



# Chance, long tails, and inference in a non-Gaussian, Bayesian theory of vocal learning in songbirds

Baohua Zhou<sup>a</sup>, David Hofmann<sup>a,b</sup>, Itai Pinkoviezky<sup>a,b</sup>, Samuel J. Sober<sup>c</sup>, and Ilya Nemenman<sup>a,b,c,1</sup>

<sup>a</sup>Department of Physics, Emory University, Atlanta, GA 30322; <sup>b</sup>Initiative in Theory and Modeling of Living Systems, Emory University, Atlanta, GA 30322; and <sup>c</sup>Department of Biology, Emory University, Atlanta, GA 30322

Edited by Terrence J. Sejnowski, Salk Institute for Biological Studies, La Jolla, CA, and approved July 23, 2018 (received for review July 21, 2017)

**Traditional theories of sensorimotor learning posit that animals use sensory error signals to find the optimal motor command in the face of Gaussian sensory and motor noise. However, most such theories cannot explain common behavioral observations, for example, that smaller sensory errors are more readily corrected than larger errors and large abrupt (but not gradually introduced) errors lead to weak learning. Here, we propose a theory of sensorimotor learning that explains these observations. The theory posits that the animal controls an entire probability distribution of motor commands rather than trying to produce a single optimal command and that learning arises via Bayesian inference when new sensory information becomes available. We test this theory using data from a songbird, the Bengalese finch, that is adapting the pitch (fundamental frequency) of its song following perturbations of auditory feedback using miniature headphones. We observe the distribution of the sung pitches to have long, non-Gaussian tails, which, within our theory, explains the observed dynamics of learning. Further, the theory makes surprising predictions about the dynamics of the shape of the pitch distribution, which we confirm experimentally.**

power-law tails | sensorimotor learning | dynamical Bayesian inference | vocal control

**L**earned 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 (1, 2), but rather a significant part of it is controlled by animals themselves and is used for exploration during learning (3, 4). 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 (1, 5–8).

Since many complex behaviors are characterized by both tightly controlled motor variability and 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 the classic Rescorla–Wagner (RW) model (9), as well as more modern approaches belonging to the family of reinforcement learning (10–12), Kalman filters (13, 14), or dynamical Bayesian filter models (15, 16). 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 (17–19). Such theories have addressed many important experimental questions, such as evaluating the optimality of the learning process (13, 20–23), accounting for multiple temporal scales in learning (7, 13, 24, 25), identifying the complexity of behaviors that can be learned (26), and pointing out how the necessary computations could be performed using networks of spiking neurons (12, 27–31).

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 (6, 22, 32, 33). One can rescue traditional theories by allowing the animal to reject large errors as “irrelevant”—unlikely to have come from its own actions (22, 34). 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 (6, 33).

Here we present a theory (Fig. 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. 2C), and the pitch variability is much smaller in a song directed at a female vs. the undirected song, suggesting that the variability and hence its non-Gaussian tails

## Significance

**Skilled behaviors are learned through a series of trial and error. The ubiquity of such processes notwithstanding, current theories of learning fail to explain how the speed and the magnitude of learning depend on the pattern of experienced sensory errors. Here, we introduce a theory, formulated and tested in the context of a specific behavior—vocal learning in songbirds. The theory explains the observed dependence of learning on the dynamics of sensory errors. Furthermore, it makes additional strong predictions about the dynamics of learning that we verify experimentally.**

Author contributions: B.Z., S.J.S., and I.N. designed research; B.Z., D.H., and I.N. developed the theory; S.J.S. and I.N. managed research; B.Z., D.H., I.P., S.J.S., and I.N. performed research; B.Z. analyzed data; and B.Z., D.H., I.P., S.J.S., and I.N. wrote the paper.

The authors declare no conflict of interest.

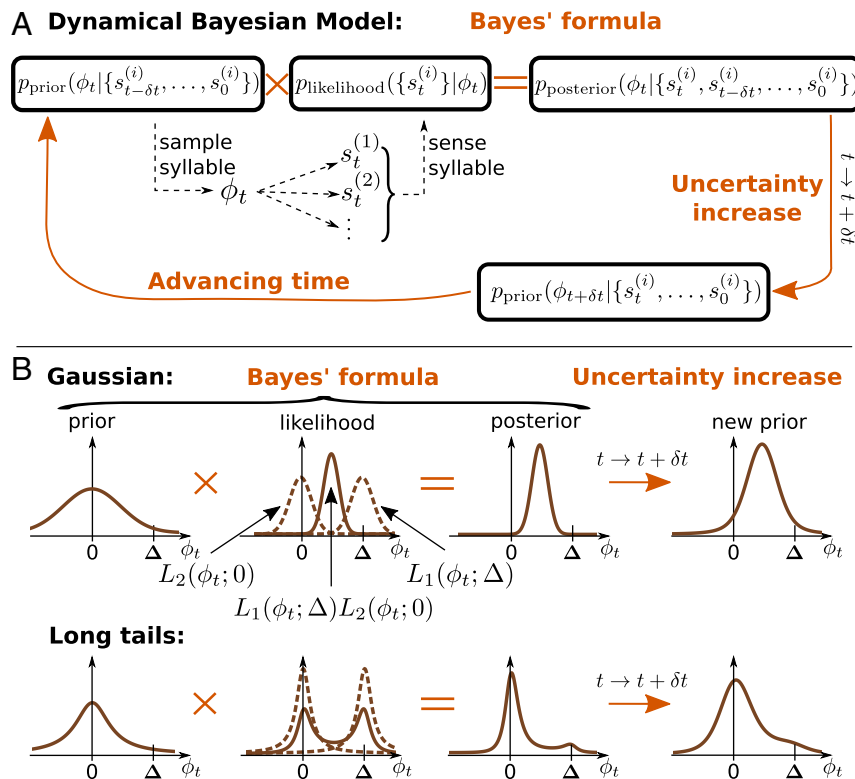
This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: The code and data reported in this paper have been deposited on GitHub and are available at <https://github.com/EmoryUniversityTheoreticalBiophysics/NonGaussianLearning>.

<sup>1</sup>To whom correspondence should be addressed. Email: [ilya.nemenman@emory.edu](mailto:ilya.nemenman@emory.edu).

Published online August 20, 2018.



**Fig. 1.** The dynamical Bayesian model (Bayesian filter). (A) A Bayesian filter consists of the recursive application of two general steps (35): (i) an observation update, which corresponds to novel sensory input and updates the underlying probability distribution of plausible motor commands using Bayes' formula, and (ii) a time evolution update, which denotes the temporal propagation and corresponds to uncertainty increasing with time (main text); here the probability distribution is updated by convolution with a propagator. These two steps are repeated for each new piece of 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,  $\phi_t$ , 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 the animal compensating its pitch toward  $+\Delta$ . 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  $\phi_t$ . 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.

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.

## Results

**Biological Model System.** Vocal control in songbirds is a powerful model system for examining sensorimotor learning of complex tasks (36). The phenomenology we are trying to explain arises from experimental approaches to inducing song plasticity (33). Songbirds sing spontaneously and prolifically and use auditory feedback to shape their songs toward 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 (33), 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. 24, 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 three 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. 2B).

**Data.** We use experimental data collected in our previous work (8, 33) to develop our mathematical model of learning. As detailed in ref. 39, we used a virtual auditory feedback system (8, 40) to evoke sensorimotor learning in adult songbirds. For this, miniature headphones were custom fitted to each bird and used to provide online auditory 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 d 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).



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 (46). Overall, this results in the distribution of motor outputs after one cycle of the model

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

where  $Z$  is the normalization constant.

We choose  $\delta t$  to be 1 d 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 *A* and *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 one affected by the headphones and the second one not affected, 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 (more on the minus sign below). Note that since  $\phi_t$  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  $\phi_t$  was the correct motor command or as a likelihood of  $\phi_t$  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  $\phi_t$  to lead to the target is negative and vice versa. While potentially confusing, this is the same convention that is used in all filtering applications—a positive sensory signal means 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 Eq. 1 as

$$p_{\text{prior}}(\phi_{t+\delta t}) = \frac{1}{Z} \int p_{\text{prop}}(\phi_{t+\delta t}|\phi_t) \times L_1(\phi_t; \Delta) L_2(\phi_t; 0) p_{\text{prior}}(\phi_t) d\phi_t, \quad [2]$$

where 0 and  $\Delta$  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. Note that the likelihoods for the shifted and unshifted modalities are

centered around  $\Delta$  and 0, respectively, and bias the learning of what should be sung toward these centers irrespective of the current value of  $\phi_t$ . 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 1 d in the Bayesian update equation above.

As illustrated in Fig. 1*B*, 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  $\Delta$ , no matter how large this error is, which conflicts with the experimental results in songbirds and other species (5, 8, 22, 36). 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  $\Delta$  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 Eq. 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 Eq. 2 with no perturbation,  $\Delta = 0$ . We have verified numerically that a wide variety of shapes of  $L_1$ ,  $L_2$ , and  $p_{\text{prop}}$  result in learning dynamics that can approximate the experimental data (*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 distributions (47) (*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 (*Materials and Methods*). Here we assume symmetric stable distributions, truncated at  $\pm 8$  semitones (*Materials and Methods*). Such distributions are characterized by three parameters: the stability parameter  $\alpha$  (measuring the proportion of the probability in the tails), the scale or width parameter  $\gamma$ , and the location or the center parameter  $\mu$  (the latter can be predetermined to be 0,  $\Delta$ , or the previous time-step value in our case). For three distributions  $L_1(\phi_t; \Delta)$ ,  $L_2(\phi_t; 0)$ , and  $p_{\text{prop}}$ , this results in the total of six unknown parameters.

**Fits to Data.** We fitted the set of six parameters of our model simultaneously to all of the data shown in Fig. 2. Our dataset consists of 23 individual experiments across five experimental conditions: four constant pitch-shift learning curves and one gradual, staircase-shift learning curve (see *Materials and Methods* for details). As mentioned previously, birds learn the best (larger and faster compensation) for smaller perturbations, here

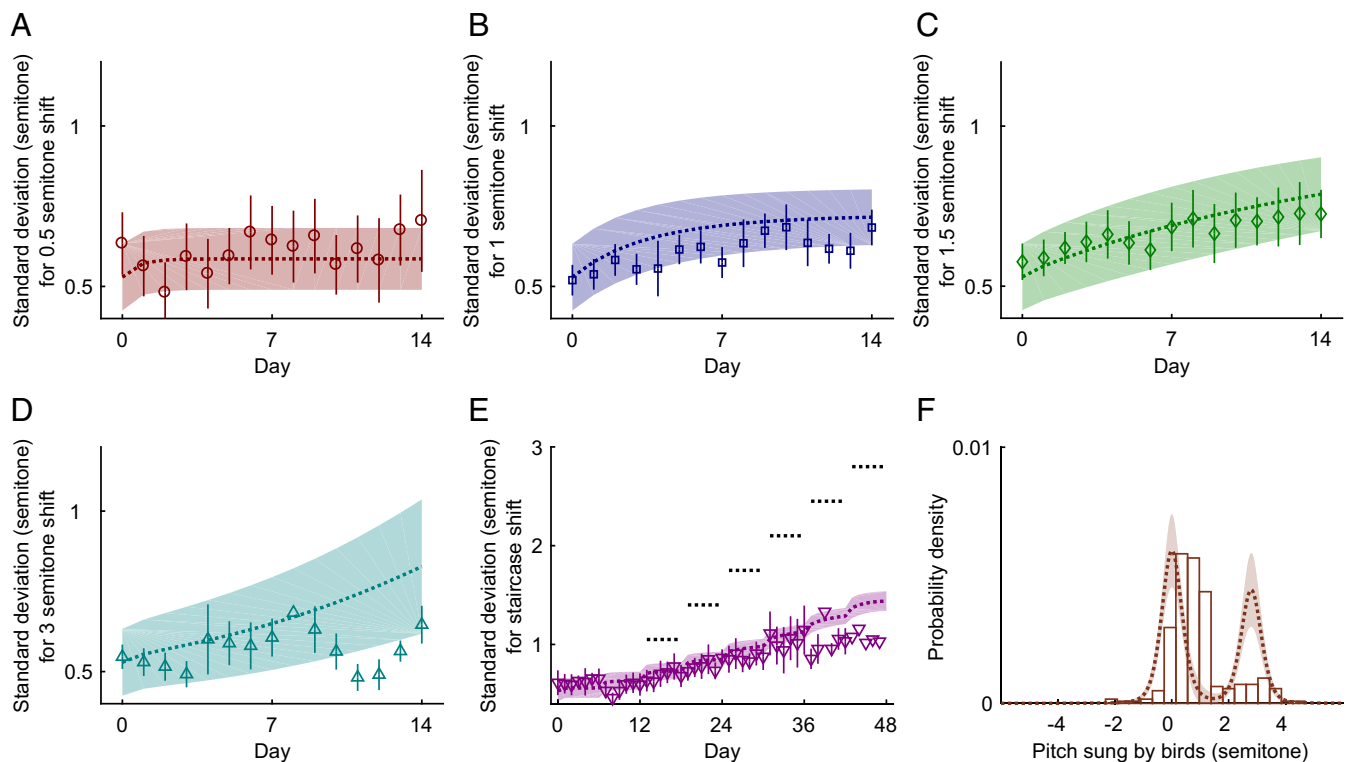
0.5 semitone (Fig. 2A). In contrast, for a large 3-semitone perturbation, the birds do not compensate at all within the 14 d of the experiment. However, the birds are able to learn and compensate large perturbations when the perturbation increases gradually, as in the staircase experiment in Fig. 2B. Importantly, the baseline distribution (Fig. 2C) 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.

**Predictions.** 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 (SD in this case) of the pitch. As shown in Fig. 3A–E, our model correctly predicted time-dependent behaviors in the SD in both single-shift (Fig. 3A–D) and staircase-shift experiments (Fig. 3E) with surprising accuracy. We stress again that no new parameter fits were done for these curves. Potentially

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

## Discussion

We introduced a mathematical framework within the class of observation–evolution models (35) 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



**Fig. 3.** Predictions of our model using the parameter values obtained from fitting the data shown in Fig. 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. The dotted lines with 1-SD bands represent model predictions. Our model correctly predicts the behaviors of the SDs of the pitch distributions. Specifically, the best-fit model lines predict increases in the SD in B, C, and E, which correspond to 1 semitone, 1.5 semitones, and the staircase shift, respectively. At the same time, the data show that the SD increases for B, C, and E (*P* value for a positive dependence of the SD when regressed on time is  $4 \times 10^{-4}$ ,  $5 \times 10^{-5}$ , and  $< 10^{-6}$  for B, C, and E, respectively). (F) Our model predicts that, at the end of the staircase experiment (mean and SD shown in Figs. 2B and 3E, respectively), the pitch distribution should be bimodal, while it is unimodal initially (compare Fig. 2C). This is also supported by the data. Specifically, a fit with a mixture of two Gaussian peaks has an Akaike's information criterion score higher than a fit to a single Gaussian by 50 (in decimal log units), which is highly statistically significant (the data here are 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 ( $0.59 \pm 0.04$  of a semitone and  $2.17 \pm 0.62$  semitones, with error bars obtained by bootstrapping), illustrating the true bimodality. Neither the data nor the models show unambiguous bimodality in other learning cases.

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 (*Materials and Methods*) that traditional linear models with Gaussian statistics (13) 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 (Figs. 2C and 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 timescales of adaptation (Fig. 2A and B) using a nonlinear dynamical equation with just a single timescale of the update (1 d). As with other key aspects of the model, this phenomenon results from the non-Gaussianity of the distributions used and specifically the many timescales 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 timescales (7, 13). While multiple timescales could be needed to account for other features of the adaptation, our model clearly avoids this for the present data. We expect that an extension of our model to include multiple explicit timescales 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. 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 (25, 39): Small overlaps result in slower and smaller learning, so that different overlaps lead to different timescales. 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.

We have chosen the family of so-called Lévy alpha-stable distributions to provide 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 a 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 (48, 49), which are not easily addressable given current experimental limitations. Like in many other behavioral modeling studies (13, 34), 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 a 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 (50–53).

## Materials and Methods

**Experiments.** The data used are taken from the experiments in ref. 33 and are described in detail there. Briefly, subjects were nine male adult Bengalese finches (females do not produce song) aged over 190 d. 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 are used. The same birds were used in multiple (but not all) pitch-shift experiments separated by at least 32 d. Changes in vocal pitch were measured in semitones, which are 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 indicate the error of the mean that accounts for variances both across individual birds and within one bird. Specifically, if  $y_i^\mu$  represents the pitch of the  $i$ th vocalization from the  $\mu$ 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 as

$$\delta y = \sqrt{\frac{\sum_\mu (\bar{y}^\mu - \bar{y})^2}{(m-1)^2} + \frac{1}{m} \sum_\mu \left[ \frac{\sum_i (y_i^\mu - \bar{y}^\mu)^2}{(n-1)^2} \right]}, \quad [3]$$

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.

**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 (47). By the generalized central limit theorem, the probability distributions of sums of a large number of i.i.d. random variables with infinite variances tend to be stable distributions (47). 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 (47):

$$f(y; \alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} du e^{-|u|^\alpha} \cos(yu). \quad [4]$$

Thus, any symmetric stable distribution is characterized by three parameters: the type, or the tail weight, parameter  $\alpha$ ; the scale parameter  $\gamma$ ; and the center  $\mu$ .  $\alpha$  takes the range  $(0, 2]$  (47). If  $\alpha = 2$ , the corresponding distribution is the Gaussian, and if  $\alpha = 1$ , it is the Cauchy distribution.  $\gamma$  can be any positive real number, and  $\mu$  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 (54), communication systems (55), and brain imaging (56), there are many algorithms to compute it approximately. We used the method of ref. 57. 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 (58).

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

**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 nine parameters of which we need to fit six from data, namely the type parameters  $\alpha$  and the scale parameters  $\gamma$ , while the center parameters  $\mu$  are predetermined: The two likelihoods are at 0 and  $\Delta$ , respectively, while the propagation kernel is centered around the previous time-step value (Eq. 1). The prior (and accordingly the posterior) is a discrete distribution with a resolution of 1,600 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 Eq. 2. As described earlier in Mathematical Model, the initial prior is given as the steady-state distribution after repeatedly applying Eq. 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. 2D. Similarly, the likelihood functions must be long tailed; otherwise the combined likelihood would not be bimodal as found empirically in Fig. 3F, and large perturbations would not be rejected—even if the prior was long tailed (e.g., 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 datasets to our disposition: the  $\chi^2$  for four adaptations of the means to the respective constant shifts (Fig. 2A), the  $\chi^2$  for the adaptation of the mean to the staircase shifts (Fig. 2B), and the log-likelihood of the observed baseline pitch probability distribution (Fig. 2C). Using Eq. 3 to define the error of the mean, we calculate  $\chi^2$  as

$$\chi^2 = \frac{1}{T} \sum_{t=1}^T \frac{(\bar{\phi}_t - \bar{y}_t)^2}{\delta y(t)^2}, \quad [5]$$

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 SE 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 are 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 datasets are very constraining, we do not think that we lose 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. Fig. 4 illustrates this by showing the quality of fit as a function of each pair of  $(\alpha, \gamma)$ , 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 (59, 60), 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 Figs. 2, 3, and 5), we have

$$\langle f(\vec{\theta}) \rangle = \frac{\sum_{\vec{\theta}} e^{-\mathcal{L}(\vec{\theta})} f(\vec{\theta})}{\sum_{\vec{\theta}} e^{-\mathcal{L}(\vec{\theta})}}. \quad [6]$$

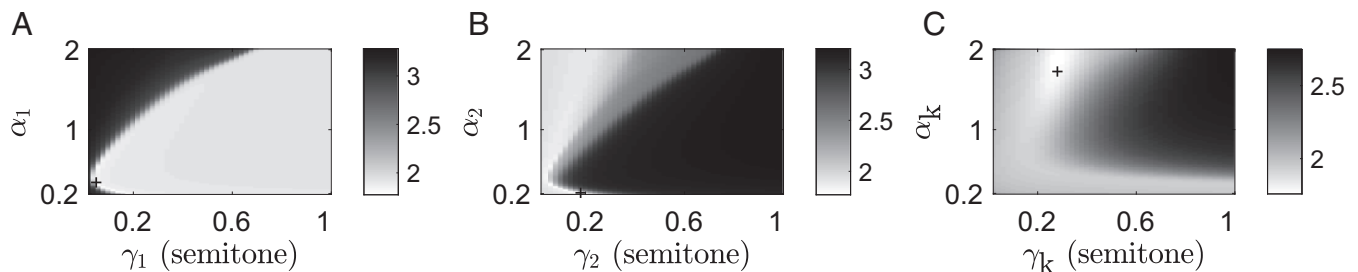
For the SDs, denoted by shaded regions in Figs. 2, 3, and 5, we have

$$\sigma_f^2 = \langle f(\vec{\theta})^2 \rangle - \langle f(\vec{\theta}) \rangle^2. \quad [7]$$

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  $\alpha$  and  $[0.01, 8]$  for each  $\gamma$ ). The extremal values avoid  $\alpha = 0, 2$  and  $\gamma = 0$ , which are singular and dramatically slow down computations. Thus, there are a total of 13 grid points for each parameter and a total of  $13^6 \approx 4.8 \cdot 10^6$  total parameter samples.

**Choice of the Shape of Distributions.** For Figs. 2 and 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 to the stable distribution model. For example, Fig. 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). The detailed shape of the distributions seems less important than the existence of the heavy tails.

**Linear Dependence on Pitch Shift in a Kalman Filter with Multiple Timescales.** 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. (13). This Kalman filter model belongs to the family of Bayes filters, which are dynamical models describing



**Fig. 4.** (A–C) 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 shades mean a better fit. Because of the logarithmic scaling, small changes in the shading 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. 2C.





$$\tau_i^{-1}d_s^i = f_i\Delta - f_i \sum_{j=1}^n (1 - \tau_j^{-1})d_s^j. \quad [17]$$

These equations are solved by

$$d_s^i = \Delta \frac{f_i \tau_i}{1 + \sum_{j=1}^n f_j (\tau_j - 1)}, \quad [18]$$

1. Shadmehr R, Smith MA, Krakauer JW (2010) Error correction, sensory prediction, and adaptation in motor control. *Annu Rev Neurosci* 33:89–108.
2. McDonnell MD, Ward LM (2011) The benefits of noise in neural systems: Bridging theory and experiment. *Nat Rev Neurosci* 12:415–426.
3. Neuringer A (2002) Operant variability: Evidence, functions, and theory. *Psychon Bull Rev* 9:672–705.
4. Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature* 433:638–643.
5. Linkenhoker BA, Knudsen EI (2002) Incremental training increases the plasticity of the auditory space map in adult barn owls. *Nature* 419:293–296.
6. Knudsen EI (2002) Instructed learning in the auditory localization pathway of the barn owl. *Nature* 417:322–328.
7. Smith MA, Ghazizadeh A, Shadmehr R (2006) Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol* 4:e179.
8. Sober SJ, Brainard MS (2009) Adult birdsong is actively maintained by error correction. *Nat Neurosci* 12:927–931.
9. Rescorla R, Wagner A (1972) *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical Conditioning II*, eds Black A, Prokasy W (Appleton-Century-Crofts, New York), pp 64–99.
10. Joel D, Niv Y, Ruppin E (2002) Actor–critic models of the basal ganglia: New anatomical and computational perspectives. *Neural Networks* 15:535–547.
11. Sutton RS, Barto AG (2012) *Reinforcement Learning: An Introduction* (MIT Press Cambridge, MA), 2nd Ed.
12. Lak A, Stauffer W, Schultz W (2016) Dopamine neurons learn relative chosen value from probabilistic rewards. *eLife* 5:e18044.
13. Kording KP, Tenenbaum JB, Shadmehr R (2007) The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nat Neurosci* 10:779–786.
14. Wolpert DM (2007) Probabilistic models in human sensorimotor control. *Hum Mov Sci* 26:511–524.
15. Gallistel CR, Mark TA, King AP, Latham PE (2001) The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *J Exper Psych Anim Behav Proc* 27:354–372.
16. Gershman SJ (2015) A unifying probabilistic view of associative learning. *PLoS Comput Biol* 11:e1004567.
17. Doya K, Sejnowski T (2000) A computational model of avian song learning. *The New Cognitive Neurosciences*, ed Gazzaniga M (MIT Press, Cambridge, MA), 2nd Ed, pp 469–482.
18. Fiete I, Seung H (2009) Birdsong learning. *Encyclopedia of Neuroscience*, ed Squire L (Academic, Oxford), pp 227–239.
19. Farries M, Fairhall A (2007) Reinforcement learning with modulated spike timing-dependent synaptic plasticity. *J Neurophys* 98:3648–3665.
20. Donchin O, Shadmehr R (2001) Linking motor learning to function approximation: Learning in an unlearnable force field. *Advances in Neural Information Processing Systems 14*, eds Dietterich T, Becker S, Ghahramani Z (MIT Press, Cambridge, MA), Vol 1, p 7.
21. van Beers RJ (2009) Motor learning is optimally tuned to the properties of motor noise. *Neuron* 63:406–417.
22. Wei K, Körding K (2009) Relevance of error: What drives motor adaptation? *J Neurophysiol* 101:655–664.
23. Beck JM, Ma WJ, Pitkow X, Latham P, Pouget A (2012) Not noisy, just wrong: The role of suboptimal inference in behavioral variability. *Neuron* 74:30–39.
24. Wei K, Körding K (2010) Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Front Comput Neurosci* 4:11.
25. Kelly CW, Sober SJ (2014) A simple computational principle predicts vocal adaptation dynamics across age and error size. *Front Integr Neurosci* 8:9.
26. Genewein T, Hez E, Razzaghpour Z, Braun DA (2015) Structure learning in Bayesian sensorimotor integration. *PLoS Comput Biol* 11:27.
27. Shadmehr R, Donchin O, Hwang EJ, Hemminger SE, Rao A (2005) Learning dynamics of reaching. *Motor Cortex and Voluntary Movements: A Distributed System for Distributed Function*, eds Riehle A, Vaadia E (CRC Press, Boca Raton, FL), pp 297–328.
28. Dayan P, Niv Y (2008) Reinforcement learning: The good, the bad and the ugly. *Curr Opin Neurobiol* 18:185–196.
29. Fischer BJ, Peña JL (2011) Owl's behavior and neural representation predicted by Bayesian inference. *Nat Neurosci* 14:1061–1066.
30. Neymotin SA, Chadderdon GL, Kerr CC, Francis JT, Lytton WW (2013) Reinforcement learning of two-joint virtual arm reaching in a computer model of sensorimotor cortex. *Neural Comput* 25:3263–3293.
31. Schultz W (2015) Neuronal reward and decision signals: From theories to data. *Physiol Rev* 95:853–951.
32. Robinson FR, Noto CT, Bevans SE (2003) Effect of visual error size on saccade adaptation in monkey. *J Neurophysiol* 90:1235–1244.
33. Sober SJ, Brainard MS (2012) Vocal learning is constrained by the statistics of sensorimotor experience. *Proc Natl Acad Sci USA* 109:21099–21103.
34. Hahnloser R, Narula G (2017) A Bayesian account of vocal adaptation to pitch-shifted auditory feedback. *PLoS ONE* 12:e0169795.
35. Kaipio J, Somersalo E (2004) *Statistical and Computational Inverse Problems* (Springer, New York).
36. Brainard MS, Doupe A (2002) What songbirds teach us about learning. *Nature* 417:351–358.
37. Nemenman I, Shafee F, Bialek W (2002) *Entropy and inference, revisited. Advances in Neural Information Processing Systems (NIPS)*, eds Dietterich T, Becker S, Ghahramani Z, (MIT Press, Cambridge, MA) Vol 14, pp 471–478.
38. Nemenman I (2011) Coincidences and estimation of entropies of random variables with large cardinalities. *Entropy* 13:2013–2023.
39. Kuebrich B, Sober S (2015) Variations on a theme: Songbirds, variability, and sensorimotor error correction. *Neuroscience* 296:48–54.
40. Hoffmann LA, Kelly CW, Nicholson DA, Sober SJ (2012) A lightweight, headphones-based system for manipulating auditory feedback in songbirds. *J Vis Exp* 69:e50027.
41. Ölveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3:e153.
42. Nizioletk C, Nagarajan S, Houde J (2013) What does motor efference copy represent? Evidence from speech production. *J Neurosci* 33:16110–16116.
43. Suthers RA, Goller F, Wild JM (2002) Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. *Proc Natl Acad Sci USA* 99:5680–5685.
44. Liu Y, et al. (2015) Selective and divided attention modulates auditory–vocal integration in the processing of pitch feedback errors. *Eur J Neurosci* 42:1895–1904.
45. Scheerer NE, Jones JA (2014) The predictability of frequency-altered auditory feedback changes the weighting of feedback and feedforward input for speech motor control. *Eur J Neurosci* 40:3793–3806.
46. Zinn-Justin J (2002) *Quantum Field Theory and Critical Phenomena* (Clarendon, Oxford), 4th Ed.
47. Nolan JP (2015) *Stable Distributions - Models for Heavy Tailed Data* (Birkhauser, Boston).
48. Hahn A, Kryslar A, Sturdy C (2013) Female song in black-capped chickadees (*Parus atricapillus*): Acoustic song features that contain individual identity information and sex differences. *Behav Proc* 98:98–105.
49. Mets D, Braind M (2018) Genetic variation interacts with experience to determine interindividual differences in learned song. *Proc Natl Acad Sci USA* 115:421–426.
50. Fiser J, Berkes P, Orbán G, Lengyel M (2010) Statistically optimal perception and learning: From behavior to neural representations. *Trends Cogn Sci* 14:119–130.
51. Buesing L, Bill J, Nessler B, Maass W (2011) Neural dynamics as sampling: A model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput Biol* 7:e1002211.
52. Kappel D, Habenschuss S, Legenstein R, Maass W (2015) Network plasticity as Bayesian inference. *PLoS Comput Biol* 11:e1004485.
53. Petrovici MA, Bill J, Bytschok I, Schemmel J, Meier K (2016) Stochastic inference with spiking neurons in the high-conductance state. *Phys Rev E* 94:042312.
54. Mittnik S, Paoletta MS, Rachev ST (2000) Diagnosing and treating the fat tails in financial returns data. *J Empir Finance* 7:389–416.
55. Nikias CL, Shao M (1995) *Signal Processing with Alpha-Stable Distributions and Applications, Adaptive and Learning Systems for Signal Processing, Communications, and Control* (Wiley, New York).
56. Salas-Gonzalez D, Górriz JM, Ramirez J, Illán IA, Lang EW (2013) Linear intensity normalization of FP-CIT SPECT brain images using the  $\alpha$ -stable distribution. *NeuroImage* 65:449–455.
57. Belov IA (2005) On the computation of the probability density function of  $\alpha$ -stable distributions. *Math Model Anal* 2:333–341.
58. Bergström H (1952) On some expansions of stable distribution functions. *Ark Mat* 2:375–378.
59. Gutenkunst R, et al. (2007) Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol* 3:1871–1878.
60. Transtrum M, et al. (2015) Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *J Chem Phys* 143:010901.