

RESEARCH ARTICLE

Control of Movement

Variable but not random: temporal pattern coding in a songbird brain area necessary for song modification

⁶ S. E. Palmer,¹ B. D. Wright,²⁺ A. J. Doupe,²⁺ and ⁶ M. H. Kao³

¹Department of Organismal Biology and Anatomy, Department of Physics, Committee on Computational Neuroscience, University of Chicago, Chicago, Illinois; ²Departments of Physiology and Psychiatry, Keck Center for Integrative Neuroscience, University of California, San Francisco, California; and ³Department of Biology & Program in Neuroscience, Tufts University, Medford, Massachusetts

Abstract

Practice of a complex motor gesture involves motor exploration to attain a better match to target, but little is known about the neural code for such exploration. We examine spiking in a premotor area of the songbird brain critical for song modification and quantify correlations between spiking and time in the motor sequence. While isolated spikes code for time in song during performance of song to a female bird, extended strings of spiking and silence, particularly bursts, code for time in song during undirected (solo) singing, or "practice." Bursts code for particular times in song with more information than individual spikes, and this spike-spike synergy is significantly higher during undirected singing. The observed pattern information cannot be accounted for by a Poisson model with a matched time-varying rate, indicating that the precise timing of spikes in both bursts in undirected singing and isolated spikes in directed singing code for song with a temporal code. Temporal coding during practice supports the hypothesis that lateral magnocellular nucleus of the anterior nidopallium neurons actively guide song modification at local instances in time.

NEW & NOTEWORTHY This paper shows that bursts of spikes in the songbird brain during practice carry information about the output motor pattern. The brain's code for song changes with social context, in performance versus practice. Synergistic combinations of spiking and silence code for time in the bird's song. This is one of the first uses of information theory to quantify neural information about a motor output. This activity may guide changes to the song.

birdsong; information theory; motor performance; motor practice; temporal coding

INTRODUCTION

The vocalizations of songbirds represent a classic learned sensorimotor skill: initially immature, variable sounds are gradually refined through extensive rehearsal, and auditory feedback is critical throughout life for song learning and modification. A variety of evidence suggests that a corticobasal ganglia circuit dedicated to song (known as the anterior forebrain pathway, AFP) is crucial for both juvenile and adult plasticity. First, lesions or inactivation of the cortexlike outflow nucleus of this circuit, the lateral magnocellular nucleus of the anterior nidopallium (LMAN; see Fig. 1*A*), cause an abrupt reduction in song variability and a failure to progress toward a good match to the target song (1–6). Second, even the subtle, trial-by-trial variations in the acoustic structure of adult song can be used to drive rapid, adaptive changes in song (7–9), and inactivating LMAN blocks this plasticity (7, 10). In addition, LMAN and the rest of the AFP exhibit singing-related premotor activity (11–13), and stimulation in subregions of these areas can alter targeted syllables in a consistent manner (14). Together, these results suggest that LMAN's role in reinforcement learning is to drive variability in a biased (7, 8, 10, 14, 15), not purely random, manner. An additional feature of this behavior and its neural substrates is that both reveal precise attention to specific locations in song. Here, we use information theory to explore the neural code in LMAN to uncover spike patterns that correlate with precise times in song.

⁺ B. D. Wright deceased 10 January 2008. A. J. Doupe deceased 24 October 2014. Correspondence: S. E. Palmer (sepalmer@uchicago.edu).

Submitted 16 January 2019 / Revised 7 December 2020 / Accepted 7 December 2020





Figure 1. Single-spike information about time in song is higher during directed song. *A*: schematic diagram of the main brain areas involved in song learning and production and corresponding areas in the mammalian brain. (AFP, anterior forebrain pathway labeled in gray; GPi, internal segment of the globus pallidus; HVC, used as a proper name; SNr, substantia nigra pars reticulata; VTA, ventral tegmental area. LMAN, RA, and X are defined in the main text.) *B*: spike rasters (*top*) and corresponding averaged firing rates smoothed with a Gaussian kernel with standard deviation (SD) = 2 ms [peri-song time histograms (PSTHs), *bottom*) for one LMAN neuron during directed (left) and undirected (right) singing show increased firing rate, more bursts, and more apparent noise during undirected singing. The PSTHs are in red for directed singing and blue for undirected, with the directed pattern overlaid in light red for comparison. *C*: two sample traces to illustrate a PSTH that carries information about time in song (upper trace) vs. a (flat) PSTH that carries no such information (second trace) The lower three traces are a schematic of the mutual information about time in song calculation, showing how the PSTH is processed to compute the average signal-to-noise (SNR) ratio across time. *D*: information rate from single-spike arrival times in LMAN during directed song (DIR, red) and undirected song (UNDIR, blue), measured using *Eq. 1*. Data are shown with a bin size, $\Delta t = 2$ ms. Lines connect data from single neurons with recordings were made during only one context. Stars and gray bars indicate the means ± standard error (SE) across all sites. ***P* < 0.001, ****P* < 0.0001.

Prior work has shown that both song and neural variability are strongly modulated by social context. When males sing courtship song directed at females, both the acoustic structure of individual syllables and syllable sequence are more stereotyped than when males sing alone ["undirected" song (2, 14, 16, 17)]. These context-dependent changes in song variability are accompanied by striking changes in the singing-related activity of LMAN neurons: LMAN neurons exhibit more variable spike timing and more frequent burst firing during undirected singing [e.g., Fig. 1B (12, 13)]. Both spike timing variability in LMAN and song variability are even higher in juvenile birds actively engaged in sensorimotor learning (3, 4, 16, 18–20). Strikingly, the same lesions or inactivation of LMAN that eliminate song plasticity also reduce song variability during undirected song to the level observed during directed singing (2, 3, 6). In addition, manipulations of the AFP circuit that specifically eliminate LMAN bursts (but not single-spike firing) also eliminate the bird's ability to change song in response to altered auditory feedback (21). The differential effectiveness of burst firing versus single spikes for driving plasticity has been demonstrated in a variety of in vivo and in vitro settings [see (22) and (23) for reviews]. Together, these results suggest that the variable LMAN bursting typical during undirected singing drives song variability and plasticity. Moreover, they support the idea that undirected song reflects a "practice" state, in which behavioral variability enables maintenance and/or optimization of song, whereas directed song reflects a "performance" state in which a male sings his current best version of song (3, 16, 24–26).

Although previous work had hypothesized that LMAN injects random variability in the motor pathway that enables song change (4, 27, 28), recent studies demonstrate that LMAN firing can target particular local variations in song output (7, 8, 10, 14, 15). Here we ask what code LMAN might be using to drive this kind of motor variation, using a novel application of information theoretic analysis to motor systems. Although information theory has been extensively used to study coding in sensory systems, very little information-theoretic analysis has been done for motor systems, with some notable exceptions in the birdsong literature (29, 30). To quantitatively dissect the neural code in LMAN for time in song, we compute the mutual information (31-33) between temporal patterns of spiking and time in the motor sequence. We find specific, temporally extended patterns of spiking carry information about song, particularly during undirected singing, and that the neural code is fundamentally different in these two behavioral contexts. We compare our analysis to Poisson models with a timevarying firing rate matched to each cell's activity and show that rate fluctuations alone do not account for the coding we see in the real data. Intriguingly, the temporal coding we observe during practice contains a high degree of spike-spike synergy; patterns of spikes carry more information than the sum of their constituent spikes.

METHODS

Experimental Procedures

Recordings.

We analyzed single-unit recordings from 28 sites in LMAN in nine adult male zebra finches, collected as part of a previous study (13) approved by the Institutional Animal Care and Use Committee at the University of California, San Francisco. Spiking activity was measured during bouts of singing both to a female bird and when the male bird was alone, as well as when the male bird was silent.

Song warping.

Songs were segmented and aligned using linear warping, as previously described (13). This warping allows us to align spike trains across renditions but does not remove variability in spectral features of song from rendition to rendition. Warping primarily affects silent periods and does slightly reduce the jitter in spike timing across trials. The difference in the correlation coefficient between the trial-averaged raw data and the time-warped data ranged between 0 and 0.12 (13). These changes did not quantitatively or qualitatively change the differences in the pattern entropy we measure between social contexts.

Defining spike patterns.

Time was binned in 0.5-, 1-, 2-, 3-, and 4-ms bins, and the particular bin size chosen did not affect our results. Data shown in the figures are for 2-ms bins. Patterns of spikes were classified as bursts when the instantaneous firing rate exceeded 200 spikes/s at any time within the window, i.e., when any two spikes were separated by fewer than two bins of silence.

Information in spikes, spike count, and spike patterns.

To quantify information contained in spikes, we computed the information about time in song using the direct method (33). For single spikes, the information about time in song is related to the average modulation of the firing rate divided by the mean rate over the entire trial. To proceed beyond single-spike information, we characterize the average entropy in temporal patterns as a function of time in song. This socalled noise entropy is subtracted from the total entropy of patterns (vocabulary size minus noise), yielding *Eq. 4* in RESULTS. To relate this equation to the information contained in single spikes computed via the peri-song time histogram (PSTH) of the cell (*Eq. 3* in RESULTS), we first write out explicitly the entropy quantities in *Eq. 4*:

$$I(\text{patterns}; t) = S(\text{patterns}) - \langle S(\text{patterns} | t) \rangle_t$$
$$= -\sum_{w_i} p(w_i) \log (p(w_i)) + \frac{1}{n_T} \sum_t \sum_{w_i} p(w_i | t) \log (p(w_i | t))$$
(1)

where n_T is the number of time bins in the song and, for brevity, we have replaced the label "pattern" with w_i , to denote the binary "word" that defines that particular pattern of spiking and silence. Noting that $\sum_{w_i} p(w_i) = \frac{1}{n_T} \sum_t \sum_{w_i} p(w_i \mid t)$, we can gather terms such that

$$I(\text{patterns}; t) = \frac{1}{n_T} \sum_{t} \sum_{w_i} p(w_i \mid t) \log\left(\frac{p(w_i \mid t)}{p(w_i)}\right)$$
$$= \sum_{w_i} p(w_i) \cdot \frac{1}{n_T} \sum_{t} \frac{p(w_i \mid t)}{p(w_i)} \log\left(\frac{p(w_i \mid t)}{p(w_i)}\right) \qquad (2)$$
$$= \sum_{i} p_i \cdot X_i$$

where the similarity to *Eq. 3* is evident if we identify the firing rate of a particular pattern, $r_i(t)$, with its time-dependent probability per unit time, $p(w_i|t)/\Delta t$. The information components in *Eq. 2* are not proper information quantities on their own, except if the absence of that particular pattern carries no information, and should be thought of as components of the total pattern information, which is a proper information quantity.

In a similar fashion to the computation of the total pattern information, the total count in the length T window can replace pattern, and this entropy can be computed both overall and as a function of time in song. Other coding schemes can be compared with the total pattern information, as partitions of the patterns into different bins (count and time of the first spike in the window, for example). All information measures are prone to sampling biases. Sampling issues are a function of both the length of the temporal window we use to evaluate spike patterns in (with longer windows, with more



bins, we have an exponential explosion in how much data we need to properly measure information, which is why we restrict our analysis to small snippets) and the vocabulary of spike patterns for that site. Entropy is a measure of that vocabulary. We need approximately 10 times the effective number of patterns in the data, which is 2^{S} . To correct for this and diagnose which sites to exclude from our analysis, we use linear extrapolation (LE) to the infinite data size limit (31, 33–36), using 50 random subsamples of the data, at each of the following data fractions: 0.95, 0.90, 0.85, 0.80, 0.75, 0.70, 0.60, and 0.50. Error bars on information quantities were estimated by extrapolating the standard deviation of the 50% data samples and dividing by $\sqrt{2}$. Here, we are assuming a 1/n "law of large numbers" scaling of the variance. We used quadratic extrapolation to diagnose sampling issues in these same data. When the quadratic correction exceeded 5% of the measured information for 100% of the data, or if the linear correction exceeded 10% of the measured information, we discarded the site from our analysis. This left us with 24 total sites, which for directed song had 12-59 repeats and for undirected song had 38-157 repeats. A larger number of repeats is necessary for neural data recorded during undirected singing, because neurons tend to have more entropy in this social context. Information is also computed for shuffled data, where all labels have been randomly reassigned. Any residual information in these shuffled data is due solely to data limitations. We exclude data when the shuffle information is not within error bars of zero. We also exclude sites for which the total number of observed patterns of spiking and silence was greater than the number of trials. These values are illustrated in Fig. 2. If we restrict our analysis to the best-sampled sites, in which there are twice as many or more trials as observed patterns, our results remain qualitatively intact.

We were also careful to check that our direct sampling method, with LE to correct for bias, agreed with the many other, sometimes more complex and sophisticated ways to quantify information in spike trains. We compared our results with that computed using the "best upper bound" or BUB method (35, 37), the unbiased entropy estimator proposed by Figure 2. Selection of sites for inclusion in our analysis. The number. M. of trials recorded for each site in each condition versus the number, L, of observed patterns of spiking and silence. This number, L, is always higher during undirected than during directed singing, reflecting the increased pattern entropy (see also Fig. 5B). Sites with fewer trials than observed patterns (gray area) were excluded from our analysis. Note that although L patterns is observed, the probability of these patterns is not uniform, so that we still sample quite well the more common patterns. The hard threshold on the number of trials is related to 2^{s} , where S is the pattern entropy. Our pattern entropy is always less than 2^{N} , so that our sampling is relatively stable for N < 7. The dashed line in A indicates the threshold for having twice as many trials as observed patterns. If we restrict our analysis to these bestsampled sites, our results remain qualitatively intact. As N bins increases from 5 shown in A to 6(B) the number of sites we use decreases from 24/28 to 15/28 (14 sites have data in both behavioral contexts).

Nemenman and colleagues or NSB method (37, 38), and the most recent estimator proposed by Pillow and colleagues (39, 40), which adds a prior on the expected sparseness of firing in the brain, the "centered Dirichlet mixture" or CDM estimator (a comparison with the CDM method is shown in Fig. 3). On average, our information values were within 8% of that quantified using any of these methods (CDM showed the largest differences), and none of our qualitative conclusions about the changes in the spike code with social context are altered by the information estimates are systematically lower than our LE estimates.



Figure 3. Comparison of our information estimation method with the centered Dirichlet mixture (CDM) method shows similar results. The information quantities estimated using the linear extrapolation method are plotted versus the centered Dirichlet mixture estimation method. Sites recorded during directed singing are in red and during undirected singing are in blue. On average, these values are within about 8% of each other across all sites included in our analysis. For those sites with higher linear extrapolation (LE) information, the CDM method's prior on the sparsity of the spiking seems to be failing. Other entropy estimation methods were closer to the LE method for these sites (data not shown).

Poisson Model

We construct a Poisson model constrained to have the same time-varying rate as a recorded cell. We sample spike trains from the model by drawing random numbers and comparing them with the firing rate in each time bin. Firing in each bin is independent of previous bins, and we can construct as many trials from this synthetic cell as we wish. We either draw 1,000 trials or the same number of trials as were collected in the real data set. Information quantities are computed for the Poisson spike trains as for the data for each neuron in our data set.

Statistical Tests

The significance of the observed differences in information measures across the two conditions was assessed using a Wilcoxon rank sum test for the total population data and a paired sign test for the sites in which both directed and undirected singing data were collected. Significance values were denoted with asterisks: *P < 0.01, **P < 0.001, **P < 0.001,

Synergy

The difference between the total pattern information and the average information carried by individual spikes and silences in the pattern was calculated as in *Eq.* 6 in RESULTS, as proposed by Bialek and coworkers (31). We also compute what fraction of the total pattern information is due to synergy, by summing over synergy values for each pattern, weighted by that pattern's probability.

RESULTS

To assess the nature of the neural code during motor practice and performance in the output nucleus of the anterior forebrain pathway, we analyzed data from single LMAN neurons in adult male zebra finches during interleaved bouts of undirected singing and courtship ("directed") singing. Song is an extended motor sequence that maintains a stereotyped structure from rendition to rendition in adult male birds. Temporal position in the motor sequence is a learned song feature that is represented in many brain areas (41, 42), and the timing of syllables and gaps is actively controlled and maintained (1, 43-50). To quantify the correlation between spiking sequences and song, we compute the mutual information between spiking patterns and time in song. We begin by quantifying the mutual information between time in song and the arrival time of single spikes both during directed singing to a female and during undirected singing, a form of song practice. Intuitively, for single spikes, this measures the amount of rate modulation during song repeats as quantified by the time-varying firing rate of the neuron (see Fig. 1). Since song is consistent from rendition to rendition in adult zebra finches, time in song can be thought of as a proxy for the particular motor gesture expressed at that time (51-54).

Single Spike Information Is Higher during Directed than Undirected Singing

As reported previously, inspection of the spike trains for single LMAN neurons shows striking differences in the reliability and precision of spikes across repeated renditions of

song depending on behavioral context, even though the song is very similar, acoustically, across conditions (12, 13). For example, in Fig. 1B, left, when a male sings songs directed to a female, spikes align across song renditions, and the average firing rate for this cell (red trace) is strongly peaked across repeated renditions of a stereotyped sequence of song elements ("motif"). The targeting of spikes to particular parts of song may reflect a refinement of the brain's firing during the task, akin to behavior-locked sharpening of response duration during song learning (55-57). When we compute the temporal jitter in the arrival times of the spikes, we find that during directed singing, spikes from this site are timed with a resolution of \sim 2.6 ± 1.6 ms, as calculated by the standard deviation in timing of spikes during peaks in firing. In contrast, for the same neuron, the spikes recorded during undirected song are more variable across renditions and have a temporal jitter of 3.9 ± 2.2 ms (Fig. 1B, right). Indeed, across the population of LMAN neurons, single spikes have an average jitter of 2.8±0.6 ms during directed singing versus 4.5 ± 3.1 ms during undirected singing, indicating that the timing of individual undirected spikes is less precisely locked to a particular time in song. On average, however, the firing pattern during undirected singing is similar to that during directed singing (Fig. 1B, blue trace, with directed pattern overlaid in light red for comparison).

To further quantify these observations, we use information theory to compute a single number that measures the correlation between the arrival time of single spikes and time in song. This reveals where in song, spiking output from LMAN is most keenly focused. Information theory has been extensively used to study coding in sensory systems, and the mutual information between neural activity and the stimulus is usually computed using data collected across repeated presentations of the sensory stimulus (31). Here, we are using the same formalism to compute the information about time during a repeated motor sequence. In this case, the songs vary slightly from rendition to rendition. Some of this variation is in acoustic features of the song, and some is in the duration of syllables and pauses between syllables. To remove the second source of variation, songs and spike trains were time-warped so that all of the motifs and spike trains are on a common and aligned time axis (see METHODS). In this way, we measure how LMAN spiking correlates with song motor output. To do so, we compute the average log signal-to-noise ratio of the firing rate,

$$I = \frac{1}{T} \int_0^T dt \left(\frac{r(t)}{\bar{r}} \right) \log_2\left(\frac{r(t)}{\bar{r}} \right), \tag{3}$$

where r(t) is the trial-averaged firing rate as a function of time during song for each recorded unit, derived from binning spikes across trials and computing their peri-song time histogram (PSTH). The procedure for computing this information from the PSTH is illustrated in Fig. 1C. Observing spikes from a neuron with a sharply peaked PSTH would indicate with high fidelity that those parts of song were currently being sung. By contrast, a spike from a neuron with a flat PSTH gives no information (no discriminability) about time in song. Thus, the mutual information that we compute quantifies how reliably peaked/inhomogeneous the firing of the neuron is during song. Observing a spike from a neuron with a high information content means that one can infer

time in song with high reliability. For the bird, this could mean that these particular parts of song are being targeted for modification or reinforcement.

We computed information about time in song for all sites, with a time resolution of 2 ms, and find that there is significantly more information in single spikes during directed song than during undirected song at each site, as shown in Fig. 1D, consistent with the greater temporal precision of single spikes in the directed context. These results indicate that in an individual trial, single spikes convey more information about time in song in the directed condition (more bits per spike) than in the undirected condition. Higher firing rates during undirected singing (more spikes per second) do not compensate for the lower information content per spike. When context-dependent differences in firing rate are taken into account, the single-spike information is still significantly lower during undirected song (Fig. 1D shows information in bits/s, Fig. 4A in bits/spike for comparison, and both measures are significantly lower in undirected singing), consistent with less temporal precision in single spiking during song practice. We also computed the single-spike information for many different temporal resolutions of the PSTH and see that both directed and undirected spikes show a steady climb in the information about time in song as resolution is increased, until jitter takes over below $\sim 1 \text{ ms}$ (Fig. 4B). This indicates that the precision of spike timing matters, down to $\sim 1 \, \text{ms}$ resolution.

Pattern Information Is Similar during Undirected and Directed Singing

The single-spike information measure assumes that all successive spikes are independent events, and thus explicitly neglects any information that might be present in temporally extended firing patterns. To analyze information from sequences of spikes, we extract short temporal strings (see METHODS) of activity from the binned spike data. We focus on five bins at 2 ms resolution for the majority of our analyses. The number of spikes in the window (count) can range from 0

to 5, with 5 representing a spike in every 2-ms time bin (Fig. 5A, top). The higher total spike count in the undirected condition (Fig. 5A, blue) reflects both the observed bursts present in these spike trains and the higher firing rate in undirected singing. The probability of observing a spiking pattern with more than two spikes in the full 10-ms window is much smaller during directed song than during undirected song. As the cartoon in Fig. 5B illustrates, individual spikes are denoted by a "1" and silences by a "0" in 2-ms bins. We then obtain the complete distribution of five-spike/no-spike patterns observed in these 10-ms windows, taken from all parts of song. The entropy of patterns for directed and undirected spiking is shown in Fig. 5C, top, as a function of the window length, from 2 to 12 ms. The pattern entropy (the "vocabulary" of LMAN spiking) is significantly greater in the undirected condition for all of the bin numbers measured.

The mutual information contained in temporal patterns about time in song computed using the direct method (33) is defined by the overall entropy of patterns minus the variability in spiking output at a particular time during song (the "noise" entropy):

$$I(\text{pattern}; t) = \sum_{i \text{ pattern}} P_i \bullet X_i \tag{4}$$

To compute the entropy of patterns at a particular time (the second term in the sum), we need to have enough samples at that time to estimate the full 2^N parameters. Care is taken to ensure that sampling bias does not affect this measurement. For five-bin temporal patterns, we can estimate pattern information for a total of 24/28 recorded sites (see METHODS). As we increase the number of bins in our temporal strings, we are restricted to sites with higher numbers of repeated song renditions, and the exact minimum number of song repetitions needed also depends on the entropy of spiking in that site (Fig. 2, e.g., for n = 6, we include 15/28 recorded sites in our analysis). In addition, a particular pattern might be very informative but also very rare. To



Figure 4. Single-spike information about time in song is lower in undirected activity across different temporal resolutions. *A*: mutual information between spikes and time in song for directed (DIR, red) and undirected (UNDIR, blue) conditions at 2 ms temporal resolution. Symbols are as in Fig. 1. *B*: information per spike increases with temporal resolution in both behavioral contexts, down to about 1 ms, where information diverges due to the sparsity of spiking aligned across trials. ***P* < 0.001, ****P* < 0.0001.



Figure 5. Information from sequences of spikes and silence become more comparable during directed and undirected song at longer time windows. *A*: an illustration (above) of counting spikes in a 10-ms window and (below) the distribution of count probabilities across all sampled time windows in song, across recording sites. Higher count events, which are mostly bursts, are not as frequent during directed singing. Error bars indicated standard deviation across all neurons. *B*: an illustration of how spiking symbols for temporal pattern calculations are defined (*top*). Spike trains are binned at 2 ms resolution. Patterns are defined in a 10-ms window. Full 5-bit patterns of spiking (1) and silence (0) are retained for computing pattern "vocabulary" as measured by the entropy for T from 1 to 6 bins. Spike pattern entropy is always significantly higher during undirected singing. (below) Mutual information between spike patterns and time in song as a function of *N* bins, for the set of neurons included in our analysis for *N* = 6. The information is expressed in bits/s so that rare patterns do not contribute significantly to the information rate. Pattern information is not significantly different between behavioral contexts for *N* = 5 or 6. ***P* < 0.001; ****P* < 0.0001; ns, not significant.

mitigate the relative contribution of highly informative but infrequent patterns, the mutual information is expressed in units of bits per second so that rare patterns do not make significant contributions to the information rate. When temporal patterns of spiking are analyzed, information about time in song increases in both contexts (Fig. 5C, bottom). The relative increase in information rates obtained for patterns versus single spikes is plotted in Fig. 6, as a function of the length of the pattern. The temporal resolution within the window does not affect count information, as expected, but greatly influences pattern information until 2ms resolution is reached, and this effect is more pronounced in undirected spiking (Fig. 7). At 2 ms resolution and for N bins = 5 or 6, the amount of pattern information in the two contexts is not significantly different (Fig. 5C, bottom, and Fig. 6), indicating that undirected spiking contains just as much information about time in song as directed song when temporal strings of at least 10 ms in length are considered in the neural code. We also see that spike patterns during undirected singing show a greater gain in information from temporal patterns over single spikes.

Pattern Information Is Conveyed by Different Spiking Sequences in the Two Contexts

Given that the total information during undirected singing increases when we take temporal patterns into account. there must be some patterns that exhibit greater precision in their timing than that of the average undirected spike (Fig. 8). These specific patterns are obscured in the full spike raster, where all patterns that occur are plotted together. Figure 8, B and C, shows firing patterns from two different LMAN neurons that are much more tightly locked to song than the average spike (top). Sharp peaks in the pattern probability (indicated by asterisks) occur at points in song where the single-spike PSTH (top) was only weakly modulated above the mean. We find that there are many such significantly peaked patterns in each recorded neuron during undirected singing. On average, 36% of the observed burst patterns in each neuron (n = 19) had at least one significant peak in its PSTH. In contrast, during directed singing, different temporal patterns tend to peak at the same time(s) in song (Fig. 8D and Fig. 9A), though one neuron out of 24 had patterns that



Figure 6. Undirected spiking displays a greater gain in information from temporal patterns over single spikes. Pattern information compared with single-spike information averaged across all sites during directed and undirected singing. Pattern information shows a significant gain over single spike coding for time in song during practice (A), but there is no significant gain in pattern information during performance (B). *P < 0.01; n.s., not significant.

peaked at different times in song during directed singing (Fig. 9B).

To characterize these informative patterns, we next assess whether the total pattern information is distributed uniformly across all possible patterns or whether it is clustered around particular patterns. The full pattern information (Eq. 4) can be decomposed into a sum over all 2^N possible strings of spikes and silence,

$$I(\text{patterns}; t) = \sum_{i \text{ patterns}} P_i \bullet X_i \tag{5}$$

where, P_i is overall probability of that pattern, and X_i is its contribution to the information computed using its PSTH using an expression analogous to Eq. 3 (see METHODS for derivation). Figure 8A shows a typical example of the contribution of pattern information to the total information. We group the observed patterns according to how many total spikes are observed in the 10-ms window (count). For example, we can compute how much information each of the twospike patterns carry about song by collecting the terms in the sum in Eq. 5 that correspond to the 10 unique two-spike patterns. There is just one zero-spike word, the pattern with all zeros, five one-spike words in which the single spike

occurs in one of each of the five bins, (two-spike words.

and so on. Grouping patterns in this way, we plot the total pattern information across the entire population of neurons decomposed according to K-spikes in the pattern in Fig. 8, E and F, in both directed and undirected song. Single-spike patterns account for most of the information observed during directed song $(70\% \pm 4\%)$. These results are shown for the average over all sites, but each individual cell also displayed the same trend. The greater precision of single-spike timing during directed song gives rise to larger information in these patterns than in undirected song. In contrast, pattern information during undirected song is dominated by two- and three-spike strings, which are primarily bursts (59% ± 5% of the total information is from two- and three-spike patterns, all of which are classified as bursts when defined by an

instantaneous firing rate of >200 spikes/s). Figure 8F shows that during undirected song, bursts are the primary contributors to information about time in song. This implies that the coding scheme used by an individual neuron during these two behavioral contexts is quite different: depending on context, the same neuron can utilize a different set of temporal patterns (Fig. 5, A and B), and these patterns contain different amounts of information (Fig. 8, A and B).

Coding Changes with Behavioral Context

To investigate directly whether the same neuron can use different coding schemes depending on social context, we eliminate timing information within the window so that we can compare the full pattern information to count information (Fig. 10). We have already shown that the complement of spiking patterns available to a neuron changes with behavioral context (Fig. 5, A and B), but we would like to quantify how much of that change in "vocabulary" is used in coding for time in song. To determine whether temporal patterns carry more information than spike count alone, we compare pattern information with count information, grouped by count within the pattern (Fig. 10, A and B.) In the directed case, a large fraction of the pattern information can be summarized by count; simply noting whether zero, one, or two spikes occurred in a 10-ms window yields more than half (64%) of the total information about time in song. The greatest difference between full pattern information and count is explained by the timing precision of the first spike in a window, for patterns containing a single spike (Fig. 10, E and *F*). In contrast, count information is consistently low in the undirected condition, even for high-count events, and accounts for a smaller fraction (39%) of the complete pattern information (Fig. 10, B and F). Moreover, higher firing rates



Figure 7. Undirected spiking shows a sharp increase in pattern information as the time resolution for resolving spikes within the window increases. Mutual information between spike patterns (solid lines) or counts (dashed lines) and time in song is plotted as a function of the temporal resolution within the window. Count is relatively unaffected by increased temporal resolution, with small gains arising from the resolving of nearby spikes into counts of 2 instead of 1. Pattern information increases around 2 ms resolution within the window, as might be expected from the measurements of temporal jitter in these spike trains.



Figure 8. Bursts during practice carry information about time in song by pointing to different parts of song, depending upon their spike pattern. A: contribution to the total pattern information from each 5-bin pattern for a particular neuron in our data set (b39b14_12). Information for undirected patterns (blue) is more spread out across patterns, whereas mostly 1-spike patterns code for time in song during directed singing (red). B-D: smoothed, average firing rates for two different lateral magnocellular nucleus of the anterior nidopallium (LMAN) neurons for single spikes (top) during undirected trials, or particular temporal patterns (middle and bottom) with two or three spikes in a 10-ms window. The black line indicates the mean firing rate, and the gray area is ±1 SD from the mean. Peri-song time histograms (PSTHs) for the same neuron during undirected and directed singing are plotted in C and D. Significant peaks in the pattern PSTHs are indicated with asterisks. The data in A are from the same neuron as in C and D. E: pattern information grouped by count for directed activity. The majority of pattern information comes from single-spike patterns. F: same as E, but for undirected activity. During undirected singing, pattern information arises predominantly from 1-, 2-, and 3-spike patterns. *P < 0.01.

during undirected singing do not yield more information in spike counts (Fig. 10*C*, *x*-axis).

To the extent to which pattern information can be summarized by simply counting spikes, the encoding is rate-like with a time resolution equal to the length of the counting window (here, 10 ms). Figure 10*C* plots the pattern information versus count information for each neuron and shows that the data roughly cluster into two groups according to social context. Information during directed singing (red circles) trends along but slightly above the rate-coding line



Figure 9. Directed patterns often code for the same part of song, redundantly. *A*: spike or pattern rate as a function of time in song. The gray area indicates the mean rate ± 1 SD. Pattern rates peak at the same parts of song. *B*: spike and pattern rates for a different site. Patterns 10100 and 10010 point to somewhat nonoverlapping parts of song in this neuron. **P* < 0.01.

(count info. = pattern info.), whereas pattern information in undirected singing (blue circles) outstrips the very small count information. This change is also evident in the ratio of pattern to count information in the two behavioral contexts (Fig. 10*D*). The change in the relative amount of count and pattern information with behavioral context indicates that single neurons switch between more rate-like coding during directed song and more fully temporal coding during undirected song.

We also consider a code in which the total count is combined with the arrival time of the first spike in the window, as illustrated in Fig. 10E, top. Total information from these three different codes is compared in the two social contexts. The gray bars in Fig. 10F, show the fraction of total pattern information that can be summarized by counting the number of spikes in the 10-ms window. Count information is a substantially larger fraction of the total pattern information during directed singing. If we keep track of both count and the timing of the first spike in the window (a measure of temporal precision in the initiation of firing in the window), most of the information in directed spiking is accounted for. However, nearly one-third of the total pattern information during undirected singing is not recovered (Fig. 10F, white region). This remaining information is the pure temporal sequence part of the code. This means that the detailed timing of spikes within bursts contributes substantially to the code for time in song during undirected singing.

To summarize, breaking down the pattern information in bursts reveals that, during undirected singing, bursts have information contained in their count, the timing of the first spike, and their within-burst pattern. In contrast, during directed singing, information about time in song could described largely by a count code plus alignment of single spikes to particular times in song, without much temporal pattern information.

A Poisson Model Does Not Explain Temporal Coding during Undirected Singing

What elements of the firing in LMAN give rise to the temporal code we observe during motor practice? Can the timevarying firing rate account for all of the observed information or do temporal correlations add significantly to the code? To address these questions, we compute the mutual information between time in song and spikes generated by a time-varying Poisson process with a rate matched to the real neurons, with varying temporal resolution (see METHODS). A Poisson model with a rate varying at the timescale of our temporal pattern window (10 ms) does not contain any information beyond the spike count in the window (Fig. 11) and serves as a check on our procedure for generating synthetic trials and computing information. We can also match the rate in the model to the real data at faster timescales (2 ms resolution) and measure the information in temporal patterns from these model responses. We find that although some



Figure 10. Temporal coding is more prevalent in undirected song, whereas coding is more rate-like in directed song. *A* and *B*: different spike patterns carry information in the two singing conditions. Information in patterns (white bars) versus information in counts (colored bars). During directed singing (*A*), most of the pattern information is contained in single-spike events. About half of this information is captured by counts. In contrast, most of the information during undirected singing is carried by higher spike count patterns, and very little information is captured by counts alone (*B*). Error bars indicate SE across recording sites. C: information from patterns plotted versus information from count for each recording site in lateral magnocellular nucleus of the anterior nidopallium (LMAN), in both behavioral conditions. The dashed line indicates perfect rate coding, where pattern information to count info. In undirected song, spike counts carry very little information, but points follow the rate coding line. Error bars indicate ±1 SD. *D*: the ratio of pattern information to count information is slightly larger than count information in undirected singing for *n* = 5 and 6. *E*: an illustration of spike pattern symbols that label both the count in the window, and the arrival bin of the first spike. *F*: summary of the contributions to total pattern information from count alone (gray portion) and from count + first-spike-in-the-bin timing (gray + hashed portion) reveals a substantial remaining amount of pure temporal sequence information (white portion) during undirected song. **P* < 0.001; ***P* < 0.001.

information about time in song is coded in this model, it fails to reproduce the qualitative changes in coding with behavioral context and does not recapitulate the observed relative ratio of count to pattern information in these data (Fig. 12). Overall, the model fails to capture the details of the temporal code for time in song, particularly during undirected song, implying that temporal correlations between successive spikes in these windows contribute significantly to temporal pattern information.

Spikes within Bursts Code for Time in Song Synergistically, More So during Undirected Singing

How can pattern information arise in undirected singing from relatively uninformative single spikes? The answer clearly has to do with the fact that subsequent spikes and silences refine an estimate of time in song, adding up independently at least. In addition, there is the potential for synergy in temporal sequences of activity. Figure 8*C* illustrates how subsequent spikes in a burst carry additional information about time in song, depending on exactly when in the window they occur. During undirected song, this typical LMAN neuron has several two-spike patterns in its vocabulary, whose probability of occurrence peaks during different parts of song. The particular time in song coded by the pattern depends on the precise temporal pattern of spiking and silence. For example, the pattern "01001" peaks at ~150 ms in the motif, whereas "10001" peaks at a completely different time in song (~450 ms). These results indicate that the larger repertoire of temporal patterns used to convey information during undirected singing is neither an artifact of a higher firing rate nor a switch to random bursting but rather is used to encode different times in song (see also Fig. 8*B*).

To dissect this more carefully, we quantify how spikes and silences combine to create the observed encoding. We compute the information from the full string of spikes and silences in an observed pattern, minus the sum of the contributions from each spike or silence individually,

$$I_{synergy} = I(\{\sigma_i\}; t) - \sum_i I(\sigma_i; t),$$
(6)

Where σ is a binary variable for a spike ($\sigma = 1$) or silence ($\sigma = 0$) and { σ_i } is the 5-bit string of spikes and silences in the temporal pattern. If this quantity is positive, the pattern synergistically codes for time in song, meaning that spikes in the pattern add information supralinearly. If negative, the spikes and silences add redundantly, meaning that additional spikes in the burst code for the same part of song. If spikes



Figure 11. Count and pattern information in a 10-ms window are equivalent in a Poisson model with 10 ms resolution. The ratio of pattern to count information is plotted versus count information for Poisson spike trains with a time-varying rate smoothed to 10 ms resolution. As expected, count and patterns are equivalent for both directed and undirected models.

and silences independently code for time in song, their synergy is zero. We compute the synergy for each pattern, then group patterns by count as in Figs. 8 and 10, and plot this value as a function of the number (K) of spikes in the pattern (Fig. 13, A and B). Most strikingly, we see significant synergy in patterns with two or more spikes. In contrast, during directed singing, most of the information from patterns derive from combinations of single spikes and silence, long silences (00000 pattern), and two-spike patterns. This relatively nonsynergistic coding in spiking during directed song

is shown in Fig. 14. Across all sites, synergy contributes more to pattern information during undirected song. This clearly indicates that burst firing carries information about song in an intrinsically temporal fashion. Indeed, if LMAN bursts were generated by a cellular mechanism intrinsic to the cell, we would expect subsequent spikes in a burst to form a redundant code for time in song, since their pattern was completely determined by burst initiation. LMAN neurons, however, are not intrinsically bursting cells (21, 58), and more recent data suggest that burst firing depends on signals from the basal ganglia nucleus Area X (59, 60). Thus, the observed bursts are the product of input activity and, to a lesser extent, recurrent circuit dynamics and may contain significant tunable temporal correlation structure. Put another way, LMAN bursts are not unitary events that can be combined into a single category or even subdivided by the number of spikes within the burst. Rather, the particular pattern of spikes within a burst conveys information about time in song and does so synergistically.

Overall, our results support the idea that LMAN activity locally modulates acoustic features of song. Bursts carry the majority of information about time in undirected song via a truly temporal pattern code. We observe that behavioral context elicits a state change in the neural code from a primarily single spike-timing code during directed singing to a temporal sequence code during undirected song.

DISCUSSION

Our analysis reveals that the precise timing of spikes and silences of single LMAN neurons carries information about time in song in both behavioral contexts but that the format of the neural code is fundamentally different. Although it has been hypothesized that the variable firing patterns in LMAN inject noise into the song motor system to facilitate



Figure 12. Time-varying rate alone does not account for the observed temporal information in lateral magnocellular nucleus of the anterior nidopallium LMAN spiking. A: the ratio of pattern to count mutual information plotted versus count information for spiking during both social contexts (blue: undirected, red: directed). Undirected spiking shows a substantial gain in pattern information over count. B: a Poisson model with the same time-varying fire rate (at 2 ms resolution) as the real neurons shows a much weaker gain in pattern information for undirected spiking, and a substantially larger gain during directed singing as compared with the real data. The inset shows results of a Poisson model where the number of trials was matched to that recorded in the data for each site and behavioral context. C: Poisson model mutual information versus data mutual information for all neurons and all behavioral contexts, and for both count (x's) and pattern (o's) information. The Poisson model has consistently more information, particularly in temporal patterns of spiking during directed singing. Real neurons have a higher variability at some points in time during song than expected in a Poisson model, and temporal correlations in spiking events lead to higher variability in spike patterns than in the model. These effects combine to yield the higher information rates in the model, even when the mean firing rates are matched.

Figure 13. Bursts code for time synergistically and overall synergy is higher during undirected singing. A: synergy (dotted bars) is plotted for combinations of spikes and silence versus spike count for undirected singing as in Eq. 4. Pattern information (dotted bars) is shown for comparison in the background (colored bars). During undirected singing, we observe more synergy in higher-count patterns, corresponding to bursts. B: averaging over all patterns, we plot the percent of pattern information that comes from synergy for undirected (blue) and directed (red) trials. The percent synergy increases with pattern length, and more substantially for undirected versus directed spiking. *P < 0.01; ***P* < 0.001.



exploration for trial-and-error reinforcement learning (4, 27, 28), we find that LMAN firing patterns come in many forms, each of which can reliably indicate particular times in song depending on their internal spiking structure. Our findings are consistent with the observation that microstimulation in LMAN can have systematic, local effects on the acoustic structure of song (14). When LMAN firing patterns are lumped together, however, they appear to be noisy, making it difficult to "read off" the code from a spike raster. Moreover, the diversity of patterns used to encode information about song changes with behavioral context: information is carried predominantly by precisely timed bursts during undirected singing, a putative practice state, versus the precise timing of single spikes during performance directed at a female.

Changes in the firing mode in individual cells and entire circuits can shape how signals are transformed within and received downstream of that brain area. Bursts, in particular, are thought to represent an important mode of neuronal signaling. In comparison with single spikes, bursts have been shown to transmit stimulus information in a way that is distinct from single spikes both in the neocortex and in the

midbrain of weakly electric fish (22, 61–64) and to enhance the reliability of information transfer in the hippocampus and in the lateral geniculate nucleus of the thalamus (23, 65, 66). In songbirds, a variety of evidence indicates that burst firing in LMAN may be particularly important for driving changes in song. First, manipulations of the AFP circuit that specifically eliminate LMAN bursts (but not single-spike firing) also eliminate the bird's ability to change song in response to altered auditory feedback (21). Indeed, it is harder to perturb song during directed singing than during undirected singing (67), perhaps because isolated spikes in LMAN do not facilitate downstream firing or plasticity as well as bursting does. The suppression of burst firing and the increased reliability and precision of spike timing in LMAN that is elicited by the presence of a female is reminiscent of the observed decline in neural variation with stimulus onset in many areas of the mammalian cortex (68).

The diversity of burst patterns deployed during undirected song may help guide downstream motor behavior toward a targeted, moment-by-moment exploration of song space through multiple mechanisms. First, particular burst patterns could specify the specific parts of song to modify as

Figure 14. Directed patterns synergy is low and accounts for less than 20% of the total pattern information across all sites. *A*: synergy (dotted bars) as a function of the number of spikes in the window, for directed spiking, as in Fig. 13. Most of the information from patterns comes from slightly redundant combinations of single spikes and silence for 1-spike patterns. Remaining patterns contain some synergy. The 00000 pattern is synergistic. Long silences code for pauses between precisely timed spikes better than uninformative short (2-ms) silences, which are scattered all throughout song. *B*: percent synergy for *N*=5 across all sites and behavioral context in our analysis. Synergy contributes a larger fraction of pattern information during practice. **P* < 0.01; n.s., not significant.



well as how to change their acoustic features by altering the firing rates of neurons in the song motor cortex analog RA [robust nucleus of the arcopallium, Fig. 1A (57)], which have been shown to correlate with the pitch, amplitude, and local entropy of individual learned song elements (20, 29, 69, 70). Indeed, altering the relative timing of inputs from LMAN and the premotor nucleus HVC is sufficient to drive jitter in the timing of RA spikes (29, 71). Perhaps even more compellingly, firing and plasticity in RA depend on burst firing in LMAN as well as the fine timing between HVC and LMAN inputs to RA (57, 71, 72).

The importance of temporal pattern coding for song has also been recently quantified in RA (30). Indeed, acute alterations of LMAN firing via microstimulation can drive systematic changes in acoustic features of song (14). Second, the specific temporal patterns within bursts could have different downstream effects via nonlinear dendritic (73) or network (74) interactions. Third, burst firing in LMAN could facilitate plasticity via NMDA receptor activation (75, 76) and increased calcium influx in RA neurons. LMAN might also impact the sequence of syllables sung more indirectly via projections onto HVC (77).

Bursting underlies the coding switch with behavioral state in this system, which is the analog to the outflow of mammalian basal ganglia circuitry. Although bursting is often thought of as a pathological signal in basal ganglia circuits in disease states, such as the increased bursting observed in Parkinson's disease (78), with therapies such as deep brain stimulation deployed to suppress it (79), here we have shown that it might play a role in shaping normal task-specific behavior, such as novelty-driven exploration (80). In addition, context-triggered switches in coding like the one seen here may be a general feature of basal ganglia circuits, which are well known to show sensitivity to contextual cues such as reward (81, 82).

More broadly, this work speaks to how reinforcement learning may be instantiated in any motor system. In contrast to models of reinforcement learning that use random noise to generate behavioral variability to thoroughly explore task space to achieve optimal behavior, avoiding the pitfalls of local minima (83), we suggest that the nervous system may actively sample motor space in a targeted manner. We find that the precise timing of spikes in temporally extended patterns can reliably signal particular times in a task and potentially direct motor exploration toward the target. Our results support theories of sensorimotor learning that permit active sampling of motor space. Such "active learning" models are more efficient, displaying an exponential improvement in the number of samples needed to reach target generalization error as compared with random or batchlearning models, and are gaining popularity in the machine learning community (84-86). Our results support their general applicability to neural systems (87).

ACKNOWLEDGMENTS

We thank Gašper Tkačik and David Schwab for useful conversations and Philip Sabes and members of the Doupe laboratory for feedback on the manuscript.

GRANTS

This work was supported by NIH Grants MH55987 and MH78824 (to A.J.D.), by the Swartz Foundation, and in part by

National Science Foundation CAREER Award 1066293 (to S.E.P.), Tufts University startup funds to M.H.K., and this work was performed in part at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

S.E.P. B.D.W., A.J.D., and M.H.K. conceived and designed research; M.H.K. performed experiments; S.E.P. and M.H.K. analyzed data; S.E.P. and M.H.K. interpreted results of experiments; S.E.P. and M.H.K. prepared figures; S.E.P. and M.H.K. drafted manuscript; S.E.P. and M.H.K. edited and revised manuscript; S.E.P. and M.H.K. approved final version of manuscript.

REFERENCES

- Bottjer SW, Misener EA, Arnold AP. Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* 224: 901–903, 1984. doi:10.1126/science.6719123.
- Hampton CM, Sakata JT, Brainard MS. An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. J Neurophysiol 101: 3235– 3245, 2009. doi:10.1152/jn.91089.2008.
- Kao MH, Brainard MS. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. J Neurophysiol 96: 1441–1455, 2006. doi:10.1152/jn.01138.2005.
- Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3: e153– 0909, 2005. doi:10.1371/journal.pbio.0030153.
- 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. J Neurosci 11: 2896–2913, 1991. doi:10.1523/jneurosci.11-09-02896.1991.
- Stepanek L, Doupe AJ. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. J Neurophysiol 104: 2474–2486, 2010. doi:10.1152/jn.00977.2009.
- Andalman AS, Fee MS. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci U* S A 106: 12518–12523, 2009. doi:10.1073/pnas.0903214106.
- Charlesworth JD, Tumer EC, Warren TL, Brainard MS. Learning the microstructure of successful behavior. *Nat Neurosci* 14: 373–380, 2011. doi:10.1038/nn.2748.
- Tumer EC, Brainard MS. Performance variability enables adaptive plasticity of "crystallized" adult birdsong. *Nature* 450: 1240–1244, 2007. doi:10.1038/nature06390.
- Warren TL, Tumer EC, Charlesworth JD, Brainard MS. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *J Neurophysiol* 106: 1806–1821, 2011. doi:10.1152/jn.00311. 2011.
- Goldberg JH, Adler A, Bergman H, Fee MS. Singing-related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: comparison to the primate internal and external pallidal segments. *J Neurosci* 30: 7088–7098, 2010. doi:10.1523/JNEUROSCI.0168-10.2010.
- Hessler NA, Doupe AJ. Social context modulates singing-related neural activity in the songbird forebrain. *Nat Neurosci* 2: 209–211, 1999. doi:10.1038/6306.
- Kao MH, Wright BD, Doupe AJ. Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. J Neurosci 28: 13232–13247, 2008. doi:10.1523/JNEUROSCI.2250-08.2008.
- Kao MH, Doupe AJ, Brainard MS. Contribution of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* 433: 638–643, 2005. doi:10.1038/nature03127.
- Charlesworth JD, Warren TL, Brainard MS. Covert skill learning in a cortical-basal ganglia circuit. *Nature* 486: 251–255, 2012. doi:10. 1038/nature11078.

- Kojima S, Doupe AJ. Social performance reveals unexpected vocal competency in young songbirds. *Proc Natl Acad Sci USA* 108: 1687– 1692, 2011. doi:10.1073/pnas.1010502108.
- Sossinka R, Böhner, J. Song types in the Zebra Finch *Poephila* guttata castanotis. Z Tierpsychol 53: 123–132, 1980. doi:10.1111/ j.1439-0310.1980.tb01044.x.
- Immelmann K. Song development in the zebra finch and other estrildid finches. In: *Bird Vocalisations: Their Relations to Current Problems in Biology and Psychology*, edited by Hinde RA. Cambridge, UK: Cambridge University Press, 1969, p. 64–74.
- Marler P. Song learning: the interface between behaviour and neuroethology. *Philos Trans R Soc Lond B Biol Sci* 329: 109–114, 1990. doi:10.1098/rstb.1990.0155.
- Tchernichovski O, Mitra PP, Lints T, Nottebohm F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291: 2564–2569, 2001. doi:10.1126/science.1058522.
- Kojima S, Kao MH, Doupe AJ. Task-related "cortical" bursting depends critically on basal ganglia input and is linked to vocal plasticity. *Proc Natl Acad Sci U S A* 110: 4756–4761, 2013. doi:10.1073/ pnas.1216308110.
- Krahe R, Gabbiani F. Burst firing in sensory systems. Nat Rev Neurosci 5:13–23, 2004. doi:10.1038/nrn1296.
- Lisman JE. Bursts as a unit of neural information: making unreliable synapses reliable. *Trends Neurosci* 20: 38–43, 1997. doi:10.1016/ S0166-2236(96)10070-9.
- Brainard MS, Doupe AJ. Translating birdsong: songbirds as a model for basic and applied medical research. *Annu Rev Neurosci* 36: 489–517, 2013. doi:10.1146/annurev-neuro-060909-152826.
- 25. Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F. For whom the bird sings: context-dependent gene expression. *Neuron* 21: 775–788, 1998. doi:10.1016/S0896-6273(00)80594-2.
- Sakata JT, Hampton CM, Brainard MS. Social modulation of sequence and syllable variability in adult birdsong. J Neurophysiol 99: 1700–1711, 2008. doi:10.1152/jn.01296.2007.
- Fee MS, Goldberg JH. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* 198: 152–170, 2011. doi:10.1016/j.neuroscience.2011.09.069.
- Fiete IR, Fee MS, Seung HS. Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. J Neurophysiol 98: 2038–2057, 2007. doi:10.1152/jn.01311. 2006.
- Srivastava KH, Holmes CM, Vellema M, Pack AR, Elemans CPH, Nemenman I, Sober SJ. Motor control by precisely timed spike patterns. Proc Natl Acad Sci U S A 114: 1171–1176, 2017. doi:10.1073/pnas. 1611734114.
- Tang C, Chehayeb D, Srivastava K, Nemenman I, Sober SJ. Millisecond-scale motor encoding in a cortical vocal area. *PLoS Biol* 12: e1002018, 2014. doi:10.1371/journal.pbio.1002018.
- Brenner N, Strong SP, Koberle R, Bialek W, De Ruyter Van Steveninck RR. Synergy in a neural code. Neural Comput 31: 15732–15741, 2000. doi:10.1162/089976600300015259.
- Osborne LC, Palmer SE, Lisberger SG, Bialek W. The neural basis for combinatorial coding in a cortical population response. *J Neurosci* 28: 13522–13531, 2008. doi:10.1523/JNEUROSCI.4390-08.2008.
- Strong SP, Koberle R, De Ruyter Van Steveninck RR, Bialek W. Entropy and information in neural spike trains. *Phys Rev Lett* 80: 197, 1998. doi:10.1103/PhysRevLett.80.197.
- Ince RAA, Mazzoni A, Petersen RS, Panzeri S. Open source tools for the information theoretic analysis of neural data. *Front Neurosci* 4: 62, 2010. doi:10.3389/neuro.01.011.2010.
- Ince RAA, Schultz SR, Panzeri S. Estimating information-theoretic quantities. In: *Encyclopedia of Computational Neuroscience*, edited by Jaeger D, Jung R. New York: Springer, 2014. doi: 10.1007/978-1-4614-7320-6_140-1.
- Treves A, Panzeri S. The upward bias in measures of information derived from limited data samples. *Neural Comput* 7: 399–407, 1995. doi:10.1162/neco.1995.7.2.399.
- Panzeri S, Senatore R, Montemurro MA, Petersen RS. Correcting for the sampling bias problem in spike train information measures. J Neurophysiol 98: 1064–1072, 2007. doi:10.1152/jn.00559.2007.
- Nemenman I, Shafee F, Bialek W. Entropy and inference, revisited. In: NIPS'01: Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic,

edited by Dietterich TG, Becker S, Ghahramani Z. Cambridge, MA: MIT Press, 2001, p. 471–478.

- Archer E, Park IM, Pillow JW. Bayesian entropy estimation for binary spike train data using parametric prior knowledge. Advances in Neural Information Processing Systems 26: 1700–1708, 2013.
- Knudson K, Pillow JW. Spike train entropy-rate estimation using hierarchical Dirichlet process priors. Advances in Neural Information Processing Systems 26: 2076–2084, 2013.
- Bouchard KE, Brainard MS. Neural encoding and integration of learned probabilistic sequences in avian sensory-motor circuitry. J Neurosci 98: 17710–17723, 2013. doi:10.1523/jneurosci.2181-13.2013.
- Margoliash D, Fortune ES. Temporal and harmonic combinationsensitive neurons in the zebra finch's HVc. J Neurosci 12: 4309– 4326, 1992. doi:10.1523/jneurosci.12-11-04309.1992.
- Ashmore RC, Wild JM, Schmidt MF. Brainstem and forebrain contributions to the generation of learned motor behaviors for song. J Neurosci 25: 8543–8554, 2005. doi:10.1523/JNEUROSCI.1668-05. 2005.
- Brainard MS, Doupe AJ. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* 404: 762– 766, 2000. doi:10.1038/35008083.
- Chi Z, Margoliash D. Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron* 32: 899–910, 2001. doi:10.1016/S0896-6273(01)00524-4.
- Glaze CM, Troyer TW. Temporal structure in zebra finch song: implications for motor coding. *J Neurosci* 26: 991–1005, 2006. doi:10. 1523/JNEUROSCI.3387-05.2006.
- Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Feher O, Ravbar P, Okanoya K, Tchernichovski O. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* 498: 104–108, 2013. doi:10.1038/nature12173.
- Lombardino AJ, Nottebohm F. Age at deafening affects the stability of learned song in adult male zebra finches. *J Neurosci* 20: 5054– 5064, 2000. doi:10.1523/jneurosci.20-13-05054.2000.
- Long MA, Fee MS. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* 456: 189–194, 2008. doi:10. 1038/nature07448.
- Nordeen KW, Nordeen EJ. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav Neural Biol* 57: 58–66, 1992. doi:10.1016/0163-1047(92)90757-U.
- Amador A, Perl YS, Mindlin GB, Margoliash D. Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature* 495: 59–64, 2013. doi:10.1038/nature11967.
- Hahnloser RHR, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 419: 65–70, 2002. doi:10.1038/nature00974.
- Laje R, Sciamarella D, Zanella J, Mindlin GB. Bilateral source acoustic interaction in a syrinx model of an oscine bird. *Phys Rev E Stat Nonlinear Soft Matter Phys* 77: 011912, 2008. doi:10.1103/PhysRev E.77.011912.
- Yu AC, Margoliash D. Temporal hierarchical control of singing in birds. *Science* 273: 1871–1875, 1996. doi:10.1126/science.273.5283. 1871.
- Jog MS, Kubota Y, Connolly CI, Graybiel AM. Building neural representations of habits. *Science* 286: 1745–1749, 1999. doi:10.1126/ science.286.5445.1745.
- Okubo TS, Mackevicius EL, Payne HL, Lynch GF, Fee MS. Growth and splitting of neural sequences in songbird vocal development. *Nature* 528: 352–357, 2015. doi:10.1038/nature15741.
- Ölveczky BP, Otchy TM, Goldberg JH, Aronov D, Fee MS. Changes in the neural control of a complex motor sequence during learning. J Neurophysiol 106: 386–397, 2011. doi:10.1152/jn.00018.2011.
- Boettiger CA, Doupe AJ. Developmentally restricted synaptic plasticity in a songbird nucleus required for song learning. *Neuron* 31: 809–818, 2001. doi:10.1016/S0896-6273(01)00403-2.
- Tanaka M, Singh Alvarado J, Murugan M, Mooney R. Focal expression of mutant huntingtin in the songbird basal ganglia disrupts cortico-basal ganglia networks and vocal sequences. *Proc Natl Acad Sci* 113: E1720–E1727, 2016. doi:10.1073/pnas.1523754113.
- Woolley SC, Rajan R, Joshua M, Doupe AJ. Emergence of contextdependent variability across a basal ganglia network. *Neuron* 82: 208–223, 2014. doi:10.1016/j.neuron.2014.01.039.

- Higgs MH, Spain WJ. Conditional bursting enhances resonant firing in neocortical layer 2-3 pyramidal neurons. *J Neurosci* 29: 1285– 1299, 2009. doi:10.1523/JNEUROSCI.3728-08.2009.
- Khosravi-Hashemi N, Fortune ES, Chacron MJ. Coding movement direction by burst firing in electrosensory neurons. *J Neurophysiol* 106: 1954–1968, 2011. doi:10.1152/jn.00116.2011.
- Middleton JW, Yu N, Longtin A, Maler L. Routing the flow of sensory signals using plastic responses to bursts and isolated spikes: experiment and theory. *J Neurosci* 31: 2461–2473, 2011. doi:10.1523/ JNEUROSCI.4672-10.2011.
- Oswald AMM, Chacron MJ, Doiron B, Bastian J, Maler L. Parallel processing of sensory input by bursts and isolated spikes. J Neurosci 21: 4351–4362, 2004. doi:10.1523/JNEUROSCI.0459-04. 2004.
- Malinow R, Otmakhov N, Blum KI, Lisman J. Visualizing hippocampal synaptic function by optical detection of Ca²⁺ entry through the N-methyl-D-aspartate channel. *Proc Natl Acad Sci U S A* 91: 8170– 8174, 1994. doi:10.1073/pnas.91.17.8170.
- Sherman SM. Tonic and burst firing: dual modes of thalamocortical relay. *Trends Neurosci* 2: 122–126, 2001. doi:10.1016/s0166-2236(00) 01714-8.
- Sakata JT, Brainard MS. Real-time contributions of auditory feedback to avian vocal motor control. *J Neurosci* 26: 9619–9628, 2006. doi:10.1523/JNEUROSCI.2027-06.2006.
- Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, Finn IM, Ferster D, Ryu SI, Santhanam G, Sahani M, Shenoy KV. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13: 369–378, 2010. doi:10.1038/nn.2501.
- Goller F, Suthers RA. Role of syringeal muscles in controlling the phonology of bird song. J Neurophysiol 76: 287–300, 1996. doi: 10.1152/jn.1996.76.1.287.
- Sober SJ, Wohlgemuth MJ, Brainard MS. Central contributions to acoustic variation in birdsong. *J Neurosci* 28: 10370–10379, 2008. doi:10.1523/JNEUROSCI.2448-08.2008.
- Murugan M, Harward S, Scharff C, Mooney R. Diminished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. *Neuron* 80: 1464–1476, 2013. doi:10.1016/ j.neuron.2013.09.021.
- Mehaffey WH, Doupe AJ. Naturalistic stimulation drives opposing heterosynaptic plasticity at two inputs to songbird cortex. *Nat Neurosci* 18: 1272–1280, 2015. doi:10.1038/nn.4078.
- London M, Häusser M. Dendritic computation. Annu Rev Neurosci 28: 503–532, 2005. doi:10.1146/annurev.neuro.28.061604.135703.

- Gütig R, Sompolinsky H. The tempotron: a neuron that learns spike timing-based decisions. *Nat Neurosci* 9: 420–428, 2006. doi:10. 1038/nn1643.
- Mooney R, Konishi M. Two distinct inputs to an avian song nucleus activate different glutamate receptor subtypes on individual neurons. Proc Natl Acad Sci U S A 88: 4075–4079, 1991. doi:10.1073/ pnas.88.10.4075.
- Stark LL, Perkel DJ. Two-stage, input-specific synaptic maturation in a nucleus essential for vocal production in the zebra finch. J Neurosci 19: 9107–9116, 1999. doi:10.1523/jneurosci.19-20-09107. 1999.
- Hamaguchi K, Mooney R. Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. *J Neurosci* 32: 11671–11687, 2012. doi:10.1523/JNEUROSCI.1666-12.2012.
- Berardelli A, Rothwell JC, Thompson PD, Hallett M. Pathophysiology of bradykinesia in parkinson's disease. *Brain* 124: 2131–2146, 2001. doi:10.1093/brain/124.11.2131.
- Modolo J, Henry J, Beuter A. Dynamics of the subthalamo-pallidal complex in Parkinson's disease during deep brain stimulation. *J Biol Phys* 34: 251–266, 2008. doi:10.1007/s10867-008-9095-y.
- Schiemann J, Schlaudraff F, Klose V, Bingmer M, Seino S, Magill PJ, Zaghloul KA, Schneider G, Liss B, Roeper J. K-ATP channels in dopamine substantia nigra neurons control bursting and noveltyinduced exploration. *Nat Neurosci* 15: 1272–1280, 2012. doi:10.1038/ nn.3185.
- Graybiel AM. Habits, rituals, and the evaluative brain. Annu Rev Neurosci 31: 359–387, 2008. doi:10.1146/annurev.neuro.29.051605. 112851.
- Saddoris MP, Sugam JA, Cacciapaglia F, Carelli RM. Rapid dopamine dynamics in the accumbens core and shell: learning and action. *Front Biosci* E5: 273–288, 2013. doi:10.2741/E615.
- Sutton RS, Barto AG. Reinforcement learning: an introduction. *IEEE Trans Neural Networks* 9: 1054, 1998. doi:10.1109/tnn.1998.712192.
- Bengio Y, Delalleau O. Justifying and generalizing contrastive divergence. Neural Comput 21: 1601–1621, 2009. doi:10.1162/neco.2008. 11-07-647.
- Mihalkova L, Mooney R. Transfer learning with Markov logic networks. ICML-2006 Workshop on Structural Knowledge Transfer for Machine Learning. Pittsburgh, PA, June, 2006. http://citeseerx.ist. psu.edu/viewdoc/download?doi=10.1.1.542.9779&rep=rep1&type=pdf.
- Settles B. Active learning literature survey. University of Wisconsin-Madison Department of Computer Sciences. 2009.
- Wu HG, Miyamoto YR, Castro LNG, Ölveczky BP, Smith MA. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat Neurosci* 17: 312–321, 2014. doi:10.1038/nn.3616.