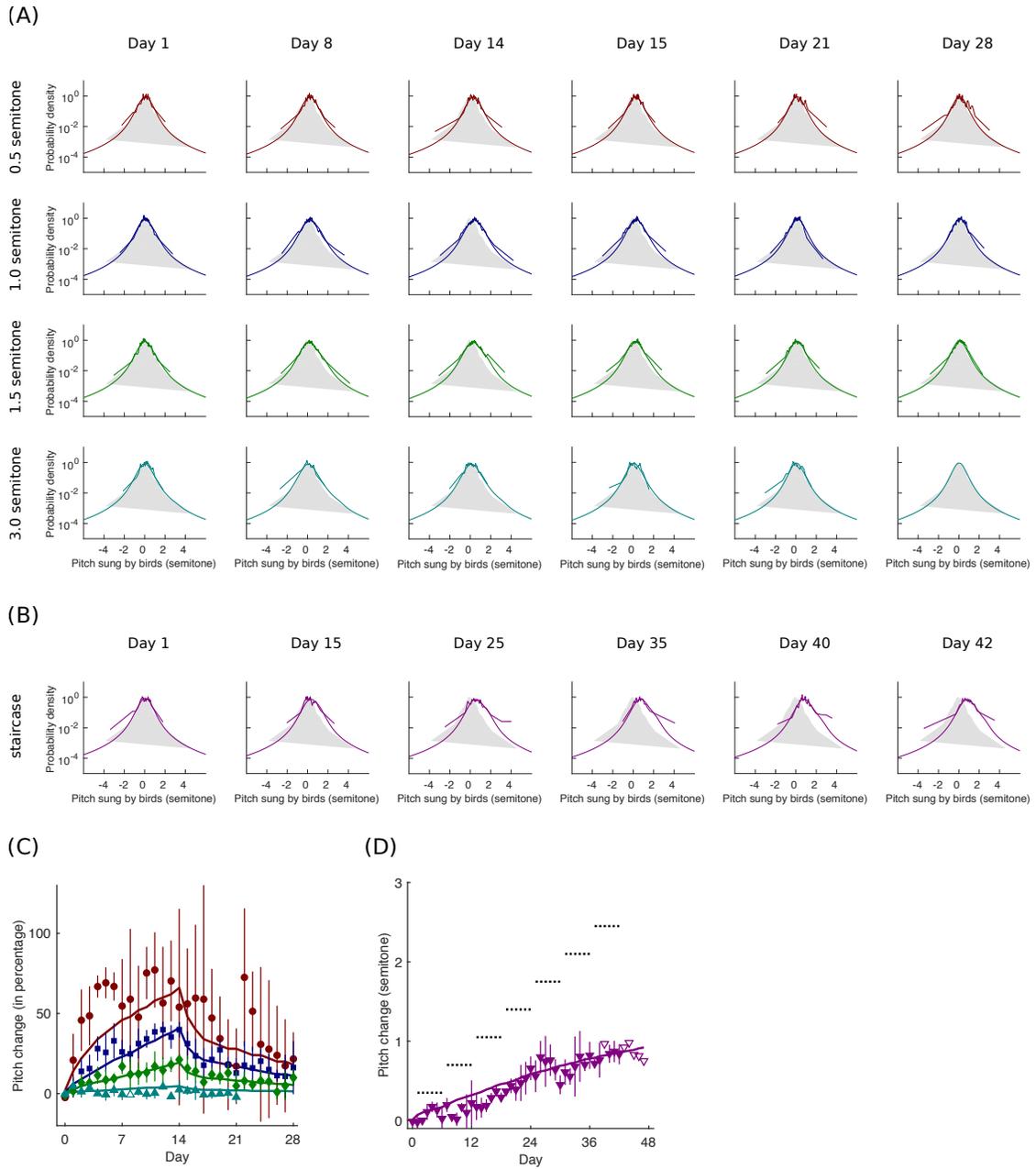


Figure 3.25: Fits of the model as in Eq. 3.6. All symbols and colors are the same as in Fig. 3.3 and Fig. 3.5.

Chapter 4 Behavioral map of a fly predicts species identity

4.1 Introduction

¹ Through millions of years of evolution, fruit flies have differentiated into more than 100 species, and their evolutionary relationships have been genetically identified and mapped to a complicated phylogenetic tree [17, 88]. For simplicity, part of the tree has been shown in Fig. 4.1A, where six species are included. From top to bottom, they are *D. melanogaster*, *D. mauritiana*, *D. simulans*, *D. sechelia*, *D. santomea*, and *D. yakuba*. From the root node to the leaves, we can see that these six species have experienced four differentiations at different time points in history, and each of the horizontal lines represent millions of years. Genetically, these differentiations have been studied thoroughly [17, 88], while little work has been done to investigate this evolutionary process in the behavioral space alone. If behaviors contain evolutionary information, then genetically different individuals must show some discrepancy in their behavioral patterns. Here, we will show that the behavioral patterns can be used to distinguish different species of flies with high accuracy.

4.2 Behavior maps of fly

If you observe a newborn baby's activity for a day, you will find that she/he will show a large amount of moving patterns from different body parts. The baby may blink the eyes, wave the arms and legs, or simply sleep without any movements. It seems at first that there are infinite amount of possible behaviors an animal can perform and there is no way to quantitatively study them. However, for simpler

¹This chapter presents part of an ongoing work in collaboration with Catalina Rivera, Damián G. Hernández, Jessica Cande, David L. Stern, and Gordon J. Berman.

organisms like flies, people have established a method to quantify and enumerate their large repertoire of behaviors [6]. In a transparent 2d chamber, a fly can move freely on a surface, and the activities of the fly is recorded as a movie. In total, we have 561 individual flies that belong to the 6 species shown in Fig. 4.1A. The analysis of these 561 movies can give a 2d heat map as shown in Fig. 4.1B [6], which represents the probability distribution of the time that is spent by flies in performing specific moving patterns (or behaviors). Different regions in Fig. 4.1B represent different behaviors, and the natures of the behaviors have been labeled according to either movements in specific body parts (eg., left or right, wings or legs) or timescales of the behaviors (eg., fast or slow). In fact, the 2d map shown in Fig. 4.1A is the average of 561 individual maps. That is, for each fly, there is a heat map like Fig. 4.1A. We expect that this heat map shows the individuality of each fly and has enough information to distinguish different fly species.

The heat map shown in Fig. 4.1B is a 201 by 201 matrix, which is difficult to deal with because of the large dimensionality. Thus, we grouped adjacent pixels into 134 regions representing more stereotyped behaviors (Fig. 4.1D). For clarity, we define a variable $p_{\mu i}$, where $\mu = 1, 2, \dots, 561$ representing fly identities and $i = 1, 2, \dots, 134$ representing behavior types. Thus, the value of the variable $p_{\mu i}$ indicates the frequency that the fly μ performs the behavior i . Besides $p_{\mu i}$, we define another variable $x_{\mu i}$, which is a transformed version of the data matrix $p_{\mu i}$. The transformation is done by first taking a log for each element in the data matrix, and then center each column, which means for each element:

$$x_{\mu i} = \log p_{\mu i} - \frac{1}{561} \sum_{\mu} \log p_{\mu i}. \quad (4.1)$$

Each heat map in Fig. 4.1D is obtained by averaging $x_{\mu i}$ within each species. It can be seen in Fig. 4.1D that each species has its own unique behavioral repertoires. For example, the *D. mauritania* performs more fast locomotions than *D. yakuba*, while *D. santomea* tends to be more idle than *D. melanogaster*.

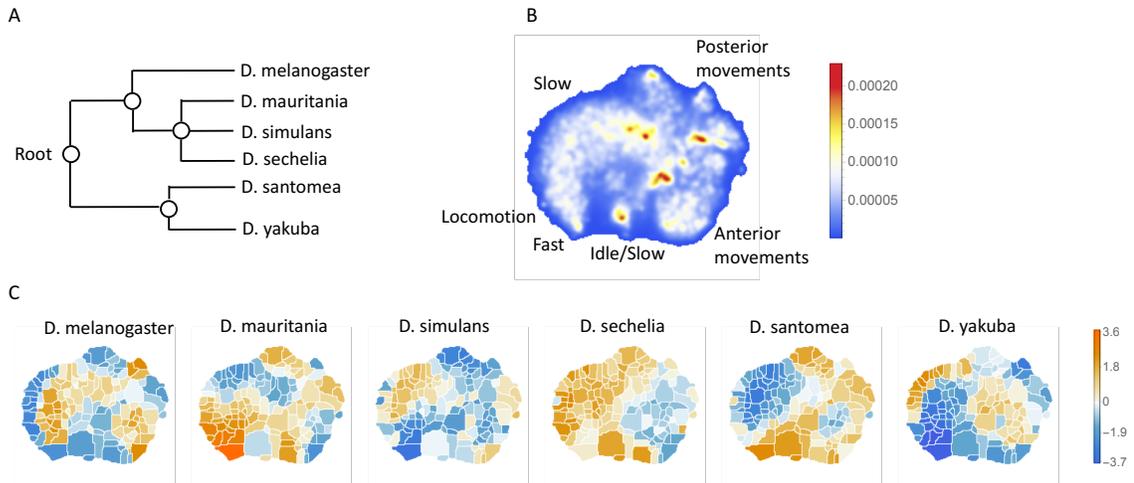


Figure 4.1: Phylogenetic tree and behavioral repertoire. A: Six species of flies are used, and their evolutionary relations are shown in this phylogenetic tree. B: This average behavioral map was obtained by taking the mean of the 561 individual maps, and each individual map was generated from a video recording of a single fly’s behavior. In principle, each pixel on the map represents a specific behavior, and its value represents how frequent a fly performs that behavior. C: The average behavioral maps for each species are shown. Here, the maps look different from the one shown in B where adjacent pixels were clustered together to show a more coarse-grained map. Each small region of those maps represents a stereotyped behavior, and the value of that region is the centered log frequency.

4.3 Embedding by t-SNE

From the matrix $p_{\mu i}$, we can define the distance matrix $d_{\mu\nu}$ of the 561 individuals by calculating the Jensen–Shannon divergence (JSD) between their behavioral frequency vectors:

$$d_{\mu\nu} = \frac{1}{2} \sum_i p_{\mu i} \log \frac{p_{\mu i}}{m_{\mu\nu i}} + \frac{1}{2} \sum_i p_{\nu i} \log \frac{p_{\nu i}}{m_{\mu\nu i}}, \quad (4.2)$$

where $m_{\mu\nu i} = (p_{\mu i} + p_{\nu i})/2$. Each element of this distance matrix is from 0 to 1, and thus, it can be used to construct the similarity matrix used in t-SNE as $1 - d_{\mu\nu}$ [63]. With this similarity matrix, we can use t-SNE to project all the 561 individuals onto a 2d plane (Fig. 4.2A). In Fig. 4.2A, B, different colors represent different species, and different symbols with the same color represent different strains within the same species. Although not perfectly, it is clear that dots with the same color and the

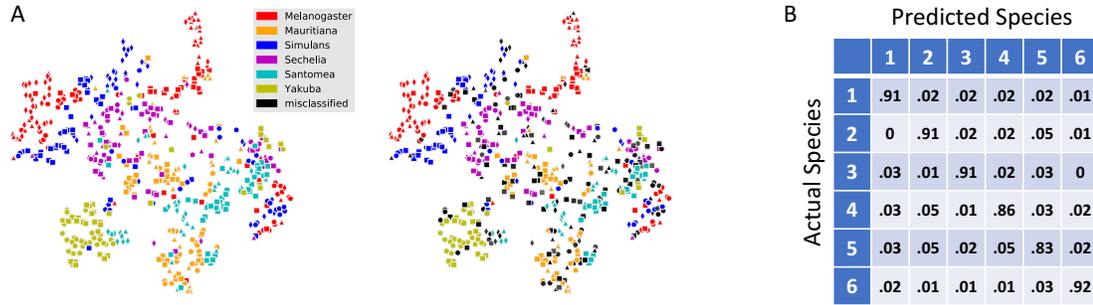


Figure 4.2: Classification of fly species solely based on the behavioral repertoire. A: In the left panel, the t-SNE embedding shows that different species roughly can be separated from each other. Each dot represents one individual fly, with different colors representing different species and different symbols with the same color representing different strains within the same species. Each element in the distance matrix (561 by 561) is the Jensen-Shannon divergence between the behavioral density maps of two individual flies. (red: *D. melanogaster*, orange: *D. mauritiana*, blue: *D. simulans*, magenta: *D. sechelia*, cyan: *D. santomea*, yellow: *D. yakuba*, black/gray: misclassified flies). In the right panel, the black/gray dots represent mis-classified individuals by logistic regression, with their transparency levels scaled linearly with the probability that they would be misclassified: the more transparent the dot is, the more times that individual is mis-classified. C: Confusion matrix for the logistic regression with each row normalized. All the values are averaged from 100 different trials. The standard error is smaller than 0.01 for the diagonal elements, and smaller than 0.005 for the off-diagonal elements. (1: *D. melanogaster*, 2: *D. mauritiana*, 3: *D. simulans*, 4: *D. sechelia*, 5: *D. santomea*, 6: *D. yakuba*)

same symbol tend to form clusters in this 2d embedding space. This suggests that behaviors do contain information that can be used to distinguish fly species and even individual flies.

4.4 Classification by logistic regression

To have a more quantitative understanding of how well behaviors can be used to characterize identities, we used logistic regression to perform a multi-class classification using the transformed feature matrix $x_{\mu i}$. The result of the classification is summarized in the confusion matrix shown in Fig. 4.2B, which illustrates the performance of the multi-class logistic regression model on the test set. We obtained an averaged test score of 0.89 on the test set (30% of the whole data set). The high

test scores shown in the confusion matrix suggest that we can characterize the fly identities reasonably well just using the behavior data.

4.5 Discussion

Biological information is mainly stored in the genome, and the properties of a new born organism is largely determined by its genome. Thus, it should not be a surprise that young flies have their own unique behavior patterns, since behaviors are a reflection of the nervous system and the nervous system is developed according to certain group of genes. Here, we show that behavior information indeed can be used to identify species. This suggests that behavior maps developed in the previous work [6] can be potentially used to study evolution as long as the associated evolutionary dynamics is carefully defined. It would be interesting to see whether there are some proper mathematical frameworks that can be constructed to model the large scale dynamics of this behavioral repertoires through millions of years.

Chapter 5 Conclusion

Carefully designed behavioral experiments can show insightful information about the underlying nervous system that produces them. This approach requires making thoughtful hypotheses about how the system might function at the macroscopic scale, and it is assumed that these hypotheses should put meaningful constraints on the nervous systems [55]. Here, we use male Bengalese finch, a type of songbird, and its singing behavior as a model system to study the general computational rules that govern the sensorimotor learning. The pitch perturbation on the bird's auditory feedback reveals rich dynamics that can be used to construct hypotheses about the sensorimotor systems and to build physical models to test these ideas. For each day, we have a whole heavy-tailed distribution of the song pitches with certain variability, which reflects the motor explorations of the animal. This distribution and its mean show interesting dynamics in various experimental settings that cannot be explained by previous sensorimotor learning models. Our Bayesian filter model with non-Gaussian statistics and multiple timescales captures all the important features of the song pitch data. Specifically, the non-Gaussian assumption on the likelihood terms explains the fact that an animal adapts better if the perturbation is small, and if the perturbation is too large, the animal cannot adapt at all, at least within the time frame of the experiments. However, if the large perturbation is imposed gradually, the animal can learn how to adapt to it very well [94]. As the two-rate model discussed in the introduction, our model assumes that there are more than one neural components in the motor system, and multiple motor memories are integrating information from past experience with different timescales. The model can explain the prolonged tails in the washout period, which is hard to describe with one-timescale models. More importantly, our model can describe the dynamics of the

whole distributions of the song pitches, which means that the model can explain how the exploration strategy of the motor system evolves with time.

Our model can make various predictions about the songbird’s singing behavior that can be tested in experiments. First, by carefully designing certain three-phase pitch-perturbation paradigms as shown in Fig. 3.6, our model predicts that the relearning of a perturbation can be faster than the learning when the animal encounter the perturbation for the first time. Moreover, if the perturbation in the third phase is set to 0, the model predicts a spontaneous rebound of the behavior toward the adaptation direction. This rebound has been observed in arm-reaching experiments, and it would be interesting if we can see the same effect in songbird’s singing behavior. Second, our model shows that smaller variance in the motor system will lead to less learning. This can be tested without any neural interventions, but it does need a relatively large sample size. For the current study, there are only 9 birds, which may not yield a clear positive correlation between motor variability and adaptation as shown in other motor systems. Third, our model can also predict the retreat of the learning when the LMAN-RA projections are blocked. This result is qualitatively comparable with the observations made in experiments where the adaptation is white-noise driven. Direct evidence in pitch-perturbation is needed. However, since the effects of LMAN-blocking in experiments might appear immediately or in only several hours, we may need to simulate the model in a hour-to-hour base, and night times should be included where no singing event happens.

Motor system adaptation happens when the animal encounters either errors in their behavior (usually continuous signals) or rewards/punishments (usually discrete signals). We have only tested our model on the error-based learning, and it would be interesting to see whether our model can be generalized to describe reward/punishment-based learning. A typical punishment-based learning experiment in a songbird is done by exposing the animal to aversive, loud, white noise signal when the pitch of its song

falls outside of a prescribed window. The motivations behind the adaptations in the perturbation-driven and punishment-driven cases are different. When facing a perturbation in its song pitch, the bird wants to correct the perturbation in order to perceive the pitch it prefers; when facing a punishment like white noise, the bird wants to avoid this punishment by changing its song pitch, but at the same time it may also want to push back this change since the song deviates the template. One potential way to model this is to add a new likelihood term with the mean as a free parameter that can be fit to the data. This mean can be either static or adaptive, and represents how far the bird is willing to be away from the pitch that triggers the white-noise punishment. This line of modeling is not very attractive, since we would just keep adding complexity to our model. However, this might be less concerning if we consider that the underlying processes that the white-noise experiments activate are very complicated, since both punishment avoidance and template matching are in play.

We have shown that our model can predict fast relearning and spontaneous rebound phenomena similar to those observed in arm-reaching experiments. Thus, it is plausible that our model can be applied to the motor system that governs the arm movement. For example, it would be interesting to check whether the distribution of arm-reaching data shows similar heavy tails as in songbird vocal behaviors. If the answer is yes, then the same non-Gaussian assumption can be made for the arm-reaching system as well, and we would like to see the same negative correlation between the perturbation size and adaptation. More importantly, we expect that the Bayesian filter model that we have established in songbirds is general and can be potentially used in many other sensorimotor systems.

The central question we are asking is: how does the brain control behaviors? Thus, the behavior studies are not just about the behavior itself; rather, the computational models we built on the behavioral level should also serve as a guide for people who

are interested in cellular or circuit level descriptions of the problem. Our results show that in order to understand the behavior data, we have to accept certain hypotheses, such as non-Gaussian statistics and multiple interacting neural components. We argue that these hypotheses should be seen as constraints on the relevant nervous systems, which means whatever the neuron activities are and however the networks are connected, similar large scale structures as proposed here should emerge. This is also why behavioral study should be the first step when we try to investigate the relations between the brain and the behaviors. Building the theory bottom-up from neurons and networks is usually hard or even impossible, and the large scale concepts and mathematical models formulated by carefully studying the behavior itself usually provide more insightful information about what exact the nervous system as a whole is doing. We hope our results are helpful for neuroscientist and even can be tested on neuron/circuit recordings in the near future.

One of the hot topics in artificial intelligence (AI) is how to build robots that behave like humans, which means that the behaviors of robots should roughly follow the same set of rules as humans'. Thus, what we achieved here might also be used as some guidance to build new AI systems. One possibility lies at the cascade model where only one variable ϕ can learn from the data, and the other variable ψ can only learn from the first variable ϕ . This structure might be used to bring stability and long-term memory to AI systems, since ψ can operate on a much longer timescale. In addition, we can assume that a specific motor system consists of a group of modules and different combinations of these modules would allow the motor system to perform different types of tasks. In this way, for a robotic system, we can have one final output module ϕ and a list of hidden modules or units ψ_i , $i = 1, 2, 3, \dots$. All the hidden units can only learn from ϕ , and their different combinations should correspond to different tasks. Thus, we might be able to build some AI systems that can be multi-tasking.

Beside building sensorimotor learning models from a single behavior, we have also

investigated how a group of behaviors contains individuality information of flies. We built a simple logistic regression model to show that fly species can be distinguished in high accuracy by using the behavioral maps as features. This suggests a new way to look at the evolution purely in the behavioral space without the need of any genetic information.

Bibliography

1. Andalman, A. S. & Fee, M. S. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences* **106**, 12518–12523 (2009).
2. Beck, J., Ma, W., Pitkow, X., Latham, P. & Pouget, A. Not Noisy, Just Wrong: The Role of Suboptimal Inference in Behavioral Variability. *Neuron* **74**, 30–39. ISSN: 0896-6273 (2012).
3. Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. *Nature neuroscience* **10**, 1214 (2007).
4. Belov, I. A. On the Computation of the Probability Density Function of α -Stable Distributions. *Mathem Model Anal* **2**, 333–341 (2005).
5. Bergström, H. On some expansions of stable distribution functions. *Ark Mat* **2**, 375–378 (1952).
6. Berman, G. J., Choi, D. M., Bialek, W. & Shaevitz, J. W. Mapping the stereotyped behaviour of freely moving fruit flies. *Journal of The Royal Society Interface* **11**, 20140672 (2014).
7. Bernacchia, A., Seo, H., Lee, D. & Wang, X.-J. A reservoir of time constants for memory traces in cortical neurons. *Nature neuroscience* **14**, 366 (2011).
8. Bolhuis, J. J., Okanoya, K. & Scharff, C. Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience* **11**, 747 (2010).
9. Bottjer, S. W., Miesner, E. A. & Arnold, A. P. Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* **224**, 901–903 (1984).
10. Brainard, M. S. & Doupe, A. What songbirds teach us about learning. *Nature* **417**, 351–358 (2002).
11. Brainard, M. S. & Doupe, A. J. Auditory feedback in learning and maintenance of vocal behaviour. *Nature Reviews Neuroscience* **1**, 31 (2000).
12. Brainard, M. S. & Doupe, A. J. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404**, 762 (2000).
13. Bromberg-Martin, E. S., Matsumoto, M., Nakahara, H. & Hikosaka, O. Multiple timescales of memory in lateral habenula and dopamine neurons. *Neuron* **67**, 499–510 (2010).
14. Buesing, L., Bill, J., Nessler, B. & Maass, W. Neural Dynamics as Sampling: A Model for Stochastic Computation in Recurrent Networks of Spiking Neurons. *PLOS Comput Biol* **7**, e1002211. ISSN: 1553-7358 (2011).

15. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51 (2012).
16. Colagiorgio, P., Bertolini, G., Bockisch, C. J., Straumann, D. & Ramat, S. Multiple timescales in the adaptation of the rotational VOR. *Journal of neurophysiology* **113**, 3130–3142 (2015).
17. Consortium, D. 1. G. *et al.* Evolution of genes and genomes on the Drosophila phylogeny. *Nature* **450**, 203 (2007).
18. Dayan, P. & Niv, Y. Reinforcement Learning: The Good, The Bad and The Ugly. *Curr Opin Neurobiol. Cognitive neuroscience* **18**, 185–196. ISSN: 0959-4388 (2008).
19. Dhawale, A. K., Smith, M. A. & Ölveczky, B. P. The role of variability in motor learning. *Annual review of neuroscience* **40**, 479–498 (2017).
20. Donchin, O. & Shadmehr, R. *Linking Motor Learning to Function Approximation: Learning in an Unlearnable Force Field* in *Adv Neural Inf Proc Syst 14* **1** (MIT Press, Cambridge, USA, 2001), 7.
21. Doya, K. & Sejnowski, T. in *The New Cognitive Neurosciences* (ed Gazzaniga, M.) 2nd, 469–482 (MIT Press, Cambridge, MA, 2000).
22. Doya, K., Ishii, S., Pouget, A. & Rao, R. P. *Bayesian brain: Probabilistic approaches to neural coding* (MIT press, 2007).
23. Eden, S. *Stroke of madness: How has Tiger Woods managed to overhaul his swing three times* tech. rep. (ESPN, 2013).
24. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429 (2002).
25. Faisal, A. A., Selen, L. P. & Wolpert, D. M. Noise in the nervous system. *Nature reviews neuroscience* **9**, 292 (2008).
26. Farries, M. & Fairhall, A. Reinforcement Learning With Modulated Spike Timing-Dependent Synaptic Plasticity. *J Neurophys* **98**, 3648–3665 (2007).
27. Fiete, I. & Seung, H. in *Encyclopedia of Neuroscience* (ed Squire, L.) 227–239 (Academic Press, Oxford, 2009).
28. Fiete, I. R. & Seung, H. S. Neural network models of birdsong production, learning, and coding. *New Encyclopedia of Neuroscience. Eds. L. Squire, T. Albright, F. Bloom, F. Gage, and N. Spitzer. Elsevier* (2007).
29. Fischer, B. J. & Peña, J. L. Owl’s Behavior and Neural Representation Predicted by Bayesian Inference. *Nature neurosci* **14**, 1061–1066. ISSN: 1546-1726 (2011).
30. Fiser, J., Berkes, P., Orbán, G. & Lengyel, M. Statistically Optimal Perception and Learning: From Behavior to Neural Representations. *Trends Cogn Sci* **14**, 119–130. ISSN: 1364-6613 (2010).

31. Galea, J. M., Vazquez, A., Pasricha, N., Orban de Xivry, J.-J. & Celnik, P. Dissociating the roles of the cerebellum and motor cortex during adaptive learning: the motor cortex retains what the cerebellum learns. *Cerebral cortex* **21**, 1761–1770 (2010).
32. Gallistel, C. R., Mark, T. A., King, A. P. & Latham, P. E. The Rat Approximates an Ideal Detector of Changes in Rates of Reward: Implications for the Law of Effect. eng. *J Exper Psych. Animal Behav Proc* **27**, 354–372. ISSN: 0097-7403 (2001).
33. Genewein, T., Hez, E., Razzaghpahan, Z. & Braun, D. A. Structure Learning in Bayesian Sensorimotor Integration. *PLoS Comput Biol* **11**, 27. ISSN: 1553-7358 (2015).
34. Gershman, S. J. A Unifying Probabilistic View of Associative Learning. *PLoS Comput Biol* **11**, e1004567 (2015).
35. Gutenkunst, R. *et al.* Universally sloppy parameter sensitivities in systems biology models. *PLoS Comp Biol* **3**, 1871–1878 (2007).
36. Hahn, A., Kryslar, A. & Sturdy, C. Female song in black-capped chickadees (*Poecile atricapillus*): acoustic song features that contain individual identity information and sex differences. *Behav Proc* **98**, 98–105 (2013).
37. Hahnloser, R. & Narula, G. A Bayesian Account of Vocal Adaptation to Pitch-Shifted Auditory Feedback. *PLoS ONE* **12**, e0169795 (2017).
38. Herzfeld, D. J. & Shadmehr, R. Motor variability is not noise, but grist for the learning mill. *nature neuroscience* **17**, 149 (2014).
39. Hoffmann, L. A., Kelly, C. W., Nicholson, D. A. & Sober, S. J. A Lightweight, Headphones-Based System for Manipulating Auditory Feedback in Songbirds. *Journal of Visualized Experiments : JoVE*. ISSN: 1940-087X. doi:[10 . 3791 / 50027](https://doi.org/10.3791/50027) (2012).
40. Inoue, M. *et al.* Three timescales in prism adaptation. *Journal of neurophysiology* **113**, 328–338 (2014).
41. Jaderberg, M. *et al.* Human-level performance in first-person multiplayer games with population-based deep reinforcement learning. *arXiv preprint arXiv:1807.01281* (2018).
42. Joel, D., Niv, Y. & Ruppin, E. Actor–critic Models of the Basal Ganglia: New Anatomical and Computational Perspectives. *Neural Networks* **15**, 535–547. ISSN: 0893-6080 (2002).
43. Kaipio, J. & Somersalo, E. *Statistical and Computational Inverse Problems* (eds Antman, S., Marsden, J. & Sirovich, L.) ISBN: 0-387-22073-9 (Springer, New York City, NY, 2004).
44. Kao, M. H. & Brainard, M. S. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *Journal of neurophysiology* **96**, 1441–1455 (2006).

45. Kao, M. H., Doupe, A. J. & Brainard, M. S. Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature* **433**, 638 (2005).
46. Kappel, D., Habenschuss, S., Legenstein, R. & Maass, W. Network Plasticity as Bayesian Inference. *PLOS Comput Biol* **11**, e1004485. ISSN: 1553-7358 (2015).
47. Ke, M. C., Guo, C. C. & Raymond, J. L. Elimination of climbing fiber instructive signals during motor learning. *Nature neuroscience* **12**, 1171 (2009).
48. Kelly, C. W. & Sober, S. J. A Simple Computational Principle Predicts Vocal Adaptation Dynamics across Age and Error Size. *Frontiers Integrative Neurosci* **8**, 9 (2014).
49. Knill, D. C. & Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences* **27**, 712–719 (2004).
50. Knudsen, E. I. Capacity for plasticity in the adult owl auditory system expanded by juvenile experience. *Science* **279**, 1531–1533 (1998).
51. Knudsen, E. I. Instructed learning in the auditory localization pathway of the barn owl. *Nature* **417**, 322 (2002).
52. Kojima, Y., Iwamoto, Y. & Yoshida, K. Memory of learning facilitates saccadic adaptation in the monkey. *Journal of Neuroscience* **24**, 7531–7539 (2004).
53. Kording, K. P., Tenenbaum, J. B. & Shadmehr, R. The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nature neuroscience* **10**, 779 (2007).
54. Körding, K. P. & Wolpert, D. M. Bayesian integration in sensorimotor learning. *Nature* **427**, 244 (2004).
55. Krakauer, J. W., Ghazanfar, A. A., Gomez-Marín, A., MacIver, M. A. & Poeppel, D. Neuroscience needs behavior: correcting a reductionist bias. *Neuron* **93**, 480–490 (2017).
56. Krakauer, J. W. & Mazzoni, P. Human sensorimotor learning: adaptation, skill, and beyond. *Current opinion in neurobiology* **21**, 636–644 (2011).
57. Krakauer, J. W. & Shadmehr, R. Consolidation of motor memory. *Trends in neurosciences* **29**, 58–64 (2006).
58. Kuebrich, B. & Sober, S. Variations on a theme: Songbirds, variability, and sensorimotor error correction. *Neuroscience* **296**, 48–54 (2015).
59. Lak, A., Stauffer, W. & Schultz, W. Dopamine Neurons Learn Relative Chosen Value from Probabilistic Rewards. *eLife* **5**, e18044. ISSN: 2050-084X (2016).
60. Linkenhoker, B. A. & Knudsen, E. I. Incremental Training Increases the Plasticity of the Auditory Space Map in Adult Barn Owls. *Nature* **419**, 293–296 (2002).
61. Liu, Y. *et al.* Selective and Divided Attention Modulates Auditory–vocal Integration in the Processing of Pitch Feedback Errors. en. *European J Neurosci* **42**, 1895–1904. ISSN: 1460-9568 (2015).

62. Ma, W. J., Beck, J. M., Latham, P. E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nature neuroscience* **9**, 1432 (2006).
63. Maaten, L. v. d. & Hinton, G. Visualizing data using t-SNE. *Journal of machine learning research* **9**, 2579–2605 (2008).
64. Marr, D. *Vision: A Computational Approach* (MIT Press, USA).
65. McDonnell, M. D. & Ward, L. M. The Benefits of Noise in Neural Systems: Bridging Theory and Experiment. *Nature Rev Neurosci* **12**, 415–426. ISSN: 1471-003X, 1471-0048 (2011).
66. Mets, D. & Braind, M. Genetic variation interacts with experience to determine interindividual differences in learned song. *Proc Natl Acad Sci (USA)* **115**, 421–426 (2018).
67. Mittnik, S., Paoletta, M. S. & Rachev, S. T. Diagnosing and Treating the Fat Tails in Financial Returns Data. *J Empir Finance. Special issue on Risk Management* **7**, 389–416. ISSN: 0927-5398 (2000).
68. Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nature neuroscience* **17**, 1661 (2014).
69. Nemenman, I. Coincidences and Estimation of Entropies of Random Variables with Large Cardinalities. *Entropy* **13**, 2013–2023 (2011).
70. Nemenman, I., Shafee, F. & Bialek, W. *Entropy and Inference, Revisited in Adv Neural Inf Proc Syst (NIPS)* (eds Dietterich, T., Becker, S. & Ghahramani, Z.) **14** (MIT Press, Cambridge, MA, 2002), 471–478.
71. Neuringer, A. Operant variability: Evidence, functions, and theory. *Psychonomic Bulletin & Review* **9**, 672–705 (2002).
72. Neymotin, S. A., Chadderdon, G. L., Kerr, C. C., Francis, J. T. & Lytton, W. W. Reinforcement Learning of Two-Joint Virtual Arm Reaching in a Computer Model of Sensorimotor Cortex. *Neural Comput* **25**, 3263–3293. ISSN: 0899-7667 (2013).
73. Nikias, C. L. & Shao, M. *Signal Processing with Alpha-Stable Distributions and Applications* ISBN: 978-0-471-10647-0 (Wiley, New York, 1995).
74. Niziolek, C., Nagarajan, S. & Houde, J. What does motor efference copy represent? Evidence from speech production. *J Neurosci* **33**, 16110–16116 (2013).
75. Nolan, J. P. *Stable Distributions - Models for Heavy Tailed Data* (Birkhauser, Boston, 2015).
76. Nordeen, K. & Nordeen, E. Deafening-induced vocal deterioration in adult songbirds is reversed by disrupting a basal ganglia-forebrain circuit. *Journal of Neuroscience* **30**, 7392–7400 (2010).
77. Ölveczky, B. P., Andalman, A. S. & Fee, M. S. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS biology* **3**, e153 (2005).

78. Petrovici, M. A., Bill, J., Bytschok, I., Schemmel, J. & Meier, K. Stochastic Inference with Spiking Neurons in the High-Conductance State. en. *Phys Rev E* **94**. ISSN: 2470-0045, 2470-0053. doi:[10.1103/PhysRevE.94.042312](https://doi.org/10.1103/PhysRevE.94.042312) (2016).
79. Pouget, A., Beck, J. M., Ma, W. J. & Latham, P. E. Probabilistic brains: knowns and unknowns. *Nature neuroscience* **16**, 1170 (2013).
80. Rescorla, R. & Wagner, A. in *Classical Conditioning II* (eds Black, A. & Prokasy, W.) 64–99 (Appleton-Century-Crofts, New York City, NY, 1972).
81. Robinson, F. R., Noto, C. T. & Bevans, S. E. Effect of visual error size on saccade adaptation in monkey. *Journal of neurophysiology* **90**, 1235–1244 (2003).
82. Roth, E., Hall, R. W., Daniel, T. L. & Sponberg, S. Integration of parallel mechanosensory and visual pathways resolved through sensory conflict. *Proceedings of the National Academy of Sciences* **113**, 12832–12837 (2016).
83. Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Distinct timescales of population coding across cortex. *Nature* **548**, 92 (2017).
84. Salas-Gonzalez, D., Górriz, J. M., Ramirez, J., Illán, I. A. & Lang, E. W. Linear Intensity Normalization of FP-CIT SPECT Brain Images Using the α -Stable Distribution. *NeuroImage* **65**, 449–455. ISSN: 1053-8119 (2013).
85. Scharff, C. & Nottebohm, F. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *Journal of Neuroscience* **11**, 2896–2913 (1991).
86. Scheerer, N. E. & Jones, J. A. The Predictability of Frequency-Altered Auditory Feedback Changes the Weighting of Feedback and Feedforward Input for Speech Motor Control. en. *European J Neurosci* **40**, 3793–3806. ISSN: 0953816X (2014).
87. Schultz, W. Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev* **95**, 853–951 (2015).
88. Seetharam, A. S. & Stuart, G. W. Whole genome phylogeny for 21 Drosophila species using predicted 2b-RAD fragments. *PeerJ* **1**, e226 (2013).
89. Shadmehr, R., Donchin, O., Hwang, E.-J., Hemminger, S. E. & Rao, A. Learning dynamics of reaching. *Motor cortex and voluntary movements*, 297–328 (2005).
90. Shadmehr, R. & Holcomb, H. H. Neural correlates of motor memory consolidation. *Science* **277**, 821–825 (1997).
91. Shadmehr, R., Smith, M. A. & Krakauer, J. W. Error correction, sensory prediction, and adaptation in motor control. *Annual review of neuroscience* **33**, 89–108 (2010).
92. Smith, M. A., Ghazizadeh, A. & Shadmehr, R. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS biology* **4**, e179 (2006).
93. Sober, S. J. & Brainard, M. S. Adult birdsong is actively maintained by error correction. *Nature neuroscience* **12**, 927 (2009).

94. Sober, S. J. & Brainard, M. S. Vocal learning is constrained by the statistics of sensorimotor experience. *Proceedings of the National Academy of Sciences* **109**, 21099–21103 (2012).
95. Suthers, R. A., Goller, F. & Wild, J. M. Somatosensory Feedback Modulates the Respiratory Motor Program of Crystallized Birdsong. en. *Proc Natl Acad Sci (USA)* **99**, 5680–5685. ISSN: 0027-8424, 1091-6490 (2002).
96. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* 2nd (MIT Press, Cambridge, USA, 2012).
97. Szepesvári, C. Algorithms for reinforcement learning. *Synthesis lectures on artificial intelligence and machine learning* **4**, 1–103 (2010).
98. Transtrum, M. *et al.* Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *J Chem Phys* **143**, 010901 (2015).
99. van Beers, R. J. Motor Learning Is Optimally Tuned to the Properties of Motor Noise. *Neuron* **63**, 406–417. ISSN: 0896-6273 (2009).
100. Vaswani, P. A. & Shadmehr, R. Decay of motor memories in the absence of error. *Journal of neuroscience* **33**, 7700–7709 (2013).
101. Wei, K. & Körding, K. Uncertainty of Feedback and State Estimation Determines the Speed of Motor Adaptation. English. *Frontiers Comput Neurosci* **4**. ISSN: 1662-5188. doi:[10.3389/fncom.2010.00011](https://doi.org/10.3389/fncom.2010.00011) (2010).
102. Wei, K. & Kording, K. Relevance of error: what drives motor adaptation? *Journal of neurophysiology* **101**, 655–664 (2009).
103. Wolpert, D. M. Probabilistic Models in Human Sensorimotor Control. en. *Human Movement Sci* **26**, 511–524. ISSN: 01679457 (2007).
104. Wolpert, D. M., Diedrichsen, J. & Flanagan, J. R. Principles of sensorimotor learning. *Nature Reviews Neuroscience* **12**, 739 (2011).
105. Wu, H. G., Miyamoto, Y. R., Castro, L. N. G., Ölveczky, B. P. & Smith, M. A. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature neuroscience* **17**, 312 (2014).
106. Zhou, B., Hofmann, D., Pinkoviezky, I., Sober, S. J. & Nemenman, I. Chance, long tails, and inference in a non-Gaussian, Bayesian theory of vocal learning in songbirds. *Proceedings of the National Academy of Sciences* **115**, E8538–E8546 (2018).
107. Zinn-Justin, J. *Quantum Field Theory and Critical Phenomena* 4th (Clarendon Press, Oxford, UK, 2002).