

2017

Network mechanisms underlying stable motor actions

<https://hdl.handle.net/2144/26503>

Downloaded from DSpace Repository, DSpace Institution's institutional repository

BOSTON UNIVERSITY
SCHOOL OF MEDICINE

Dissertation

**NETWORK MECHANISMS UNDERLYING STABLE
MOTOR ACTIONS**

by

WILLIAM A. LIBERTI III

B.S. Boston University, 2012

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2017

© 2017 by
WILLIAM A. LIBERTI III
All rights reserved

Approved by

First Reader

Timothy Gardner, Ph.D.
Associate Professor of Biology

Second Reader

Ian Davison, Ph.D.
Assistant Professor of Biology

Third Reader

Alberto Cruz-Martin, Ph.D.
Assistant Professor of Biology

Te occidere possunt sed te edere non possunt nefas est

Acknowledgments

This body of work is comprised of a few bricks of success held together by a hidden mortar made from countless failures. Every result described in this thesis was accomplished with the help and support of fellow lab mates and collaborators. Especially I would like to thank Jeff Markowitz, Greg Guitchounts, and Daniel Leman.

I received tremendous support from my thesis committee- especially Ian Davison and Alberto Cruz-Martin. I want to thank my advisor Tim Gardner for teaching me how to think- really, helping me learn how to exercise some control over how and what to think, to balance exploration with being a finisher. Being conscious and aware enough to choose what to pay attention to and to choose how to construct meaning from experience.

I want to thank my friends in Biology, and the Graduate program in Neuroscience- to Shelly and Sandi especially- for letting a wayward soul into the fold.

Finally, I am blessed to have an amazing, talented family who not only provided unwavering support, but also surprisingly fruitful scientific collaborations and insight.

NETWORK MECHANISMS UNDERLYING STABLE MOTOR ACTIONS

WILLIAM A. LIBERTI III

Boston University School of Medicine, 2017

Major Professor: Timothy Gardner, Ph.D., Associate Professor of
Biology

ABSTRACT

While we can learn to produce stereotyped movements and maintain this ability for years, it is unclear how populations of individual neurons change their firing properties to coordinate these skills. This has been difficult to address because there is a lack of tools that can monitor populations of single neurons in freely behaving animals for the durations required to remark on their tuning.

This thesis is divided into two main directions- device engineering and systems neuroscience. The first section describes the development of an electrode array comprised of tiny self-splaying carbon fibers that are small and flexible enough to avoid the immune response that typically limits electrophysiological recordings. I also describe the refinement of a head-mounted miniature microscope system, optimized for multi-month monitoring of cells expressing genetically encoded calcium indicators in freely behaving animals. In the second section, these tools are used to answer basic systems neuroscience questions in an animal with one of the most stable, complex learned behaviors in the animal kingdom: songbirds. This section explores the functional organization and long-term network stability of HVC, the songbird premotor cortical microcircuit that controls song.

Our results reveal that neural activity in HVC is correlated with a length scale of $100\mu\text{m}$. At this mesoscopic scale, basal-ganglia projecting excitatory neurons, on average, fire at a specific phase of a local 30Hz network rhythm. These results show that premotor cortical activity is inhomogeneous in time and space, and that a mesoscopic dynamical pattern underlies the generation of the neural sequences controlling song. At this mesoscopic level, neural coding is stable for weeks and months. These ensemble patterns persist after peripheral nerve damage, revealing that sensory-motor correspondence is not required to maintain the stability of the underlying neural ensemble. However, closer examination of individual excitatory neurons reveals that the participation of cells can change over the timescale of days- with particularly large shifts occurring over instances of sleep. Our findings suggest that *fine-scale drift of projection neurons, stabilized by mesoscopic level dynamics dominated by inhibition, forms the mechanistic basis of memory maintenance and and motor stability.*

Table of Contents

Acknowledgements	v
Abstract	vi
Table of Contents	viii
List of Figures	xiii
List of Abbreviations	xvi
1 Introduction	1
1.1 Motivation	1
1.2 Research outline	4
1.3 Background	6
1.3.1 General Motor Learning/ Maintenance	6
1.3.2 Zebra Finch Song Behavior	7
1.3.3 Song Circuit overview	9
1.3.4 Vocal Motor pathway	10
1.3.5 Anterior Forebrain Pathway (AFP)	10
1.3.6 HVC physiology	11
2 A carbon-fiber electrode array for long-term neural recording	15
2.1 Introduction	15
2.2 Results	18
2.2.1 Fire Sharpening	18
2.2.2 Acute recordings	19

2.2.3	Single-unit recordings in freely behaving birds	20
2.2.4	Simultaneously recorded traces	24
2.2.5	Yield and single unit stability	24
2.3	Methods	28
2.3.1	Carbon microthread arrays	28
2.3.2	Surgical Procedure	31
2.3.3	Electrophysiological Recording	32
2.3.4	Data Analysis	33
2.3.5	Sorted multi-unit and single unit stability using spike features and spike train statistics	34
2.3.6	Imaging	35
3	An open source, wireless capable miniature microscope system	36
3.1	Introduciton	36
3.2	Methods	37
3.2.1	General Miniature Microscope Design	37
3.2.2	Software Design	40
3.2.3	Near-real-time Software Details	40
3.2.4	Auditory Feedback	43
3.2.5	CMOS Sensor and Wireless Transmitter	43
3.2.6	Commutator Design	46
3.2.7	Surgical Procedure	47
3.2.8	Song Alignments	47
3.2.9	Calcium Imaging ROI Analysis	48
3.2.10	Wireless imaging quality quantification	48
3.3	Results	48

3.3.1	Optical Recording of Neurons Expressing Genetically Encoded Calcium Indicators	48
3.4	Discussion	50
3.4.1	Advantage of 3D Printed Housing for Rapid Microscope Assembly and Prototyping	51
3.4.2	All Optical Physiology with Multi-Channel Light Delivery and Recording	52
3.4.3	Low Latency, Open Source Acquisition Software	52
3.4.4	CMOS Trade-Offs and Deficits	53
3.4.5	Use of Off-the-Shelf Components for Ease of Construction	53
4	Mesoscopic Patterns of Neural Activity Support Songbird Cortical Sequences	56
4.1	Introduction	56
4.2	Results	60
4.2.1	The representation of time in HVC projection neurons is spatially correlated over a 100 μm length scale	60
4.2.2	Inhibitory neuron activity is reflected in a 30 Hz local field potential and this field potential is also correlated over a 100 μm length scale	62
4.2.3	The HVC LFP reveals that on average, HVC _X and interneurons fire at different phases of a 25-35 Hz inhibitory rhythm	66
4.3	Discussion	69
4.4	Materials and Methods:	73
4.4.1	Subjects	73
4.4.2	Viral labeling and calcium imaging.	73
4.4.3	Electrophysiology.	74

4.4.4	Analysis of calcium imaging data	75
4.4.5	Bootstrap tests of imaging data.	76
4.4.6	Analysis of electrophysiology data	77
4.4.7	Song alignments	78
4.4.8	Spike sorting.	78
4.4.9	Spectral density images	79
4.4.10	Identifying bursts in projection neurons and interneurons for burst-triggered averaging.	79
4.4.11	Computing firing pattern distance	79
4.4.12	Determining whether interneuron troughs are synchronous. . .	79
4.4.13	Phase locking.	80
4.4.14	Statistical tests	80
5	Unstable neurons underlie a stable learned behavior	82
5.1	Introduction	82
5.2	Results	85
5.2.1	Song is supported by stable mesoscopic dynamics	85
5.2.2	Multi-unit ensembles and single inhibitory neurons are stable .	86
5.2.3	Individual projection neurons are unstable over days.	89
5.2.4	Close inspection reveals the fine microstructure of song also changes over intervals of sleep and increases in consistency through- out the day	93
5.3	Discussion	95
5.4	Methods	98
5.4.1	Subjects	98
5.4.2	Splaying microfiber electrodes	98
5.4.3	Microfiber electrodes for multi-unit recordings	98

5.4.4	Tracheosyringeal nerve cut	99
5.4.5	Calcium imaging	99
5.4.6	Virus Information	100
5.4.7	Song alignments	101
5.4.8	Spectral density images	101
5.4.9	Acoustic change post-nerve cut	101
5.4.10	Acoustic change overnight	102
5.4.11	Local field potentials and single units	102
5.4.12	Multi-unit electrophysiology	103
5.4.13	Calcium imaging	103
5.4.14	Controls for stability of the imaging plane and photobleaching	105
5.4.15	Statistical analysis	106
6	Conclusions	107
6.1	Potential cellular mechanisms for reinforcement learning	107
6.2	Engineering limitations, and future directions.	109
6.2.1	Identifying the downstream projection targets of projection neurons during imaging	109
6.2.2	Structured Illumination Microscopy	111
6.3	ROI based, near-Real-Time Brain Machine Interface (BMI) experi- ments	114
6.4	The role of sleep	116
	References	120
	Curriculum Vitae	133

List of Figures

1·1	Song structure of the male zebra finch	8
1·2	Song circuit overview	9
2·1	Array assembly.	19
2·2	Wide-field imaging of a coated carbon fiber.	20
2·3	Electrode impedances.	21
2·4	Average waveforms of isolated single units recorded acutely with 16 channel carbon fiber arrays.	21
2·5	Single-unit recording in a singing bird.	22
2·6	Clusters for the chronic signal	23
2·7	Chronic recording stability in the singing bird.	23
2·8	A principal cell recorded in HVC.	25
2·9	Intracellular-like traces in singing birds with carbon fibers.	26
2·10	An HVC interneuron recorded on sessions 107 days apart.	27
2·11	Example interneuron raster	28
3·1	A custom 3D printed head-mounted fluorescence microscope.	38
3·2	Performance of closed loop feedback based on near real-time audio or image processing using custom software and GUI	41
3·3	Wireless microscope acquisition system and performance.	44
3·4	3D printed active commutator system for chronic neural recording in small animals.	46
3·5	Images and in-vivo video collected from the microscope.	49

4.1	Three hypothetical models for the spatiotemporal organization of the song premotor code in HVC.	59
4.2	Calcium imaging of HVC projection neurons in awake, behaving birds	61
4.3	Stereotyped network activity during vocal production in songbird premotor	63
4.4	The phase of the 25-35 Hz LFP can be used to study the spatial structure of inhibition.	64
4.5	A stereotyped spatiotemporal 30 Hz pattern underlies premotor cortical activity during song.	64
4.6	LFPs in the 25-35 Hz band and calcium events are correlated over a 100 μm length scale.	65
4.7	Inhibitory neuron firing patterns and three models for projection neuron/ inhibitory neuron relationships.	67
4.8	Spike-field analysis reveals an alternating 30 Hz rhythm in excitatory and inhibitory cells.	68
5.1	Approach to measuring the stability of neural firing patterns underlying a highly stable behavior.	84
5.2	Activity patterns from multi-unit ensembles and single interneurons are highly stable.	85
5.3	Drift in multi-unit firing patterns is not accelerated by unilateral TS nerve transection.	88
5.4	Projection neurons drift over a five day imaging session	90
5.5	Histological verification of virus used for calcium imaging.	91
5.6	Examples of stable and unstable cells in the same imaging region. . .	92
5.7	Examples of amplitude traces from multi-peaked traces.	93
5.8	Projection neuron activity preferentially changes over periods of sleep.	94

6.1	Cell-type specific genetic labeling	110
6.2	Multi-wavelength imaging in a 1P miniature microscope.	111
6.3	Context related changes in HVC projection neurons	113
6.4	Structured Illumination Integration	113
6.5	BMI experimental outline	116
6.6	Spontaneous activity during sleep	118

List of Abbreviations

AFP	Anterior Forebrain Pathway
BMI	Brain Machine Interface
CAD	Computer-Aided Design
DLM	dorsolateral anterior thalamic nucleus
FWHM	Full-Width at Half Maximum
CNC	Complementary metal-oxide-semiconductor
HVC	Computer numerical control
HVC _{RA}	HVC neuron that projects to RA
HVC _X	HVC neuron that projects to Area X
HVC _{AV}	HVC neuron that projects to Avalanche
HVC _I	HVC interneuron
ISIH	Inter-spike interval histogram
LFP	Local Field Potential
LMAN	lateral magnocellular nucleus of the anterior neostriatum
RSV	Rous sarcoma virus
SNR	signal-to-noise ratio
SD	standard deviation
STL	STereoLithography
UVA	nucleus uvaeformis
VMP	vocal-motor pathway

Chapter 1

Introduction

1.1 Motivation

From speech production to athletic performance, implementing reproducible motor gestures is integral to expressing meaningful actions. While it is clearly important to have the ability to execute learned skills with high precision, often over the timescales of years, it is not clear how centrally controlled motor systems maintain this ability. A simple theory is that the individual neurons involved are as stable as the motor skill itself, where reproducible actions result from stereotyped patterns of neural activity. However, it is clear that we rely on external feedback to guide learned skills. Neurons must somehow acclimate to changes in our behavioral state and environment to produce desired motor outputs. These conditions can vary greatly, even over the course of hours, suggesting that individual neurons may need to significantly change their firing properties over these timescales to produce the same motor output. Neuronal circuits must constantly evaluate the tradeoffs between stable operating regimes and plastic adaptation to changing environmental demands. This balancing act is not trivial. In most cases, individual neurons remain physically integrated with surrounding circuitry for several decades. While it is clear that axons and dendrites can be highly dynamic even in adult animals (Holtmaat and Svoboda, 2009), it is not clear to what extent individual neurons maintain stable activity patterns over extended time periods in the face of changing environmental demands.

Even given constant environmental conditions it is not clear how reproducible

the timing and participation of cells in central motor networks are when performing the same action. In fact, the stability of neural tuning for motor systems is not necessarily expected. Vast numbers of neurons in motor cortex converge to relatively few muscles, suggesting that a given muscular activation pattern could be produced by many, potentially redundant, patterns of neural activity. In parallel, single neurons in motor cortex have been shown to switch tuning properties in a task-dependent manner (Koralek et al., 2012). While the production of stereotyped movements is important for many motor tasks, it is equally critical to maintain the ability to flexibly adapt this motor plan in response to changing environmental conditions to achieve dynamic motor objectives.

Neurophysiological evidence suggests that at a mesoscopic scale, neural populations exhibit similar activity patterns over time. In mice, surface cortical layers maintain heterogeneous, log-normally distributed firing rates that remain largely stable across learning (Barth and Poulet, 2012). Sensory cortical areas contain sparsely active neurons interspersed within networks of largely unresponsive cells (Hromádka et al., 2008) (Lefort et al., 2009) (Yassin et al., 2010) and this functional architecture is largely preserved over days to weeks (Margolis et al., 2012), although it is unclear what mechanisms direct these embedded cells (Hromádka et al., 2008). In the hippocampus, earlier electrophysiological methods tracking tens of cells emphasized stable neural tuning over a timescale of a week, whereas recent studies tracking thousands of cells using new optical techniques revealed that 75 – 85% of the cells change their tuning properties within the same timeframe (Ziv et al. 2013). In the whisker motor cortex, individual neurons in mice trained on an object detection task were not active reliably, but the relationship between ensemble measurements and behavior remained stable (Huber et al., 2012a). These studies support the view that for stable behaviors, individual neurons involved can show substantial changes in participation.

Adaptive neural tuning is likely driven by homeostatic processes- neurons retain a tendency toward a relatively stable equilibrium between competing interdependent network elements (Turrigiano, 2008) . Homeostatic tuning has been observed in vivo, particularly in the visual system using chronic multielectrode recordings to monitor firing rates in the visual cortex of freely behaving rats during chronic monocular visual deprivation (Hengen et al., 2013) (Mrsic-Flogel et al., 2007). Similar tuning has been seen in the Barrel cortex (Holtmaat and Svoboda, 2009) and hippocampus (Hirase et al., 2001) (Goold and Nicoll, 2010). Intriguingly, it seems as though much of these homeostatic processes are occurring over periods of sleep (Hengen et al., 2013) (Mrsic-Flogel et al., 2007) (Hengen et al., 2016).

To address how motor networks achieve stable behavioral outputs, it is critical that we study an experimentally tractable animal model that can learn to produce stereotyped motor sequences. Invertebrates such as the larval zebrafish present an appealing model, given recent advances in volumetric video-rate imaging of calcium dynamics in large numbers of neurons during fictive swimming (Ahrens et al., 2012). Yet, it is not clear if larval zebrafish are capable of more complex motor behaviors. Rodents produce stereotyped sequences in their grooming rituals (Colonnese et al., 1996) and to a certain extent their overall behavioral performance (Wiltschko et al., 2015). However, these grooming rituals and behavioral gestures remain challenging to quantify (Wiltschko et al., 2015). While rodents produce ultrasonic vocalizations that are easy to quantify, (Holy and Guo, 2005) it is unclear if ultrasonic vocalizations are learned or even centrally controlled. One promising alternative is to study skilled reaches in mice (Azim et al., 2014). Unfortunately, for all of these behaviors, the central motor circuits underlying each behavior remain undefined.

Songbirds present a unique opportunity to study the neural mechanisms underling learned, stereotyped motor sequences. Birdsong is an easily quantifiable sequential

behavior that is both learned and reproducible on a millisecond time-scale. Zebra finches recapitulate the same song for years (Lombardino and Nottebohm, 2000a) and the neural circuitry that produces song is extremely well studied and well-defined (Nottebohm et al., 1976) (Nottebohm and Arnold, 1976)

Even when environmental conditions and behavioral performance variability are constrained, questions about coding spatiotemporal organization and stability at the single neuron level have been notoriously difficult to address given the technical challenge of stably recording from single neurons over timescales relevant for learning. A major push in this thesis is to develop technologies to maintain chronic recordings of the same neurons over days. I will describe devices developed for extracellular recordings (*Chapter 2*) and for optical imaging of genetically encoded calcium indicators (*Chapter 3*) as well as their deployment to study a behavior that is well-constrained and preserved for years (*Chapter 4 & Chapter 5*). Using these techniques, we have been able to offer insights that reconcile these apparently contradictory views on motor tuning. Our experimental evidence suggests that the brain encodes learned behaviors coarsely at a mesoscopic level that is highly stable, but individual neurons embedded within these networks can change their firing patterns. Possibly, the stability of a memory may be rooted in the robustness of network patterns, where single neurons can adapt to optimize motor output.

1.2 Research outline

This thesis is divided into two parts of approximately equal length. The first part describes the development and deployment of new tools to study motor stability in awake, behaving animals. The second section focuses on the characterization of mesoscopic and fine scale network dynamics of the song premotor cortical zone HVC (used here as a proper name) in zebra finches through in vivo electrophysiology and

calcium imaging.

PART 1: *Engineering tools for long term neural recording in small animals*

- **Chapter 2** discusses the development of the carbon fiber electrode array that enabled the experiments described in *Chapter 4 & Chapter 5*. The array is comprised of a 16 channels that self-splay upon entering the brain like the bristles of a paintbrush, providing densely sampled recordings with a minimally invasive interface.
- **Chapter 3** discusses the development of a 3D printed, wireless capable miniature microscope for single-photon fluorescence imaging in freely behaving animals, that enabled experiments described in *Chapter 5*. The device is made from 3D printed parts and off-the-shelf components. These microscopes weigh less than 2g, can be configured to image a variety of fluorophores, and can be used wirelessly or in conjunction with active commutators. Microscope control software, based in Swift for macOS, provides low-latency image processing capabilities for closed-loop, or BMI, experiments.

PART 2: *Systems level description of HVC dynamics*

- **Chapter 4** describes a series of experiments where we recorded both single units and local field potentials (LFPs) in the freely behaving zebra finch using self-splaying carbon fiber electrodes and miniature head-mounted microscopes. We reveal that excitatory and inhibitory neurons were entrained to opposite phases of a prominent 30Hz LFP, and that nearby projection neurons are active close in time, strongly suggesting that local inhibition gates excitatory neuron firing.
- **Chapter 5** describes that in the song premotor nucleus, HVC, mesoscopic patterns of inhibition are stable for months, but that interspersed populations of

projection neurons coding for song can change from day to day. The most dramatic shifts occur over intervals of sleep. In contrast to the transient participation of excitatory neurons, ensemble measurements dominated by inhibition persist unchanged even after damage to downstream motor nerves. These observations offer a principle of motor stability: spatiotemporal patterns of inhibition can maintain a stable scaffold for motor dynamics while the population of principal neurons that directly drive behavior shift from one day to the next.

- Finally, I conclude with a discussion of future directions based on these experiments in **Chapter 6**.

1.3 Background

1.3.1 General Motor Learning/ Maintenance

Successful sensorimotor learning depends on the generation of a skilled motor behavior that consistently yields outcomes that are expected or favorable. Motor skill learning is canonically defined as the repetitive, trial-and-error process used to form and increase the speed and accuracy of a new motor behavior (Diedrichsen and Kornysheva, 2015) (Shmuelof and Krakauer, 2014) (Shmuelof et al., 2014). Learning a new motor skill begins with a rapid initial improvement followed by more moderate refinements over a longer time period (Karni et al., 1998). The early stages involve exploration of a range of behaviors and outcome-based evaluation and selection. Later in the learning process, repetition-based refinements of the task dominate, driving the formation of a highly stereotyped movement with minimal trial-to-trial variability.

The process of updating a system by reinforcing internal states that led to favorable outcomes is a fundamental tenant of reinforcement learning (Kaelbling, 1996) (Andrew, 1998) (Sutton and Barto, 1998). Framing neurobiological learning in this computational framework has contributed greatly to understanding how biological

networks process information (Niv, 2009) and, in turn, these studies have inspired algorithms for machine learning (Sutton and Barto, 1998). Reinforcement learning explicitly requires exploration. The network probes the consequences of various actions and registers or updates their values. This process allows the network to adaptively and contextually regulate the expression of the probed actions. There is increasing evidence that the brain implements computations predicted by reinforcement learning theory (Daw et al., 2006) (Niv, 2009) (Eshel et al., 2015) (Schultz et al., 1997).

Neuronal circuits have to achieve an ongoing trade-off between attempts to achieve stable operating regimes while evaluating and updating output in order to retain plasticity mechanisms for adapting to changing environmental demands. A critical step in addressing the detailed biological mechanisms in reinforcement learning is to have a well constrained behavior that naturally converges around a set point, where stability in light of environmental challenges is actively maintained.

1.3.2 Zebra Finch Song Behavior

Songbirds present themselves as an excellent model to study the neural activity that produces stable behavior. While the vocal repertoire of the zebra finch is diverse and complex (Woolley et al., 2005), the stereotypical male courtship song provides a unique opportunity for reproducible hypothesis testing, making it one of the most studied and best understood models for motor learning (Williams, 2004). The neural circuits that underlie song behavior are well defined, extensively studied, and in key respects homologous to the cortico-basal ganglia circuits underlying sensory-motor learning in mammals.

For the male zebra finch, singing behavior is a complex, learned, stereotyped motor action that is behaviorally relevant, and, once learned, is performed with great precision. Each song bout consists of many repetitions of a canonical or dominant motif that spans 0.5-1 seconds (Williams, 2004) with an occasional minor variation

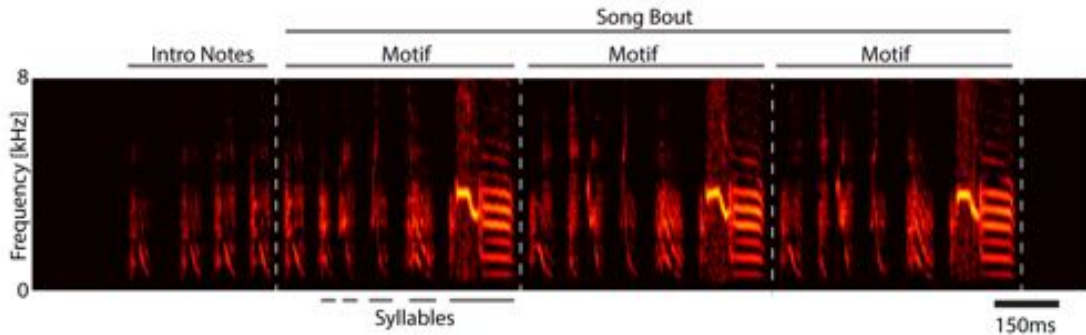


Figure 1-1: Song structure of the male zebra finch Spectrogram demonstrating the behavior of the adult male zebra finch songbird, including the song bout, and its comprising motifs. Each motif, in turn, is comprised of vocal elements called syllables.

(Sossinka and Böhner, 1980). Adult finches will produce hundreds and thousands of motifs per hour in isolation (Pytte et al., 2011). A motif is then comprised of 3-7 smaller vocal elements called syllables that range from tens to hundreds of milliseconds in duration, and each syllable is highly stereotyped (Glaze and Troyer, 2006). The stereotyped nature and repetition frequency of zebra finch song make them well-suited for the study of sequential motor behaviors.

The zebra finch is considered a *closed-ended learner*, which means that they do not learn new vocal elements after the first 90 days after hatching, the period of flexible vocal exploration (Funabiki and Funabiki, 2009). However, adult zebra finches do retain some plasticity in adulthood, and they are able to modify both the spectral and the temporal characteristics of individual syllables, but only to a limited extent. Finches can learn to change features of their song in response to experimentally-delivered conditional auditory feedback (CAF) (Ali et al., 2013) (Andalman and Fee, 2009). Additionally, the stereotypy and spectral structure of song degrades significantly after deafening in the adult, an effect that is reversed by lesions of the basal ganglia (Brainard and Doupe, 2001) (Hamaguchi et al., 2014). Interestingly, there is a dramatic change in variability when zebra finches direct their songs to females.

Even in juveniles, the presence of a female leads to a sharp reduction in spectral variability and duration (Aronov and Fee, 2012) (Kojima and Doupe, 2011). A possible mechanism for this effect is a reduction of variability in spiking in the song motor circuit involved in trial and error learning. (Hessler and Doupe, 1999) (Kao, 2006).

1.3.3 Song Circuit overview

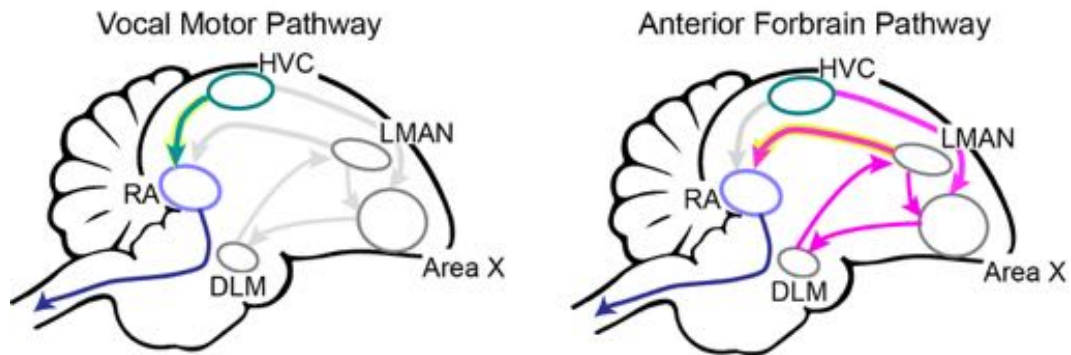


Figure 1.2: Song circuit overview. *a*, *Left*, Vocal Motor Pathway (VMP), which acts as a simple motor hierarchy. *Right*, the Anterior Forebrain Pathway (AFP), which is involved in trial-and error vocal learning. It is important to note that both pathways include the nuclei HVC and RA.

The neural circuitry underlying the zebra finch’s song is one of the best described in neuroscience. This circuit is distributed throughout the brain in small computational engines called *nuclei*, that are distributed into two primary pathways: the *vocal motor pathway* (VMP) that mediates adult, stereotyped vocal performance, and the *anterior forebrain pathway* (AFP) that underlies vocal exploration and learning (Aronov et al., 2008) (Oliveczky et al., 2005). Of critical interest are two nuclei. The first is the premotor nucleus, *HVC*, which is the shared origin of the VMP and the AFP. The second is the *robust nucleus of the arcopallium* (RA), a critical component of the descending motor pathway where both the VMP and the AFP converge. From RA, neurons project to the hypoglossal nucleus *nXII*, in the brainstem and neurons

in the caudal portion of nXII directly innervate the muscles controlling the syrinx (Olveczky et al., 2005) (Vicario and Nottebohm, 1988). RA has a weak projection to the nXII and another projection to brainstem nuclei that controls breathing (Sturdy et al., 2003) (Krützfeldt and Wild, 2004). Both pathways are described in some detail below.

1.3.4 Vocal Motor pathway

At the apex of the VMP, there is a strong non-topographic projection from HVC to RA (Foster and Bottjer, 1998). Physiological evidence suggests that HVC neurons projecting to RA (HVC_{RA}) provide temporal control of song through addressing disparate, discrete locations within RA. In turn, different RA neurons innervate different muscle groups in the bird's vocal organ (Fee and Goldberg, 2011b). Supporting this, unilateral lesions of RA in the adult result in severe deficits in vocal production (Ashmore et al., 2008), whereas unilateral lesions of HVC mostly affect song sequencing (Williams, 2004). This effect is relatively subtle compared to bilateral lesions that cause adults to revert to juvenile-like subsong (Aronov et al., 2008). A one-to-many projection from HVC to RA would allow individual HVC_{RA} neuron bursts to coordinate the activity of multiple muscles at a single point in time. In this way, the VMP is simply a motor control hierarchy.

1.3.5 Anterior Forebrain Pathway (AFP)

Anatomically, in the AFP, projection neurons begin in HVC, which sends axons to the striatopallidal nucleus *Area X*. This area shares strong physiological similarities to the mammalian basal ganglia (Goldberg et al., 2010) (Fee and Goldberg, 2011b). Pallidal neurons in *Area X* project to the dorsolateral division of the medial thalamus (DLM). DLM, in turn, projects to LMAN, which sends bifurcating axon collaterals to both RA and back to *Area X* (Nixdorf-Bergweiler et al., 1995) (Vates and Nottebohm, 1995).

Lesions of LMAN or Area X have not been shown to have a pronounced effect on the structure or timing of adult song (Olivezky et al., 2005) (Scharff and Nottebohm, 1991) (Sohrabji et al., 1990). However, lesions of either DLM or LMAN in juveniles and adults result in a significant decrease in trial-to-trial song variability and song stereotypy increases (Aronov et al., 2008) (Fee and Goldberg, 2011b) (Kao, 2006) (Olivezky et al., 2005). Alternatively, bilateral lesions of HVC in adult birds leads to a regression of song to a state that resembles juvenile-like babbling (Aronov et al., 2008). Together, these results suggest that the AFP, specifically the LMAN \rightarrow RA projections, provides a source of trial and error variability, where the evaluation of this variation could form the basis of trial and error learning. Still, the basic mechanisms of how the AFP injects noise into RA, and how beneficial noise is incorporated into the VMP, remain to be directly observed.

These theories have been tested explicitly in studies where the bird must shift the pitch of a targeted syllable to avoid playback of aversive white noise (Andalman and Fee, 2009) (Tumer and Brainard, 2007). If the AFP is inactivated after multiple days of learning, the pitch will revert back to the outcome of the previous day's learning. (Tumer and Brainard, 2007) This observation strongly implies that the AFP rapidly generates fluctuations to explore advantageous variations and mediates fast learning and that VMP learns new motor programs slowly, potentially over periods of sleep (Andalman and Fee, 2009).

1.3.6 HVC physiology

Projection neurons in HVC can be broadly classified by their downstream projection targets, including basal ganglia-projecting neurons (HVC_X) that project to the AFP and are thought to be involved in trial-and-error learning and motor projecting neurons (HVC_{RA}), which in turn project to brainstem respiratory and vocal nuclei that directly drive the vocal organ (Krützfeldt and Wild, 2004). A third, much larger

projection neuron type that innervates the auditory region avalanche (HVC_{AV}) also sparsely populates the nucleus. Individual HVC neurons fire a brief, 10ms volley of action potentials (where firing rates can exceed 100Hz!) at the same point in a song over many repetitions (McCasland, 1987), sparsely coding for points in time throughout the song. (Hahnloser et al., 2002b). HVC_{RA} neurons burst at a single point in a song with ≈ 1 ms onset jitter between repetitions of the same song. HVC_X cells appear to share the same sparse coding properties of HVC_{RA} neurons, though they tend to burst more than once per motif (Kozhevnikov and Fee, 2007a). HVC_I neurons produce dense, complex, repeating patterns. Given that the downstream neurons in RA burst at many points throughout the song (Leonardo and Fee, 2005), it has been proposed that there is a many-to-one projection from HVC_{RA} neurons. This convergence would allow single HVC_{RA} neurons to act as the ticks of a clock with each 'tick' triggering the recruitment of multiple muscle groups as needed throughout the song.

Mild cooling of HVC with a bilaterally implanted Peltier device results in linear stretching of spectral components of a song, whereas cooling of RA does not affect song timing, suggesting that the neural dynamics generated by intrinsic circuitry within HVC controls moment-to-moment timings of acoustic features in the syllables of adult song. (Long and Fee, 2008) (Fee and Long, 2011). Also, dense sampling of projection neurons reveals largely uniform coverage in time (Picardo et al., 2016a) (Hahnloser et al., 2002b)). These results have suggested that a ballistic propagation of synaptically connected premotor neurons in HVC form a synaptic chain that forms a basic clock, dictating song timing in a top-down manner. However, direct excitatory synapses are sparse (Kornfeld et al., 2017) and excitatory neuron coupling is low (Mooney and Prather, 2005a). Moreover, it is clear that inhibition plays a large role in coordinating and stabilizing these dynamics (Markowitz et al., 2013a).

In a closely related species, the bengalese finch, cooling HVC alters the tran-

sition probabilities between syllables, reducing the number of repetitions of long-repeated syllables and increases the randomness of syllable sequences. In addition, HVC neurons in this species that project to the basal ganglia (HVC_X) display context-dependent activity correlated to syllable repetition and transition (Wang et al., 2008). Additionally, in finches, stimulation of HVC during singing can terminate, restart, or rearrange song syntax (Vu et al., 1998) (Wang et al., 2008). Furthermore, UVA, an upstream input to HVC, displayed activity locked to the offsets of song motifs and song bouts. Taken together, this evidence suggests that timing and probabilistic sequencing of motor action can share the same localized neural circuit within HVC and that it can be influenced by its upstream inputs.

Previous studies of HVC have only characterized the activity of small numbers of single neurons recorded serially (Hahnloser et al., 2002b) (Kozhevnikov and Fee, 2007a) Long et al., 2010; Yu and Margoliash, 1996). Until the work described in this thesis, no group has directly observed the relationship between single cells and network activity during the production of song. Rectifying the mismatch between models describing HVC dynamics and the lack of experimental data to validate them is a critical aim of this thesis. Even now, we are still only taking the first steps toward a long journey of understanding HVC at single cell resolution, and on a network scale. For example, basic questions about network function, such as how and to what degree the stability of neuronal activity is balanced with plasticity in brain circuitry, has been largely ignored- primarily because addressing these challenges requires longitudinal measurements from the same neurons in vivo, which has been difficult to achieve using classical neurophysiological techniques. The work described in the following chapters only scratches at the surface of these mysteries.

Part I

Engineering tools for long term
neural recording in small animals.

Chapter 2

A carbon-fiber electrode array for long-term neural recording

This chapter is a reproduction of previously published work (Guitchounts et al., 2013), and contains critical contributions from *Dr. Jeffery Markowitz*, as well as *Gregory Guitchounts*, who are equal co-authors on this manuscript.

2.1 Introduction

A powerful approach to the study of learning involves tracking neural firing patterns across time. Optical methods for stable recording are developing rapidly (Harvey et al., 2012) but the temporal resolution of electrical recordings remains unsurpassed, and chronically implanted microelectrodes are central to scientific studies of neural circuit function in behaving animals, and central to the development of intracranial brain-machine interfaces in humans (Donoghue, 2008). In primate motor cortex, relatively large neurons can be tracked for weeks (Dickey et al., 2009; Fraser and Schwartz, 2012; Tolias et al., 2007) using commercially available electrode arrays. This has enabled researchers to study the stability of motor tuning (Chestek et al., 2007; Rokni et al., 2007) and the process of memory formation (Kentros et al., 2004; Thompson and Best, 1990); and it has provided the basis for brain machine interface technologies (Koralek et al., 2012). Over time, chronically implanted electrodes are severely limited by a tissue reaction that eventually encapsulates the electrode, killing neurons in the vicinity of the electrode. The limitations of this tissue response are

particularly acute if the goal is recording from densely packed neurons in small brains. For a silicone array whose cross-section is $15 \times 200 \mu\text{m}$, histological markers of gliosis and neuron death reveal tissue damage extending up to $300 \mu\text{m}$ or more from the implant (Biran et al., 2005). This length-scale of tissue damage does not prohibit long term recording from pyramidal neurons in primate cortex whose large polarized dendrites and large somas (up to $100 \mu\text{m}$) produce a strong signal for extracellular recording. However, the length scale of tissue damage becomes prohibitive when recording from many cell types in smaller organisms. For example, in the songbird nucleus HVC (used here as a proper name), somas are only $8\text{-}15 \mu\text{m}$ in diameter (Mooney and Prather, 2005) and closely-packed in clusters making soma-soma contact (Scott et al., 2012). Dendrites in HVC are also compact ($40\text{-}100 \mu\text{m}$ radius), and spherical in shape rather than polarized (Katz and Gurney, 1981; Lewicki, 1996; Mooney and Prather, 2005). To isolate the weak signal generated from these cells, electrodes with small recording surfaces are advanced with motorized microdrives, allowing μm -scale control over electrode positioning. The absence of any report of single neuron isolation in HVC with a fixed chronic electrode implant underscores the difficulty of recording small cells with an implant whose damage length scale is large relative to the target neurons. Examples of multi-day recordings in mouse hippocampus can be found (Kentros et al., 2004; Koralek et al., 2012) that employ movable tetrode microwires, but recording methods that make this process more efficient are needed. Ultimately, to facilitate long term recordings from densely packed neurons in small animals, and to improve the longevity of human and primate neural interfaces, electrode designs are needed that reduce chronic tissue damage. The cross-section of implanted electrodes must be minimized to reduce chronic disruption of the blood brain barrier (Biran et al., 2005; Polikov et al., 2005). Orthogonal pressures are driving an increase in the number of recording sites per implant, but increasing the

density of recording sites may be counter-productive if the implant size also increases.

Recently proposed carbon fiber ultramicroelectrodes promise to reduce damage upon implantation, and may partially evade immune rejection (Kozai et al., 2012). Glass insulated carbon fibers have been used for cyclic voltammetry and extracellular recording for some time (Armstrong-James and Millar, 1979; Garris et al., 1994), but the essential advance of the proposed microthread electrode involves doing away with the large-diameter glass insulation in favor of a thin ($1\ \mu\text{m}$) layer of parylene deposited over a small ($3\text{-}7\ \mu\text{m}$) diameter carbon fiber. This also serves to dramatically reduce the stiffness of the implant, another factor hypothesized to contribute to tissue damage (Subbaroyan et al., 2005). The small profile over the full length of the electrode, in principle, could provide for minimal chronic tissue damage and neuron death. The small scale of the proposed carbon microthread electrode makes implantation challenging. Practical methods for building and implanting high channel-count carbon fiber electrodes remain undefined. The present study describes one such design.

The 16 channel carbon fiber array described here has a final cross-section of approximately $26\ \mu\text{m}$, on par with single micro-wires in commercial microwire arrays ($30\text{-}50\ \mu\text{m}$ wires in all Tucker Davis microwire arrays, for example.) The principle innovations of this work arise from carbon fiber manipulations using the elements of fire, air and water: burning partially submerged carbon fibers leads to a consistent electrode tip preparation. Drawing carbon fibers through an air-water interface leads to surface-tension driven bundling of the fibers, resulting in an implantable array. To test the design, we examined the neural representation of song execution in zebra finches- a uniquely tractable test-bed for assessing chronic electrode stability. Zebra finch song is a stereotyped natural behavior produced by a distinctive learned pattern of neural activity. As the bird sings, each of three cell types in pre-motor nucleus HVC interneurons, basal ganglia-projecting neurons, and motor output neurons dis-

charge in a stereotyped, characteristic pattern with minimal spike-time jitter (<1 ms for motor output neurons) between renditions of the same song (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007).

The various cell types in HVC are small (8-15 μm), and synchronous elevated firing rates of interneurons during song make isolation of single units challenging. Current approaches to recording the three cell types in HVC require the use of high impedance electrodes positioned close to individual cells using a motorized microdrive (Fee and Leonardo, 2001). Neurons isolated in this manner in HVC are typically not recorded for a time-scale longer than tens of minutes. Using fixed implants of the carbon fiber arrays described here, we recorded unambiguous single units in HVC over timescales of 4-12 hours, and sorted multi-unit clusters over time-scales of months. We find that the precise temporal patterns recorded in HVC are stable over the time-scales of our recordings. This feature provides a means of validating signal stability independently of spike waveforms, providing a useful test-bed to assess chronic recording methods.

2.2 Results

2.2.1 Fire Sharpening

The underwater firing process exposed 89 ± 17 (SD) μm of insulation as measured by widefield and confocal microscopy ($n = 34$ tips), leaving sharpened, uniform tips (see SEM images in Figure 2·1c and fluorescent parylene images in Figure 2·2). The torching process produced tips with an average impedance of $1.26 \text{ M}\Omega$ ($n = 210$ tips) (Figure 2·3a). Impedances measured in vivo after implanting showed a wide range of values in the weeks following implant (Figure 2·3b).

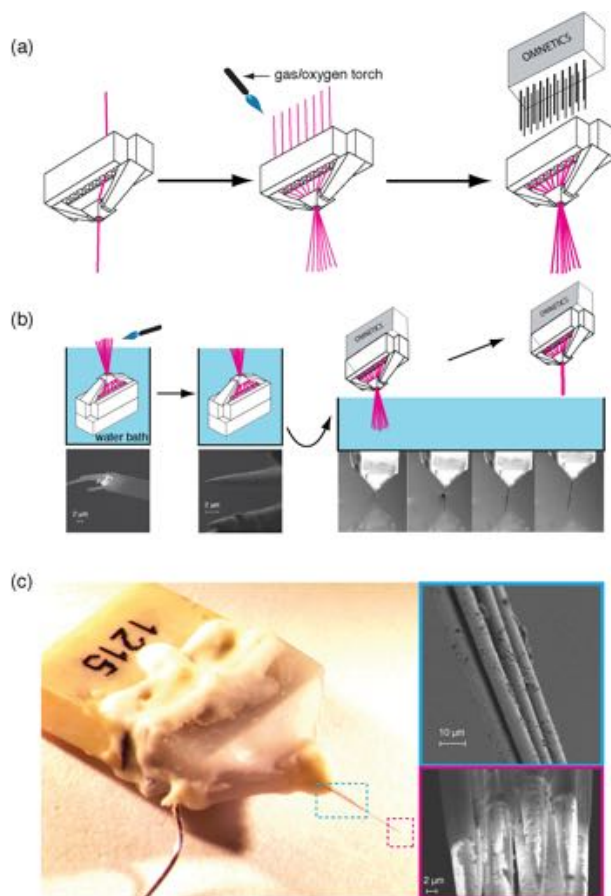


Figure 2-1: Array assembly. **a**, *Left*, Diagram of the 3D-printed plastic block with wells for 16 carbon fibers. Fibers are threaded through the wells on the top of the block and exit through the hole on the bottom. *Middle*, To expose the connector-side ends of the fibers from parylene, the fibers are heated by passing them through a gas/oxygen torch. *Right*, The wells are then filled with conductive silver paint to make electric contact with an Omnetics connector, which slides into the plastic block wells. **b**, Fire-sharpening of electrode tips. *Left*, the assembled array is lowered into a water bath with the tips of the carbon fibers protruding above the surface of the water. *Bottom*, SEM image of a blunt cut carbon fiber electrode, with insulation frayed near the tip. *Middle*, A gas/oxygen torch is passed over the surface of the water, burning the carbon and the insulating parylene down to the water surface. *Bottom*, After passing the torch over the exposed tips, the carbon fiber tapers to a sharp point. *Right*, The array is then taken out of the water, with the tips pointing down; surface tension acts to bring the carbon fibers into a single tight bundle. **c**, The assembled array. The 16-electrode bundle is approximately $26\ \mu\text{m}$ in diameter (upper right). The immersion tip burning process exposes approximately $89\ \mu\text{m}$ of carbon (lower right).

2.2.2 Acute recordings

To initially assess the viability of carbon fiber electrode arrays for recording at various depths we recorded extracellular signal acutely in anesthetized or awake head-fixed

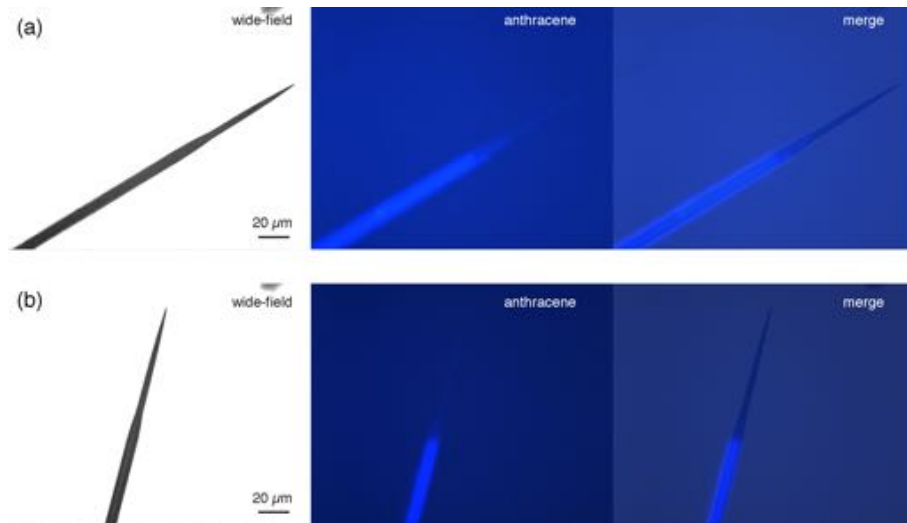


Figure 2.2: Wide-field imaging of a coated carbon fiber. **a**, *Left*, wide-field image of a carbon fiber coated in fluorescent Parylene-C, after it had been fire-sharpened. *Middle*, UV filter image for the same fiber. *Right*, Merged images of the wide-field and fluorescence view shows that the fire sharpening method exposes an average of $89 \pm 17 \mu\text{m}$. **b**, Another example fiber.

birds ($n=4$ cells from 4 birds). Figure 2.4 shows average waveforms from well-isolated neurons and spike trains recorded in auditory area Field L in awake head-fixed birds with an SNR of 9.18 and 3.00 (Figure 2.4a and 2.4d, respectively); in premotor nucleus HVC in an anesthetized bird with an SNR of 21.64 (Figure 2.4b) and in basal ganglia nucleus Area X in an anesthetized bird with an SNR of 3.50 (Figure 2.4c). Carbon fiber arrays were thus able to measure signals from a range of cell types across a variety of brain regions, including a recording zone 3.0 mm deep (Area X).

2.2.3 Single-unit recordings in freely behaving birds

Our primary goal was to develop an electrode capable of tracking single units and small multi-unit clusters over extended periods of time. Thus, to assess the reliability and longevity of single cell signal, we implanted 16 channel carbon fiber arrays into the premotor nucleus HVC ($n=10$ birds), an area that produces precise neural firing patterns that ultimately drive muscular sequences to produce song (Long et al., 2010).

Spikes were aligned to all renditions of a given vocal element for a single day

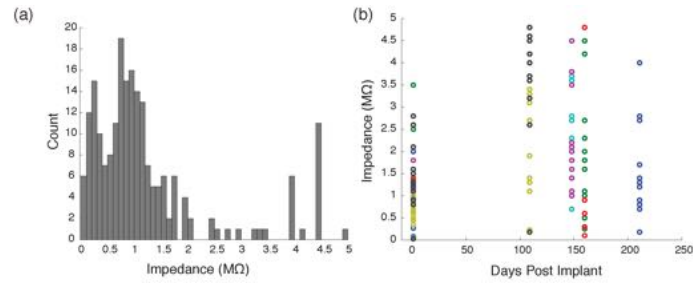


Figure 2-3: Electrode impedances. **a**, Histogram of the fire-sharpened pre-implant electrode impedance ($n = 210$ fibers; median = $1.0\text{M}\Omega$). **b**, Impedance of fibers in 7 implanted arrays measured at various time points after implanting. Colors indicate fibers grouping into arrays. The pre-implant impedances (in saline) in corresponding colors are shown at Day 0.

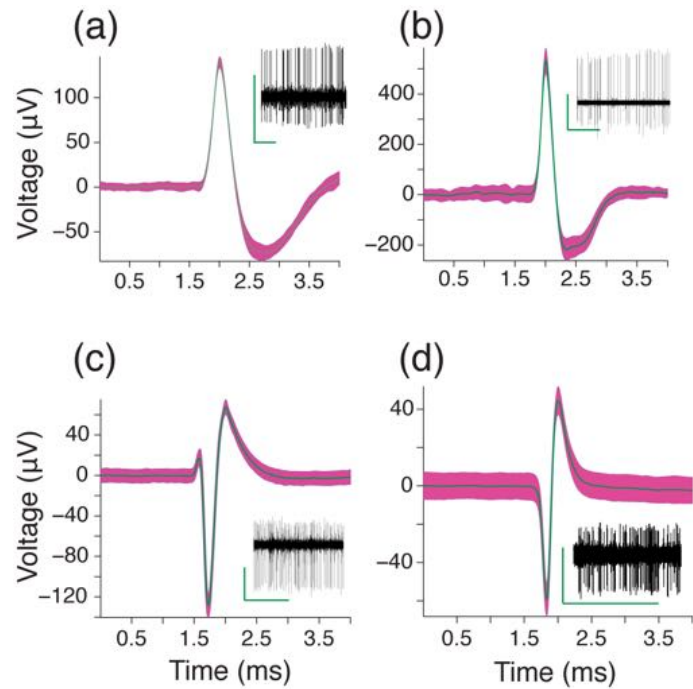


Figure 2-4: Average waveforms of isolated single units recorded acutely with 16 channel carbon fiber arrays. **a**, A unit recorded in auditory area Field L in an awake head-fixed bird (SNR=9.18). **b**, A unit from the pre-motor nucleus HVC in an anesthetized bird (SNR=21.64). **c**, An isolated unit found in the basal ganglia of an anesthetized bird (SNR=3.50). **d**, A unit recorded in Field L of an awake, head-fixed bird (SNR=3). Insets show corresponding spike trains. Inset scale bars: (A) $200\ \mu\text{V}$, 50 s; (B) $400\ \mu\text{V}$, 100 s; (C) $100\ \mu\text{V}$, 1 s; (D) $100\ \mu\text{V}$, 5 s. Pink indicates standard deviation.

and displayed as a raster plot (Figure 2-5). Figure 2-7 shows one such raster from a putative HVC interneuron (classified as single-unit by the standard criterion, see

Section 2.3.5) recorded over a period of 15 days. Average waveforms and ISI histograms from the 1st, 7th and 14th days are consistent throughout the period (Figure 2.6). The average firing patterns are also stable over these time-scales.

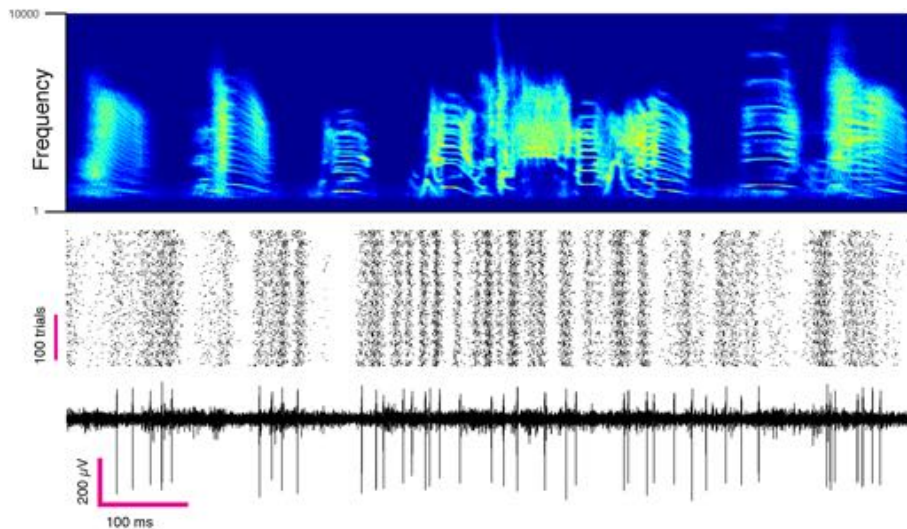


Figure 2-5: Single-unit recording in a singing bird. Example of a putative interneuron recorded in the pre-motor nucleus HVC aligned to song. *Top*, the time-frequency histogram of aligned renditions of the same song motif. *Middle* and *Bottom*, spike raster from a single unit aligned to song and a raw trace from the same channel.

Additionally, in some cells we found distinctive waveforms and discharge patterns characteristic of principal neurons; that is, sparse high-frequency bursting aligned to a single point in the birds song (Figure 2-8). The cell recorded for this figure met the standard criterion for single unit isolation. Prior recordings of this neuron type have required high impedance electrodes mounted on motorized microdrives that allow for fine positioning of the electrode in the vicinity of the neuron (Hahnloser et al., 2002). This is the first report of HVC projection neuron recordings from immobile chronic implants. The yield and longevity of all recorded neuron types are reported below.

On two occasions, a portion of the sharpened tip of an electrode apparently entered a cell, yielding intracellular-like traces that were stably held for 12 hours in one case and 36 hours in the next (Figure 2-9). The intracellular-like cells recorded in area

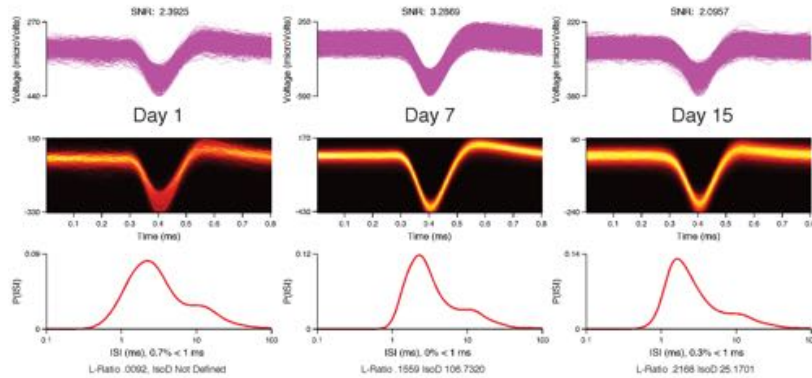


Figure 2-6: Clusters for the chronic signal shown in previous figure. *Top row*, overlaid spike waveforms. *Middle row*, spike waveform histograms. *Bottom row*, spike ISIHs. The SNR changed from 2.39 on Day 1 of recording (left column) to 3.29 on Day 7 (middle column) and 2.09 on Day 14 (right column).

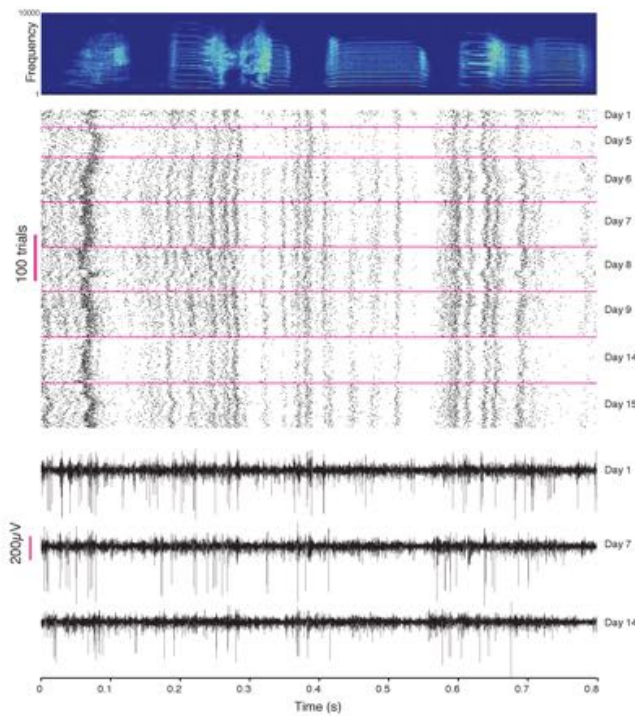


Figure 2-7: Chronic recording stability in the singing bird. *Top*, A putative HVC interneuron (single unit by the standard criterion) in bird recorded over 15 days. *Bottom*, raw traces from the same channel on days 1, 7 and 14. Signal fading (as on day 14) indicates periods of partial loss of cell isolation.

X and HVC were characterized by high amplitude spikes and positive subthreshold voltage ramps prior to spikes or bursts and stereotyped hyperpolarizing potentials

following the burst in the area X cell. In these rare recordings a portion of the uninsulated ($80\ \mu\text{m}$) tip must have remained outside the cell (Angle et al., 2012).

2.2.4 Simultaneously recorded traces

Of particular interest in multi-electrode recordings is not only the longevity of single-unit signal, but the ability to record multiple signals at once. In chronic implants of carbon fiber arrays, high quality multi-unit signal was often present on the majority of electrode contacts, though multiple rigorous single units were not recorded simultaneously on any implant. With the small diameter and proximity of electrodes, individual neurons were occasionally visible on multiple channels simultaneously. Features of correlated signal across channels (i.e. tetrode effect) are commonly necessary to isolate densely packed neurons, but these features were not used here. However, we have observed examples of channels with common signal on two channels from birds implanted with 16 channel arrays. This examples illustrate the potential of improving single unit isolation based on multi-electrode features in future carbon fiber electrode designs.

2.2.5 Yield and single unit stability

Single units defined by standard criteria were defined as containing (1) adequate SNR (> 1.1) and (2) a minimal fraction of ISIs shorter than a refractory period of 1 ms (< 5). After discarding clustered units that did not meet these criteria, we recorded an average of 5.3 neurons per bird as defined by the standard criterion ($n=6$) that ranged from as few as 2 neurons to as many as 8. This count includes both putative interneurons and projection neurons, classified according to their firing patterns.

Four rigorous single units ($n=3$ interneurons and $n=1$ projection neuron) ($\text{SNR} > 2.8$ and 0% ISI violations) were found in a set of $n=4$ implants. Finally, 16 projection neurons that did not meet the rigorous single unit criteria were recorded in at

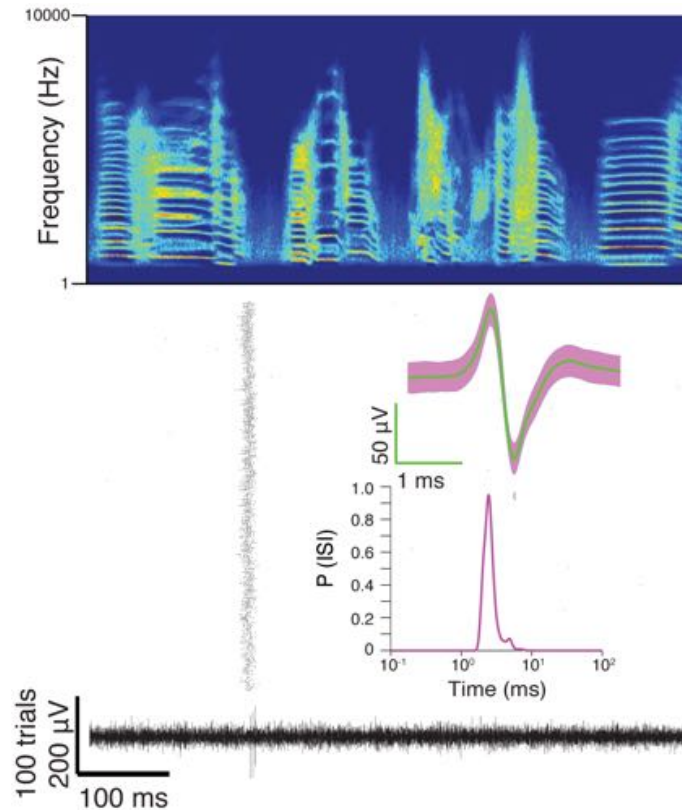


Figure 2-8: A principal cell recorded in HVC. *Top*, Song-aligned spike raster of a putative RA-projecting neuron. *Bottom*, The raw voltage trace from rendition 199 out of 500 song renditions recorded across 1 hour and 26 minutes. Insets show the average waveform with SD and ISI distribution.

total of $n=10$ implants ($\text{SNR} < 2.8$, though for this population all cells had 0% ISI violations except two, which had 0.3%). These cells produce a single burst one or two stereotyped times in the song, and single unit isolation could be confirmed in spite of the low SNR values. For this population of cells, we computed the fraction of contaminating spikes (spikes occurring within 100 ms of the burst, but outside a 10 ms window around the average burst time). The fraction of contaminating spikes ranged from $F=0.425$ in the most marginal case to $F=0.008$ in the best case, with an average of $0.128 \pm .135$ (SD). Examples of these raster plots are illustrated in Figure 2-8.

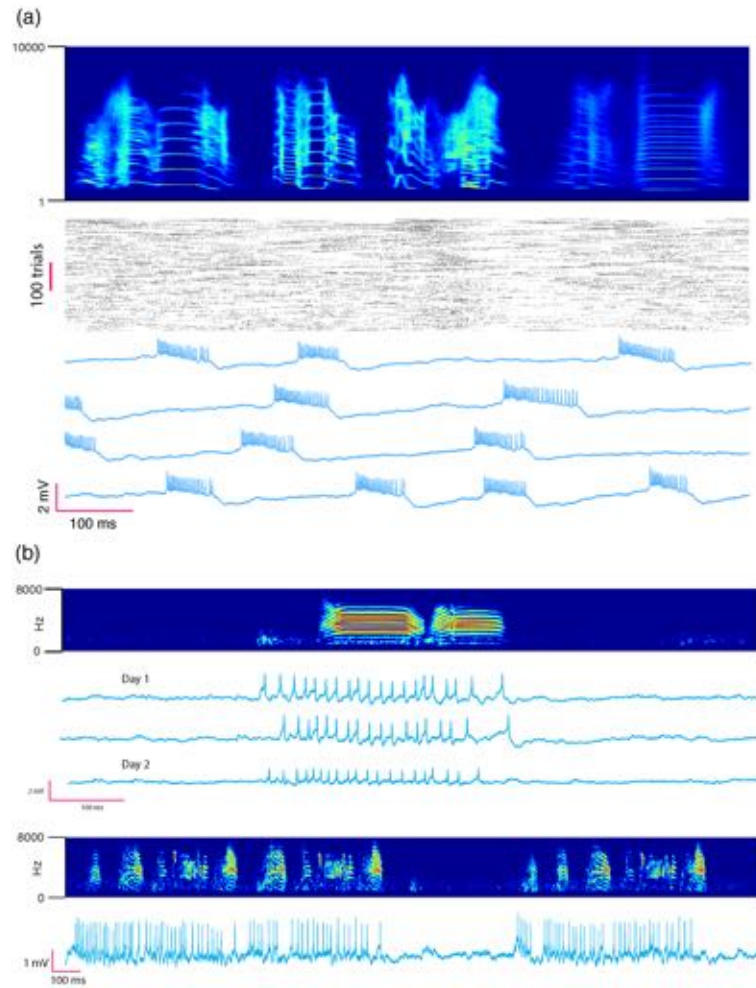


Figure 2-9: Intracellular-like traces in singing birds with carbon fibers. **a**, a bursting cell with high amplitude positive peaks recorded from a bird implanted in Area X. *Top*, sonogram of the song; *Middle*, spike raster; *bottom*, low-pass filtered traces showing prominent LFP modulation during spiking. **b**, A similar cell recorded from a bird implanted in HVC. Unfiltered traces during a call, across two days (top); and during singing (bottom).

To assess the stability of isolated units, we followed previously described methods (Dickey et al., 2009; Fraser and Schwartz, 2012). From one day to the next, a stable waveform indicated continuity of recording from a single neuron. However, considering the population of all neurons recorded, waveform shapes were not unique, nor were ISIHS. One solution is to increase the dimensionality of the waveform shape by considering projections of the waveform on additional channels. However, in HVC

of adult songbirds, a more powerful approach is possible based on the unique spike patterns produced by different neurons in HVC (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007). For each neuron observed here, spike timing patterns were unique and stable across time. This is true both for clusters consisting of small numbers of cells, single interneurons (Figure 2·10), and projection neurons (Figure 2·8).

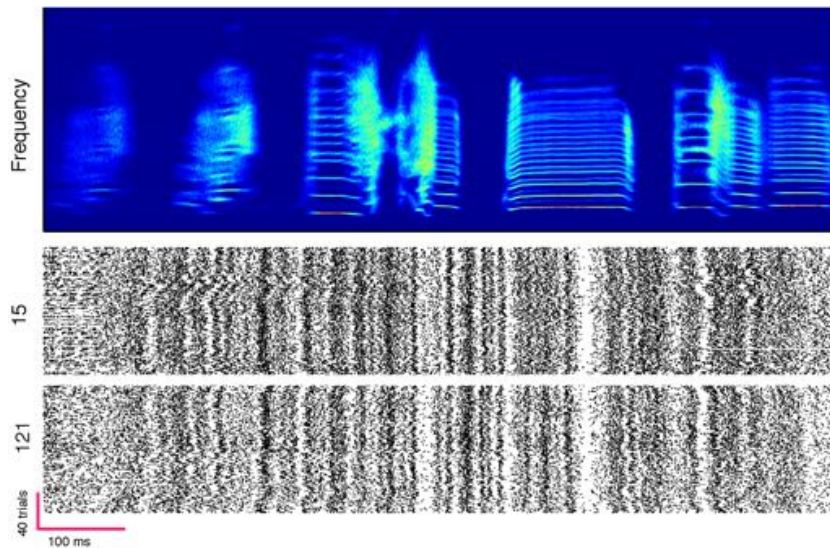


Figure 2·10: An HVC interneuron recorded on sessions 107 days apart. Analysis on waveform shape, ISI distribution and firing pattern classifies the recordings as the same sorted cell, based on the standard criterion. Days post implant are shown on the left of each raster.

Of the 27 neurons that passed the standard SNR and ISIH quality criteria ($n=6$ implants), we analyzed the 18 standard quality cells that were stable for more than one day (for the other 9, no suitable clusters on the same electrode channel were found after the first recording session). Of this set of 18, the cluster quality varied, with 5/18 having $< 1\%$ ISI violations and the rest $< 5\%$. Of this total, 11/18 were stable for one week, 6/18 for two weeks, and 4/18 for 30 days. Projection neurons were not recorded for more than 2 days.

For the rigorous single units ($n=4$) we analyzed all putative cell types. The longevity of each single unit recording is shown for each of 4 neurons, which ranges

from 4 to 12 hours. (Outside of this time-scale, the cells fell below the threshold for rigorous single units, though they were still isolatable based on standard criteria that allowed for some error.)

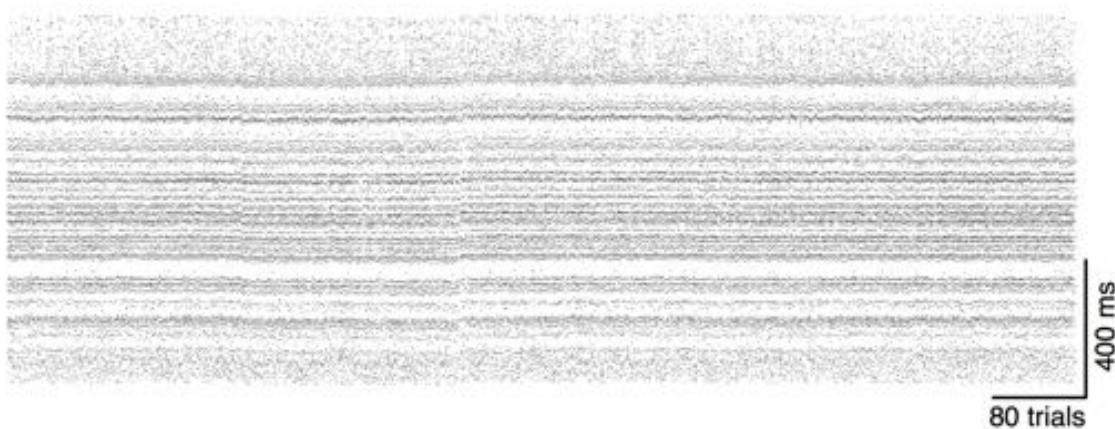


Figure 2-11: Spike raster for example neuron from Figure 2-10. All trials included in the analysis for the bottom example from Figure 2-10 are shown in order from left to right.

2.3 Methods

2.3.1 Carbon microthread arrays

4.5- μm diameter carbon fiber threads (Goodfellow USA, Grade UMS2526) form the basis of the array. (Young's modulus of 380 GPa compared to tungsten's 400 GPa; volume resistivity 1000 $\mu\Omega\text{cm}$ compared with 5.4 $\mu\Omega\text{cm}$ for tungsten.) Epoxy sizing was removed by heating fibers at 400°C for 6 hours (Lin et al., 1998) using a Paragon SC2 kiln (Paragon). A 1 μm layer of Parylene-C (di-chloro-di-p-xylylene) (Kisco) was then deposited using an SCS Labcoter 2 (PDS-2010, Specialty Coating Systems) using 2.3g of parylene and factory settings as follows: Furnace, 690°C; Chamber Gauge, 135°C; Vaporizer, 175°C; and Vacuum at 15 vacuum units above base pressure. The integrity of the coating was verified through bubble testing initially, and defects were never found. Fourteen coated fibers and 1 coated or uncoated fiber (for the reference) were

threaded through a custom plastic block designed in SolidWorks (Dassault Systems) and 3-D printed using stereolithography (Realize Inc) (one channel out of the possible 16 was used for the microphone trace and one was shorted with the reference, making a total of 14 electrophysiology channels) (Figure 2.1a, left). The block contains 16 wells (18x27 mils) that funnel the carbon fibers through a small aperture (12 mils in diameter) (block design files are available upon request). At this stage the carbon fibers exiting the aperture are splayed. On the top side, the block was cut to fit the straight tails of an Omnetics connector (A79038-001, Omnetics). Fibers at the mating end of the carbon were briefly passed through a gas/oxygen torch to remove insulation for making electrical contact (Smith Equipment) (Figure 2.1a, middle). The fibers were then connected to an Omnetics connector using conductive silver paint (Silver Print part no. 842-20G, MG Chemicals), which was spread into the wells housing the carbon fibers. The connector was pressed into the silver-filled wells and glued to the plastic block using light bonded dental acrylic (Flow-It ALC, Pentron Clinical) (Figure 2.1a, right). Initially, the carbon fiber bundle was blunt cut with surgical scissors (Fine Science Tools) or a razor blade to expose the insulation. This resulted in widely varying impedances, often as high as 4 M Ω (measured at 1kHz with a Bak Electronics IMP-2 impedance tester). All impedance measures reported here are in phosphate buffered saline solution (Sigma Aldrich D5773 SIGMA). We next tried embedding the wires in carbowax (polyethylene glycol) and cutting using a vibratome, but this also did not reduce impedance. We then tried grinding the tips of the fibers on a spinning plate (a modified hard drive) for up to 30 minutes, but this did not produce desirable impedance either. To produce a consistent low-impedance tip, we developed a process of fire-sharpening with a gas/oxygen torch (flame 4.5 mm across, 7.5 mm in length). The key innovation involves holding the array underwater while burning exposed tips (Figure 2.1b). With the array secured in a water bath with car-

bon fiber tips protruding above the surface of the water (Figure 2·1b, left), the carbon was burned down to the surface of the water with the torch (Figure 2·1b, middle). The water acted as a flame retardant/insulator, providing control over the amount of Parylene-C taken off of the tip of the carbon. This had two desirable effects: (1) sharpening the tip of the bundle, and (2) lowering the tip-impedance to an acceptable range for extracellular recording (Figure 2·3). By comparison, we found blunt-cutting of the Parylene insulated electrodes to produce widely varying impedance values and unreliable signals. Figure 2·1b, left panel, shows one blunt-cut electrode, revealing a carbon recording surface recessed from the cut Parylene surface. A recent study found it necessary to surface treat parylene-insulated carbon fibers with PEDOT for recording chronic extracellular signals (Kozai et al., 2012). The large surface area of the exposed carbon in the fire sharpened electrode may explain why chronic signals were found in the present study without additional surface modifications required in the previous study. In a final step, the array was slowly drawn out of the bath with the wires facing down. In this step, surface tension pulls together the splayed fibers into a single bundle, and this bundle remains together after the fibers dry (Figure-2 1b, right), allowing the entire bundle to be implanted. Figure 2·1c shows the (approximately 250 mg) assembled 16 channel array. This final weight is comparable to commercially available electrode arrays with a similar number of contacts (300 mg for a 16 channel Omnetics TDT array, 140 mg for a 16 channel H-style probe from Neuronexus, and 130 mg for a 16 channel microwire array from Microprobes). In the final construction, the wires converged in one bundle with a bundle diameter of 26 μm (Figure 2·1c). Each fiber terminated in uninsulated carbon, in a conical profile. The length of this cone was 89 ± 17 (SD) μm . The diameter of the bundle, along the full length of the array, is smaller than single wires in most commercial arrays (33-50 μm for all Tucker Davis microwire arrays). Considering the time required to insulate

the carbon, burn the fibers, and test impedance, array construction typically takes 3-4 hours for an experienced electrode builder. In a sample of 16 arrays constructed by an expert, an average of 92 ± 17 (\pm indicates SD) of the wires were functional ($n = 210$ fibers), where functional electrodes are defined as having an impedance lower than $4 \text{ M}\Omega$. For an experienced electrode builder, the failure rate is low.

2.3.2 Surgical Procedure

All procedures were approved by the Institutional Animal Care and Use Committee of Boston University (protocol number 09-007). Zebra finches ($n = 14$; 4 for acute experiments, 10 for chronic HVC recordings and 2 for intracellular-like recordings) (> 120 days post-hatch) were anesthetized with 4.0% isoflurane and maintained at 1-2% isoflurane during the aseptic surgical procedure. The analgesic Meloxicam ($120 \mu\text{L}$) was injected intramuscularly into the breast at the start of the procedure and the animal was placed into a stereotaxic instrument. Feathers were removed from the scalp and a Betadyne solution applied. Bupivacaine ($50 \mu\text{L}$) was then injected subcutaneously into the scalp before an incision was made along the AP axis. The skull over HVC was localized using stereotaxic coordinates (30 head angle; 0.7 mm AP, 2.3 mm ML, 0.4-0.7 mm DV), and the the outer bone leaflet removed at the location of HVC with a dental drill. The lower bone leaflet was carefully removed with an ophthalmic scalpel, similar to implant procedures for recording with microdrives (Long et al., 2010), exposing a hole of $100 \mu\text{m}$ diameter. A minimal durotomy was performed using an ophthalmic scalpel (typical durotomy was less than $50 \mu\text{m}$.). Electrodes were mounted on a digital manipulator attached to the stereotax and slowly lowered through the durotomy. During insertion into the brain, the carbon fibers would occasionally begin to visibly splay. After lowering the array to the appropriate depth, the position in the song nucleus HVC was verified using antidromic stimulation from a bipolar electrode implanted in downstream Area X (Hahnloser et

al., 2002). After verifying the position of the array, the craniotomy was covered with the silicone elastomer Kwik-Sil (World Precision Instruments) and the array was glued into place using light-bonded acrylic (Flow-It ALC, Pentron) along the entire length of the electrode shank, such that no portion of the carbon fiber bundle was left exposed or loose. The same surgical procedure was followed for the acute recordings in Area X and HVC. For acute recordings in auditory area Field L, adult (> 120 DPH) female zebra finches were anesthetized with 4% isoflurane (maintained at 1-2%) and a custom-made stainless steel head-post was glued to the scalp over the left hemisphere. The birds were then given several hours to recover, after which they were head-fixed in a soundproof chamber. Recordings were made from spontaneously active cells in Field L. Following the acute recordings, birds were euthanized using 200 μ L sodium pentobarbital (250 mg/kg) injected into the breast muscle.

2.3.3 Electrophysiological Recording

Acute recordings were performed using an RZ-5 BioAmp Processor and Medusa Pre-Amplifier (Tucker-Davis Technologies). Data was sampled at 25 kHz with filter cutoffs set to 300 Hz and 20 kHz. (Data from the acute recordings were not impacted by any aliasing due to the proximity of the high frequency filter and the sampling rate. See Figure 2.2.) To record from behaving birds, we used the Intan Technologies 16 channel multiplexing headstage (RHA2116 with unipolar inputs) paired with the RHA2000-EVAL board for acquisition at 25 kHz. These head stages were configured with a fixed 11kHz lowpass filter. To send signals from the headstage to the evaluation board (RHA2000-EVAL), a custom flex PCB cable was designed that connected the headstage to a commutator (9-Channel SwivElectra, Crist Instruments), which then passed signal to the RHA2000-EVAL. All data was analyzed off-line using a series of custom MATLAB (Mathworks) scripts. During the experiment birds were recorded continuously for five days each week and left off the flex-PCB cable for

two days a week. To record singing, a miniature microphone (Knowles Acoustics; Digi-Key catalog # EM-23046-P16) was glued to the Intan headstage and recorded through an electrode channel, following a protocol that is now available online. For the data recorded here, the reference electrode for all channels and microphone was an uninsulated carbon fiber bundled along with 15 insulated fibers. Prior to spike sorting, this reference signal was occasionally replaced offline with a common average reference subtraction (Ludwig et al., 2009). In preliminary tests, referencing to one of the insulated electrodes also works fine. In practice, the referencing is accomplished by bridging the reference channel on the Intan headstage to an arbitrary electrode pin on the Intan headstage. (Note that the online protocol for microphone recording referenced above should be modified to ground one of the electrodes, so that the reference signal can be recorded. Subtracting the reference from the microphone signal offline will provide a cleaner microphone signal, though in practice referencing the mic to the brain works well most of the time, since the two signal amplitudes are of very different scales.)

2.3.4 Data Analysis

Song bouts were detected by looking for threshold crossings in the average power of the microphone trace between 2-6 kHz. Song syllables were clustered using previously defined methods (Poole et al., 2012). We used a manual cluster cut to train a support vector machine (Cortes and Vapnik, 1995), which subsequently identified all instances of that syllable automatically. To analyze single-unit activity, voltage traces were aligned to singing and high pass filtered with an 800 Hz cutoff (2nd order Butterworth filter). (Quiroga .6745 et al., 2004). After detecting positive- and negative-going threshold crossings, we took a 1.1 ms window centered on the threshold crossing and re-centered on the absolute minimum after interpolating by a factor of 8 using cubic splines. Features of the aligned spike windows were computed using principal

components analysis (Abeles and Goldstein, 1977). We fit a mixture of Gaussians model to the features using the Expectation Maximization algorithm (Dempster et al., 1977), and to detect the number of components in the mixture we fit models with 2-7 components, and chose the best model by minimum description length (Rissanen, 1978). We assessed the quality of clustering using signal-to-noise ratio (SNR), defined as the peak-to-peak voltage of mean spike waveform divided by the six times the standard deviation of the signal with all spikes removed. Only units exceeding a SNR of 1.1 were included for additional analysis (Ludwig et al., 2009). Additionally, inter-spike-interval histograms (ISIHS) were checked for refractory period violations, and the L-ratio and Isolation Distance were used to assess the quality of unit isolation (Schmitzer-Torbert et al., 2005). The spike-sorting analysis described above applies an accepted standard for sorted single units that results in signals of varying degrees of isolation. In addition to this analysis, we applied a more stringent analysis for unambiguous single units which required a minimum SNR of 2.8 and 0% ISI violations, which is sortable based on amplitude-threshold alone. For these rigorous single units, we also required stability of firing pattern. In what follows we report both on standard single units, or sorted multi-unit, and rigorous single units, making clear which criteria applies to each statement.

2.3.5 Sorted multi-unit and single unit stability using spike features and spike train statistics

A critical question in chronic studies is whether two different clusters on separate days represent the same neuron. To approach this problem quantitatively, we used methods developed in (Dickey et al., 2009; Fraser and Schwartz, 2012; Tolias et al., 2007). In brief, the similarity between mean waveforms and inter-spike-interval histograms were used as measures of the likelihood that clusters across recordings sessions represent the sorted multi-unit ensemble site. For waveform similarity, we

used the Fisher- transformed peak normalized cross-correlation; for ISIH similarity, we used the Jensen Shannon divergence (Lin, 1991), a commonly used probability distance measure. In HVC of an adult songbird, it is also possible to confirm recording stability by examining raster patterns across time. For neurons classified as stable by the methods described above, we found consistent firing patterns across time. This point is discussed in greater detail in the results section.

To verify the stability of rigorous single units, we continuously tracked the spike height, spike width, and change in firing pattern across bouts of singing. The change in firing pattern was assessed by computing the average smoothed instantaneous firing rate (Gaussian window, $\theta = 5$ ms) (Leonardo and Fee, 2005) over a 25 trial sliding window and taking the root-mean-square error between successive window averages. Synchronous changes in spike features and the firing pattern indicated that the single-unit signal had been lost or contaminated by other cells.

2.3.6 Imaging

Scanning Electron Microscopy images of the carbon fiber arrays were taken at the Boston University Photonics Center using a Zeiss Supra 55VP Field Emission Scanning Electron Microscope. Confocal images were taken with an Olympus 1X70 microscope. For fluorescence imaging of electrode tips, we coated UMS carbon fibers with Parylene- C containing 1% anthracene (141062 Sigma Aldrich) which fluoresces in ultraviolet light. Images were acquired using Olympus MagnaFIRE software and measurements of the length of the carbon fiber tip exposed from Parylene were made in Adobe Illustrator CS6 by merging the fluorescence with wide-field images.

Chapter 3

An open source, wireless capable miniature microscope system

This chapter is a reproduction of previously published work (Liberti et al., 2017), and contains critical contributions from *L. Nathan Perkins*, who is an equal co-author on this manuscript.

3.1 Introduction

Optical recording of neural activity in the brains of behaving animals has become an essential method in systems neuroscience. Through the use of genetically encoded calcium indicators, the principles of learning can be studied in large ensembles of neurons over timescales of weeks and months (Cai et al. 2016), (Ziv et al. 2013), (Liberti et al. 2016). An increasingly widespread and powerful method employs miniature head-mounted fluorescence microscopes to record cellular resolution activity in freely moving animals (Ghosh et al. 2011), (Barbera et al. 2016; Park et al. 2011). A variety of miniature head-mounted microscopes are available commercially, and the technique has been adopted by many labs (Betley et al. 2015), (Rubin et al. 2015), but these off-the-shelf devices currently lack a number of desirable features such as easy modification, wireless interfacing, color sensors, and flexible real-time analysis software. Some exciting custom built and/or open-source options have emerged (Cai et al. 2016), (Barbera et al. 2016), but the need remains for simple modular designs that use off the shelf parts and rapid prototyping tools such as a 3D printing. In our

particular application, none of the existing options were sufficiently lightweight for the small animal models employed in our lab. In response, we developed a flexible platform for miniature microscope design and fabrication that could address a variety of experimental needs. New features described in this project include an open-source 3D printed housing for easy experiment-specific reconfiguration, and wireless telemetry. For small animals such as juvenile mice or small songbirds that cannot easily carry the extra weight of the wireless transmitter and battery, we also describe a torque-sensing motorized commutator based on 3D printed parts and low cost hardware. A wired configuration connected through the commutator enables an ultralight configuration for recording (Fee and Leonardo 2001) .

In addition to the microscope, this project describes open-source control software capable of low-latency image processing and feedback for closed-loop experiments. Such experiments can trigger stimulation or other feedback in response to either patterns of recorded neural activity or external analog inputs. As a proof of concept, we demonstrate this capability by performing a closed-loop experiment in which auditory feedback is triggered with precise latency on particular syllables of a zebra finch’s song.

We intend for this project to act synergistically with other emerging open-source neurophotonic efforts, to bring these new tools to a larger user base and provide a platform for innovative optical recording techniques across the wider scientific community.

3.2 Methods

3.2.1 General Miniature Microscope Design

The microscope design was partially based on a previously described optical pathway Figure 3-1A (Ghosh et al. 2011). These miniature microscopes use a gradient

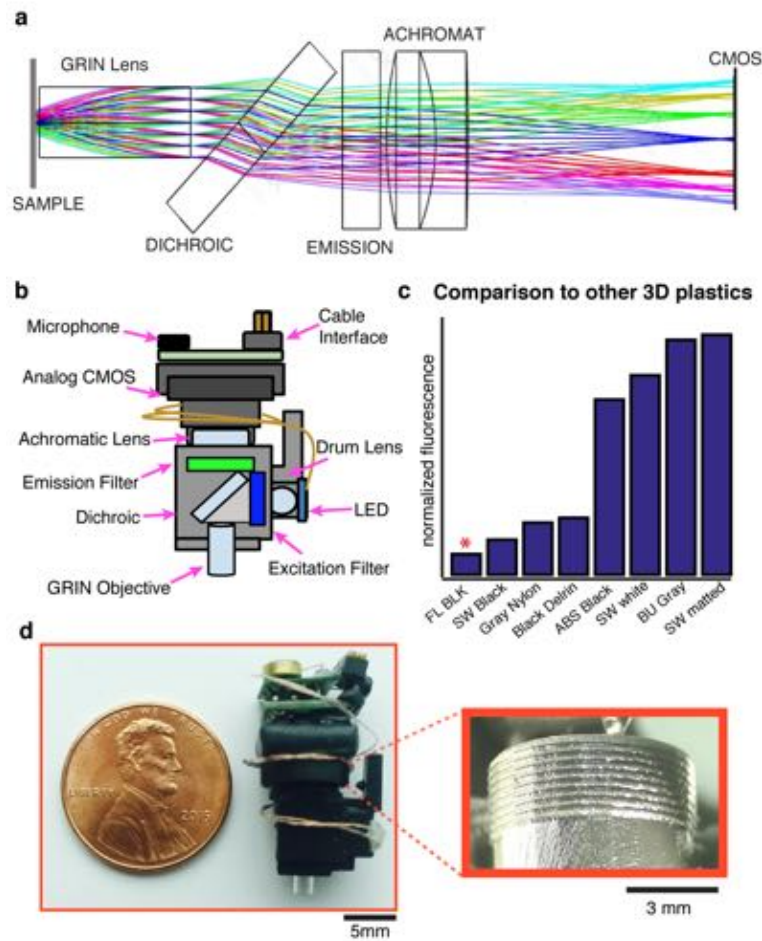


Figure 3-1: A custom 3D printed head-mounted fluorescence microscope. **A.** Optical layout of emission pathway for miniature microscope. **B.** Microscope schematic. Microscope body is lightweight and robust; CAD design is easily modified. **C.** A wide range of 3D printers and plastics were surveyed to maximize resolution and minimize autofluorescence. The red asterisk indicates our final choice of material: *Formlabs Black resin*. Autofluorescence of the current design is 1/2 the autofluorescence of black Delrin, one material used for machined microscope designs. **D.** Photograph of the microscope produced on a consumer grade 3D printer (Form 2), with inset showing the 3D printed focussing threads with a pitch of 0.34 mm.

refractive index (GRIN) lens as a high quality, short focal length objective lens. This design takes advantage of the off-axis focusing properties of these lenses, enabling fine focus adjustments post implant. We performed our original experiments with a green fluorescence indicator (GCaMP6) and chose our standard filters accordingly. A blue LED produces excitation light (470 nm peak, LUXEON Rebel), powered by a micro-

controller. A drum lens (Edmund Optics, NT45-549) directs the LED emission, which passes through a 4 mm x 4 mm excitation filter (Chroma, bandpass filter, 480/40 nm, 4 mm x 4 mm x 1.05 mm), deflects off a dichroic mirror (Chroma, 500 BS, 4 mm x 4.8 mm) and enters the imaging pathway via the gradient refractive index (GRIN) objective lens (GRINtech, GT-IFRL-200, 0.245 pitch length, 0.45 numerical aperture, or Edmund Optics 1.8mm Dia, 670 nm DWL, 0.23 pitch length, VIS Coated, GRIN Lens). Fluorescence from the sample returns through the objective, the dichroic, the emission filter (Chroma, bandpass filter, 535/50 nm, 4 mm x 4 mm x 1.05 mm) and an achromatic doublet lens (Edmund Optics, NT45-207, $f = 15, 12.5$ or 10 mm) that focuses the image onto the CMOS. This sensor is described in greater detail below. The frame rate of the camera is 30 Hz, and the field of view, when using a 15 mm achromat, is approximately $800\mu\text{m} \times 600\mu\text{m}$.

Various 3D printing resins were screened for their light blocking capacity, minimum print resolution, and auto-fluorescence in response to imaging wavelengths. We surveyed ten 3D printer and plastic combinations to find a low cost, high quality print option. We chose to build our microscope using a commercially available, consumer grade desktop 3D printer (Formlabs Form 2, resin type FGPBLK01 and FGPBLK02). The material has half the autofluorescence as Delrin in the green emission channel (480 nm excitation, 510 nm emission). Other materials surveyed include black acrylic-like resin (Shapeways), opaque acrylic-like resin (Shapeways), Black Delrin (DuPont), Black Nylon (Shapeways), ABS-black (Realize inc), ABS-P430 Ivory (UPrint) and VeroBlack Matte (Objet) Figure 1C. With the Form 2 printer, parts can be printed with extremely small feature size ($25\mu\text{m}$ layers, $140\mu\text{m}$ laser spot size), which allows for printing high-resolution threads used to adjust the focal length Figure 1D. The microscope housing is provided in the supplement, in the common STL format.

3.2.2 Software Design

Custom image acquisition software running on the macOS operating system leverages native AVFoundation frameworks to communicate with a USB frame grabber and capture a synchronized audio-video stream Figure 3.2. Video and audio are written to disk in MPEG-4 container files with video encoded at full resolution using either H.264 or lossless MJPEG Open DML codecs and audio encoded using the AAC codec with a 48 kHz sampling rate. Data on the regions of interest, fluorescence and triggered feedback are simultaneously written to a table (CSV) that includes corresponding frame numbers. In addition, the software communicates with a microcontroller (Arduino 2560 Mega) via a USB-to-serial connection. The microcontroller board allows the software to control LED brightness, and provides an interface with other analog and digital signals.

3.2.3 Near-real-time Software Details

The software is able to perform near real-time analysis on both the video stream as well as two independent analog channels. Sampled at 48 kHz, these can be used for analog inputs such as audio, TTL pulses, electrophysiology or data about behavioral context. These channels are used for low-latency data collection and triggering, with precise video alignment. For our songbird experiments, one of these channels is used for audio recording and song detection. In our experiments this has been particularly advantageous. We are able to activate the microscope's LED exclusively during periods when the animal is singing, restricting photobleaching and streamlining data collection. For this functionality, the audio stream is processed through a short-time Fourier transform to identify spectral content consistent with singing.

Either the analog input or the video stream can be the substrate for near-real time computational analysis. This allows feedback to be selectively triggered during

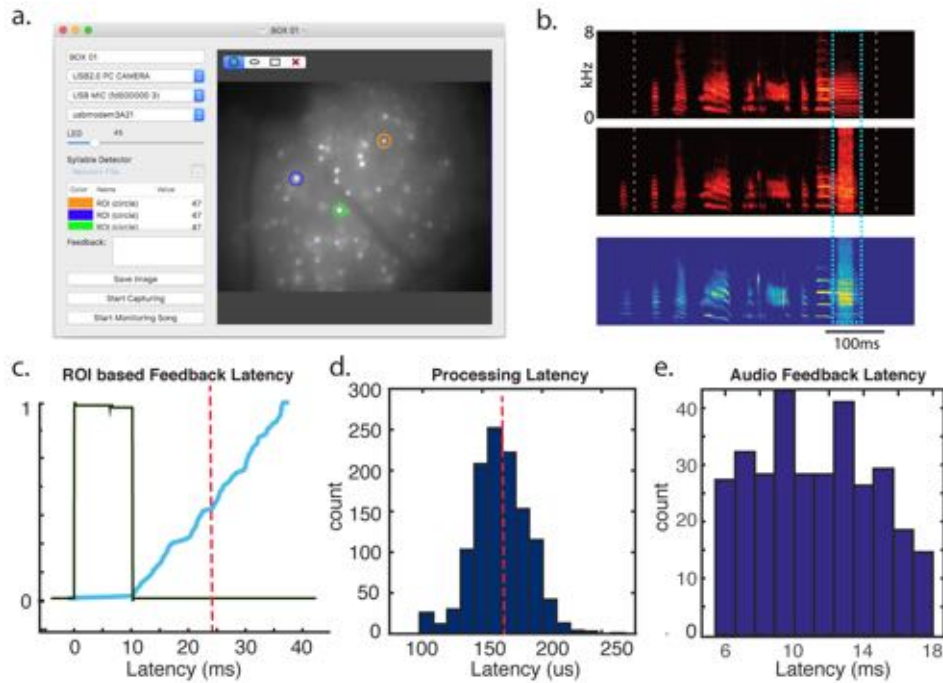


Figure 3.2: Performance of closed loop feedback based on near real-time audio or image processing using custom software and GUI. A. Image of user interface of acquisition software written in Swift for macOS. The software allows for low-latency ROI tracking and microscope control. It also interfaces with a microcontroller with 54 digital I/O pins, 16 8-bit analog inputs and an ADC with two tightly synchronized 16-bit 48 kHz analog inputs (typically for audio). **B.** Example of feedback contingent on features of the audio input: white noise is triggered at a specific syllable of song. White dotted lines mark a single motif of song, blue indicates the target syllable. *Top:* a 'catch' trial, where no feedback was delivered. *Middle:* a 'hit' trial, where a 50ms white noise pulse was delivered. *Bottom:* spectral density image of all song aligned trials (including hit and catch trials), demonstrating the reliability of the white noise pulse. **C.** Example of feedback contingent on ROI tracking. Black: voltage driving an LED light flash that is recorded in the field of the CMOS; blue: the cumulative probability density function (CDF) of a brief TTL pulse triggered by the software in response to the LED flash processed through the entire acquisition system. Event detection was based on ROI analysis on a Mac Mini computer. Latency of the full loop from camera to Arduino based TTL output is approximately $23.9 \text{ ms} \pm 7.9 \text{ ms}$ (95% confidence interval), with the jitter comparable to the frame rate of the camera. In this test, the LED was not synchronized to the onset of the frame, as would be the case for spontaneous video recording of neural activity. This represents the experimentally relevant performance of the system, intrinsically limited by the 33 ms frame rate of the camera. **D.** Of the total latency, image processing to extract fluorescence from ten cell-sized regions of interest contributes an average of only 0.17 ms; much of the ROI feedback latency is a reflection of the frame rate and acquisition time. **E.** Latency of auditory-based feedback, where a TTL pulse followed detections of a specific syllable structure (shown in B), had a latency of $12 \text{ ms} \pm 6 \text{ ms}$.

activity of interest on the video or analog channels. In the case of video-analysis, the uncompressed video stream is processed to extract fluorescence information for predefined regions of interest. Figure 3-2 shows an example of how the software can be used with a predefined triggering rule that applies feedback contingent on fluorescence signals measured in specific ROIs. In this paradigm, the camera can be used as a brain machine interface (BMI) in which songbirds control sounds directly through the measured calcium signals in the brain (Clancy et al. 2014). Other applications of the software could include closed-loop stimulation experiments that seek to electrically or optically disrupt patterns of activity in real-time.

Figure 3-2C depicts time between LED onset and feedback in our software test bed. The average time to trigger a feedback stimulus is $23.9 \text{ ms} \pm 7.9 \text{ ms}$ (95% confidence interval), including frame exposure, digitization, transmission, data acquisition and software processing. This latency reflects the intrinsic limitations of the 30 FPS (33 ms) frame rate as well as time spent capturing and decoding video, and communicating an output TTL with the Arduino controller through the serial port. Data acquisition, region of interest processing, and feedback triggering all occur within an average of 1.7 ms after frame capture (95% confidence: <11 ms). This variability stems primarily from video encoding and storage, where the codec and write buffering necessitate increased processing time for a subset of frames. An additional source of variability in processing time, which has a smaller impact and is shown Figure 3-2D, relates to the number and complexity of the defined regions of interest. As the number and size of the regions of interest increases, the processing time increases. In tests involving ten cell-sized regions of interest, the average processing time was 0.17 ms with a range of 0.11 ms - 0.22 ms (95% confidence interval).

3.2.4 Auditory Feedback

To test the acquisition pipeline for triggering on behavior, we used the integrated microphone on the head mounted microscopes to detect specific syllables of a bird’s song and deliver white noise auditory feedback at specific delays after the syllable was produced. White noise bursts are a common aversive stimulus for differential reinforcement in songbirds because it obscures normal feedback that the bird experiences (Andalman and Fee 2009), (Tumer and Brainard 2007). Custom Swift software was used for online detection of target syllables. White noise bursts were delivered at precise times relative to the target syllable during the bird’s vocalization. The duration of white noise bursts was long enough to overlap with the remainder of the targeted syllable (50 ms). This feedback was delivered with minimal trial to trial jitter, on the order of $12 \text{ ms} \pm 6 \text{ ms}$ latency from the target syllable, using a neural-network-based syllable detector (Pearre et al, 2017).

3.2.5 CMOS Sensor and Wireless Transmitter

The choice of CMOS imaging sensor is flexible. Benchtop systems and existing commercial solutions have employed a High Definition (HD) resolution sensor with on-board digitization in the microscope. However, due to out-of plane fluorescence, the resolution of single photon imaging in-vivo is typically much lower than the Nyquist sampling of a HD CMOS. At HD resolution, spatially oversampled images can result in a cumbersome large file size that hinders data processing on conventional lab computers; as a result, it is common for processing pipelines to start with spatial downsampling. Given the limited value in HD resolution for our application, we chose to start with a lower resolution sensor having only 640 x 480 pixels, a design choice that provided other benefits. Specifically, this lower resolution frame can be transmitted through a variety of low-cost, compact wireless transmitters Figure 3-3.

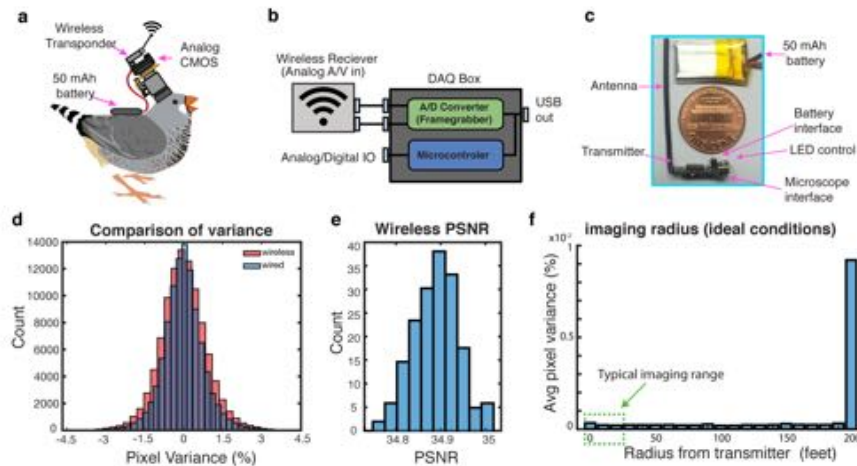


Figure 3-3: Wireless microscope acquisition system and performance. Signals from the camera can be relayed with an off-the-shelf wireless transmitter. **A.** Diagram of the system, mounted on a songbird. The microscope LED, CMOS and transmitter together draw approximately 100 mA, at the typical input voltage of 3.5 v and the typical in-vivo imaging LED brightness. **B.** The wireless acquisition system uses a wireless receiver, frame-grabber and digitizer to acquire synchronized video and two channels of 48 kHz audio. **C.** Image of wireless transmitter and 50 mAh battery. **D.** Comparison of pixel noise in the wired and wireless conditions. **E.** Histogram of per-frame PSNR values of the wireless condition, as compared to the wired condition (at 3 meters). **F.** Variance of the mean pixel value, over 100 continuously recorded frames (per distance) using a 100mW transmitter. High variance indicates signal degradation due to transmission loss or interference. Over distances relevant for typical neuroscience applications in an indoor laboratory setting, these transmitters are subject to minimal signal degradation.

The NTSC video protocol we use here is a mature technology incorporated into many low cost devices such as wireless spy cameras or miniature cameras for remote control airplanes. The same format also allows for analog transmission via flexible tethers for use in small songbirds or juvenile mice. Minimal degradation over distances used in realistic experiments Figure 3-3D,E and normal behavior of animals wearing the system (Supplemental Video 5) demonstrate that the system can be appropriately implemented in-vivo. However, the greatest utility will be in animals where a wired option is not possible or feasible, such as in behaving bats or crows, or where a data cable might restrict normal behavioral gestures.

The camera circuit employed for the microscopes in this study is an off-the-shelf

integrated color CMOS camera system (3rd eye electronics, MC900). It uses a 1/3 inch color CMOS (OV7960 TSV). Each on-chip pixel is 6 μm , and the signal to noise ratio (SNR) for the camera is 52 dB, with a sensitivity of 0.008 Lux. The camera circuit is also available with an integrated microphone (MC900A). The circuit draws less than 70 mA with an operating voltage of 3.3-5 v. A total of five wires run from the microscope: one wire each for camera power, ground, audio, NTSC analog video, and LED power. These wires run through a custom flex-PCB interconnect (Rigiflex) up to a custom-built active commutator (described in greater detail below). The NTSC video signal and analog audio are digitized through a USB frame-grabber. The integrated video frame-grabber and analog-to-digital converter (AV2USB V1.4, or DM420) converts two channels of 48 kHz analog data (in the case of songbirds, used for audio) and a synchronized NTSC video stream for reading by the host computer. While the color NTSC cameras provide wireless and color imaging, it potentially results in lower SNR in the GFP band. This may not be ideal for some applications; fortunately, the CMOS is easily changed for users that do not require color or wireless recording.

Off-the-shelf wireless transmitters are available for the NTSC format audio/video, and many weigh less than 0.6 grams (for example, BOSCAM TX24019, or many others) Figure 3-3C These transmitters perform reliably over distances of 5 meters or more, a measure confirmed in Figure 3E. The small scale and relatively low power requirements of the transmitter make them an attractive modification for untethered recording in freely moving mice or other organisms. In our tests, we powered the device with a lightweight, consumer grade Lithium Polymer (LiPo) battery. The choice of battery depends on weight constraints of the animal species: we have tested small, 1 gram 50 mAh batteries for functional use up to 30 minutes, and 3 gram 105 mAh batteries for up to an hour, at average imaging LED intensities. The microscope

weighs 1.8 grams and the 50 mAh battery and transmitter adds an additional 2 grams. This combination was burdensome for most zebra finches, although they are able to carry the complete system. The wireless system is ideal for larger animals, such as bats, rats or crows, that can carry the additional weight. We have had success in using this system in mice (unpublished observations).

3.2.6 Commutator Design

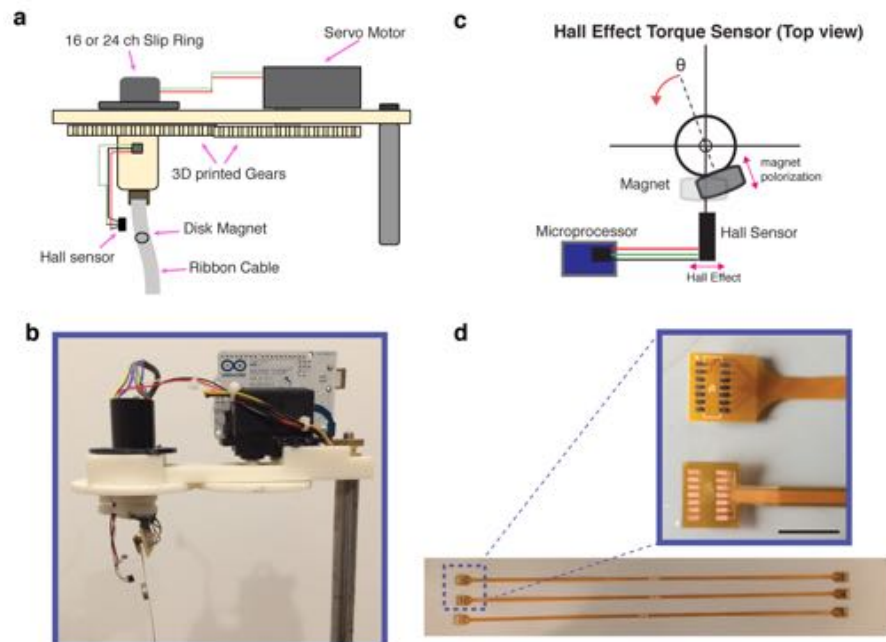


Figure 3-4: 3D printed active commutator system for chronic neural recording in small animals. **A.** Schematic of the 3D printed commutator. **B.** An image of an assembled commutator. **C.** These devices use the deflection of the magnetic field of a disk magnet located on a flex PCB cable to detect torque via a hall sensor. A feedback circuit mediated by a microcontroller corrects the deflection by rotating a slip ring via a servo-driven gearbox with a 1:1 ratio. **D.** Example of two different flex cables designs with 7-9 conductors, weighing under 0.25 grams. The additional wires are present for electrical/optical stimulation or other head-mounted accessories. Scale bar indicates 9 mm.

To image in songbirds, which were unable to carry the extra weight of the battery and wireless transmitter, we developed a torque-sensing 3D printed commutator (Figure 4A), as well as lightweight (0.1 g) low-noise, flexible PCB cables. The com-

mutator was loosely inspired by previous designs (Fee and Leonardo 2001). These commutators allowed for 24/7 longitudinal recording without battery replacement or frequent handling of the animals. The ability to record longitudinally from animals may prove useful in larger animals, such as mice, as it avoids the complexity of battery management and charging. The active commutator consists of a gear assembly (ABS-P430 Ivory, UPrint), along with seven electronic components that can be purchased for a total cost of under \$100. These additional components include a servo and a resistor, a slip ring commutator, a disk magnet and a hall sensor Figure 4C. The mechanical designs, editable stereolithography designs, and software to operate the commutator is included in the supplement.

3.2.7 Surgical Procedure

To label neurons in the song-related premotor nucleus HVC with a calcium indicator, birds received three 250 nl injections of lentivirus packaged with either GCaMP6s or GCaMP6f under a Rous sarcoma virus (RSV) promoter into the song premotor cortex. Virus is prepared as described previously (Liberti et al. 2016), and constructs are available in Addgene. To guide the injection of virus, the boundaries of HVC were determined by fluorescence targeting of a DiI retrograde tracer injected into the downstream song motor nucleus Area X.

3.2.8 Song Alignments

For all neural recording modalities, trials were aligned to song using previously described methods (Poole et al. 2012), using the Euclidean distance in spectral features between the data and a template song in a sliding window. Local minima in the Euclidean distance were considered candidate hits, which were then plotted in 2 dimensions for the user to perform a cluster cut. No time warping was applied to any data (Glaze and Troyer 2006).

3.2.9 Calcium Imaging ROI Analysis

Calcium imaging data was analyzed as described previously (Markowitz et al., 2015). In brief, raw imaging data was motion corrected using a previously published algorithm (Guizar-Sicairos et al. 2008). Then, regions of interest (ROIs) were manually selected, and for each frame, pixels intensities were averaged for all pixels in the region. ROI traces were converted to $\Delta F/F_0$ by estimating F_0 as the 12th percentile in an 800 ms sliding window.

3.2.10 Wireless imaging quality quantification

Wireless imaging data quality was assessed by splitting the analog NTSC video signal into two paths: one to a wired frame-grabber, and the other through a wireless transmitter and receiver, to a second frame-grabber. This allows for explicit evaluation of the image quality for the same exact video stream, enabling frame-by-frame comparison of the wireless and wired conditions. The differences between frames encompasses both signal degradation due to wireless encoding, and noise introduced due to interference from other devices, giving a practical measurement of signal degradation. These experiments were performed in realistic lab conditions with a clear path from transmitter to receiver, but no additional steps were taken to mitigate potential sources of noise.

3.3 Results

3.3.1 Optical Recording of Neurons Expressing Genetically Encoded Calcium Indicators

To provide information about the stability of excitatory cells, we used our miniature microscopes to perform optical imaging of genetically encoded calcium indicators Figure 3.5 (Markowitz et al. 2015), (Liberti et al. 2016). Electrophysiology methods are

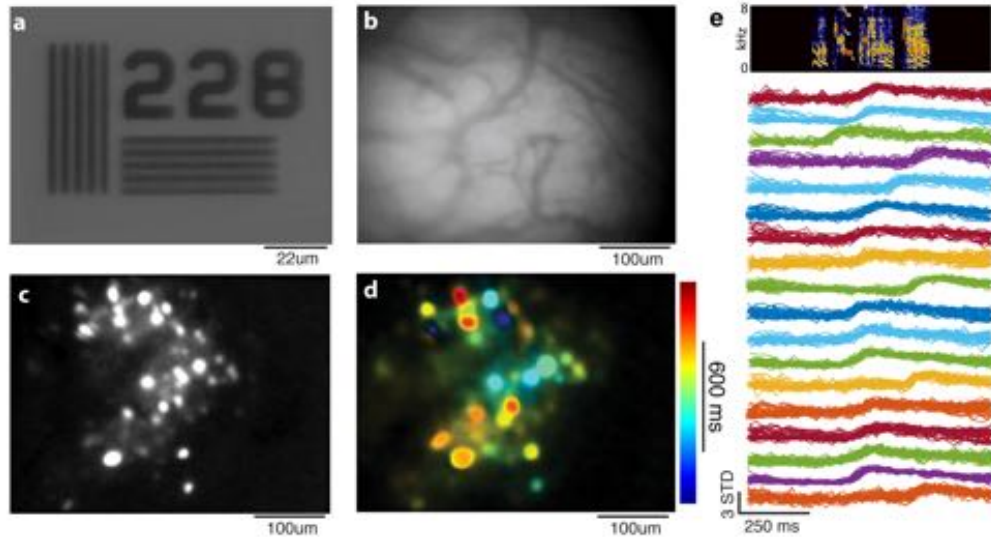


Figure 3-5: Images and in-vivo video collected from the microscope. **A.** Image taken by microscope of a High-Frequency NBS 1963A Resolution Test Target, showing 228 lines per mm. **B.** In-vivo widefield image of blood vessels over premotor area HVC in a zebra finch. **C.** Maximum intensity projection of $\Delta F/F_0$ video from a bout of singing. Imaging depth is 150-200 μm below the surface of intact dura. **D.** Time-intensity plot, where each pixel is colored by its center of mass in time. **E.** Stereotyped single neuron calcium traces recorded in singing birds using GCaMP6, aligned to song. *textitTop*: spectrogram of a single song rendition; *textitBottom*: calcium traces from 18 ROIs over 50 song-aligned trials from a single bird. Vertical scale bar indicates standard deviation.

typically unable to track individual neurons over long time periods. Excitatory projection cells in this region are extremely difficult to continuously record for timescales longer than a single day, perhaps due to limitations of recording from smaller cells (Guitchounts et al. 2013). As a proof of concept, we optically recorded the fluorescence transients from neurons expressing the calcium indicator GCaMP6 in the songbird premotor cortex HVC. Consistent with previous extracellular recordings and calcium imaging studies in HVC, projection neuron calcium activity patterns were highly stereotyped and stable across song trials within a day of singing Figure 3-5A. Microvasculature can be clearly seen above HVC on the surface Figure 3-5B. Cells were found that produced stereotyped calcium transients at all time-points within song, and their spatiotemporal organization was defined by a 100 μm length-scale

clustering (Markowitz et al. 2015). Within a song rendition, dozens of cells can be recorded simultaneously. Because these microscopes are lightweight and inexpensive, they can be chronically implanted and left in place for weeks at a time. This provides a paradigm for stable longitudinal recordings.

3.4 Discussion

Cellular resolution optical imaging in behaving animals is a foundational method in modern neuroscience; allowing researchers to longitudinally track cells in sparsely active networks like the songbird premotor region HVC and the rodent hippocampus with high spatial resolution. Through the use of genetically encoded calcium indicators (Chen et al. 2013), (Dana et al. 2016), the principles of learning, information encoding, etc, can be studied in large ensembles of neurons at cellular resolution, over periods of weeks and months. Typically, optical experiments utilize either benchtop two-photon imaging systems in head-restrained animals (Dombeck, 2010), (Minderer et al. 2016), (Rickgauer et al. 2014), or single photon imaging in freely moving animals through the use of a head-mounted miniature fluorescence (single photon) microscope (Cai et al. 2016), (Ghosh et al. 2011), (Barbera et al. 2016, Park et al. 2011). While the axial resolution of multiphoton microscopy is vastly superior (Helmchen and Denk 2005), head-mounted microscopes are often the only way to optically observe neural populations during naturalistic behaviors (Resendez et al. 2016).

The motivation for this project was that existing commercially-available miniature microscopes proved too heavy to consistently evoke undirected song, a learning-intensive motor behavior in zebra finches. With extensive screening and training of birds, it is possible to evoke song in a head-fixed preparation in the presence of a mate (Picardo et al. 2016), but this approach can be low yield since few birds will sing head-fixed, and the method may preclude the study of the mechanisms of motor

maintenance that occur during undirected singing in the absence of a female. More generally, there are many applications in neuroscience where even a wire tether may restrict natural behavior and prevent interrogation of the underlying neural mechanisms (Wiltschko et al. 2015), (Yartsev and Ulanovsky 2013). Songbirds wearing commercially available microscopes rarely sing; possibly due to the microscope weight or bulky cables used to stream data from the CMOS. With the torque sensing commutator and lightweight tether, our present design is sufficiently unobtrusive for zebra finches; we routinely gather 400-1,000 song trials per day. This microscope system, in conjunction with behaviorally triggered acquisition has allowed us to perform around-the-clock studies of brain activity in a substantial number of animals without constant human supervision, enabling us to gather densely sampled longitudinal recordings. The tool will enable further studies of learning at cellular resolution in zebra finches (Figure 3·5). This development history and outcome underscores a limitation of commercial solutions: while off-the-shelf versions may provide adequate functionality for some species or experimental paradigms, many experiments require user-defined customization. It is likely that modifications described here, such as wireless interfacing, custom filter sets, and color sensors, as well as the modular design of this miniature microscope will inspire innovation and enable data collection in new experimental models or species.

3.4.1 Advantage of 3D Printed Housing for Rapid Microscope Assembly and Prototyping

Existing miniature microscope designs require assembly of carefully machined parts, typically milled from Peek or Delrin plastic (Cai et al. 2016), (Ghosh et al. 2011). These machined parts have the advantage of precise tolerances, but the disadvantage of higher cost and longer design timescales relative to 3D printed materials. 3D printed microscopes can be produced at low cost with geometries that are not

possible with CNC milling, we take advantage of a single-piece design that reduces weight by eliminating metal bolts and allowing thin walls. Because 3D printed parts avoid the constraints of machined parts, these microscopes can be lighter, more easily constructed, and readily reconfigurable to accommodate design variations.

3.4.2 All Optical Physiology with Multi-Channel Light Delivery and Recording

While we ran initial tests with filters optimized for imaging GCaMP6, different filter sets can be used to accommodate newer, red-shifted indicators. With some modifications, these cameras can be adapted to incorporate two independent color channels at the expense of SNR within each channel. Optimization of filter sets and incorporating multi-wavelength spectral peak LEDs can provide additional information about cells in the imaging plane, using fluorescent proteins or tracers with well-isolated spectral profiles to differentially label subpopulations of cells. Examples include GCaMP6 combined with infrared Alexa dyes for retrograde labelling. This additional channel can allow disambiguation of specific neuron types within an imaging field. Alternatively, adding a second pass to the dichroic and excitation filter can allow full field optical stimulation outside of the imaging bands.

3.4.3 Low Latency, Open Source Acquisition Software

In our tests, the latency of triggering off fluorescence activity was limited by the framerate of the microscope, about $23.9 \text{ ms} \pm 7.9 \text{ ms}$ Figure 3-2C. For many experiments, this response latency is acceptable. In the songbird, for example, the sensory-motor latency from pre-motor neuron activity in HVC to auditory sensory consequences processed in the basal ganglia is a minimum of 32-50 ms (Andalman and Fee 2009). Based on these estimates and published accounts of conditional feedback experiments in songbirds, the time-delay should provide a learnable brain machine

interface (Oliveczky et al. 2005), (Tumer and Brainard 2007), (Sakata and Brainard 2008), (Sober and Brainard 2009). However it remains to be seen whether the timing jitter in the current system is acceptable for brain machine interface experiments in a system as precise as the songbird. For the zebra finch, relative jitter between premotor commands and auditory feedback is just a few milliseconds. For some experiments lower latency and lower variability may be desirable, and with the introduction of faster frame rate cameras and deterministic real-time operating systems, this latency and latency variability will decrease.

3.4.4 CMOS Trade-Offs and Deficits

In our tests, we use an off-the-shelf CMOS that outputs analog video. In terms of resolution, shot noise and frame rate, the NTSC analog camera is less versatile than many new digital sensors. These sensors do have the advantage of being easily broadcast wirelessly with high signal fidelity and low power consumption, which allows real-time data streaming for wireless BMI experiments, and also allows adjustment of imaging parameters on-the-fly. These sensors have been adequate for calcium imaging experiments- although other experiments may benefit from higher resolution digital sensors (Deisseroth and Schnitzer 2013), (Cai et al. 2016).

3.4.5 Use of Off-the-Shelf Components for Ease of Construction

Designing and constructing optical equipment poses engineering challenges: many researchers and educators, especially in resource-limited circumstances, may be intimidated by constructing their own systems. This can especially be the case with custom ASIC or chip design requiring considerable technical skill. With this in mind, we have aimed to present a design that is built from low cost, off-the-shelf components. The design requires minimal maintenance, enabling longitudinal use. This microscope design will allow labs with little electronics experience to enter the field

of awake behaving imaging and to build their own microscopes, while providing room for electronics savvy experimentalists to iterate and develop novel imaging and acquisition back ends.

The future development of these devices relies greatly on a combined multidisciplinary effort, involving biological, chemical, mechanical, and materials engineering, among others. In particular, neuroscience stands to benefit from advances in consumer grade electronics. The cost, availability, size and quality of electronics used in this project has been driven by cellular and telecommunications industries (Deisseroth and Schnitzer 2013). This sector will likely continue to drive rapid innovation in miniaturization of electronics, nanoscale 3D printed components (Sun and Kawata 2004) and optics (Gissibl et al. 2016), and high-fidelity wireless technology- all of which stand to increase the quality of neuroscience instrumentation. We hope that the microscope presented here will be further enhanced by these efforts and allow maximum integration with other emerging open-source neurophotonics projects.

Part II

Network Level Description of HVC

Chapter 4

Mesoscopic Patterns of Neural Activity Support Songbird Cortical Sequences

This chapter is a reproduction of previously published work (Markowitz et al., 2015), and contains critical contributions from *Dr. Jeffrey Markowitz*, who is an equal co-author on this manuscript.

4.1 Introduction

For many sensory systems, features of sensory inputs are represented topographically in the brain. Examples include orientation tuning in visual cortex, pitch tuning in auditory cortex, and the representation of touch in somatosensory cortex. Motor cortical circuits also contain maps for distinct muscle groups and preferred direction of movement (Georgopoulos et al., 2007; Naselaris et al., 2006), but it is not known whether higher-order features such as time in a movement sequence or serial order are represented topographically. The specific question of how time is mapped in space is particularly tractable to address in the songbird premotor nucleus HVC (used as a proper name), which is known for producing extremely precise, learned temporal sequences. Projection neurons in HVC fire sparse bursts of action potentials that are precisely aligned to singing. One subtype (HVC_{RA}) drives downstream areas controlling the vocal organ (Hahnloser et al., 2002a), while another (HVC_X) projects to the basal ganglia, which is hypothesized to guide exploratory trial and error learning (Fee and Goldberg, 2011a). These cells can be compared with a larger class of 'time

cells' observed in rodents that fire at specific moments during a stereotyped behavior, and may be involved in motor sequence generation, navigational planning or episodic memory (Crowe et al., 2014; Harvey et al., 2012; Long et al., 2010; MacDonald et al., 2011; Merchant et al., 2013; Pastalkova et al., 2008). In addition to these sparsely firing cells, HVC contains inhibitory interneurons that produce dense, stereotyped firing patterns during singing (Guitchounts et al., 2013).

Neither the interaction between inhibitory neurons and projections neurons nor the spatiotemporal organization of HVC activity have been observed in singing birds. In spite of decades of detailed electrophysiology, these aspects of song coding have been impossible to address since imaging and high channel count multi-electrode physiology in singing birds have been technically infeasible until recently. In parallel with these experimental limitations, dynamical models of the HVC network have generally been framed in terms of the neuronal connectivity that supports sequence generation, but the spatial organization of cellular activity has been considered irrelevant (Li and Greenside, 2006; Long et al., 2010). However, a number of experimental studies have suggested that HVC activity is spatially organized. Multi-unit activity in HVC is strongly modulated on a 30-100 ms timescale in singing birds (Crandall et al., 2007; Schmidt, 2003), which could suggest correlated activity in nearby cells. In anesthetized animals, calcium imaging has revealed that spontaneous firing in HVC_X neurons is correlated over space ($\sigma=263 \mu\text{m}$) (Graber et al., 2013). Retrograde tracer injections and stimulation experiments in HVC show a high degree of lateral connectivity along the rostrocaudal axis (Nottebohm et al., 1982; Stauffer et al., 2012). Newborn HVC_{RA} projection neurons form soma-soma contacts with existing HVC_X neurons at the moment of integration into the song circuit (Scott et al., 2012). Finally, multi-electrode recordings in anesthetized birds demonstrated a spatial anisotropy in spike-spike coherence between different cell types: projection neurons were correlated

in the rostrocaudal direction and inhibitory neurons in the mediolateral direction (Day et al., 2013). Taken together, these studies suggest that HVC may represent time with some degree of spatial organization, but direct measurements in singing birds are lacking.

In addition to ignoring spatial correlations, models of HVC have largely ignored temporal correlations between inhibition and excitation (Li and Greenside, 2006; Long et al., 2010) (though some exceptions can be found (Bertram et al., 2014; Gibb et al., 2009)). Yet, excitatory and inhibitory neurons are known to be interconnected in HVC (Mooney, 2000; Mooney and Prather, 2005b), HVC_X neurons fire bursts of rebound spikes after inhibition in slice (Daou et al., 2013), and HVC_X spiking is anti-correlated with inhibitory neurons in anesthetized songbirds (Day and Nick, 2013). These observed correlations do not necessarily apply to the singing state. Indeed, in singing birds there appears to be a high diversity of interneuron firing patterns (Kozhevnikov and Fee, 2007b), and this diversity suggests that net inhibition on any given projection neuron could be a tonic signal that is relatively unmodulated in time. However, if inhibitory inputs to projection neurons contain synchronous bursts or pauses in firing, the impact on projection neuron firing times could be profound. As a result, the question of how excitatory and inhibitory cells interact in HVC is intrinsically related to the question of whether or not there is a spatial organization or synchrony among neighboring cells in the HVC microcircuit.

Here we provide what is, to our knowledge, the first experimental study of how ensemble activity in HVC is organized in space and time in singing zebra finches. To do this, we developed new methods for multi-channel electrophysiology and calcium imaging in singing birds. We address activity patterns of projection neurons and inhibitory interneurons, and investigate the rules that relate the two cell classes together. First, we frame the question of how HVC activity is spatio-temporally or-

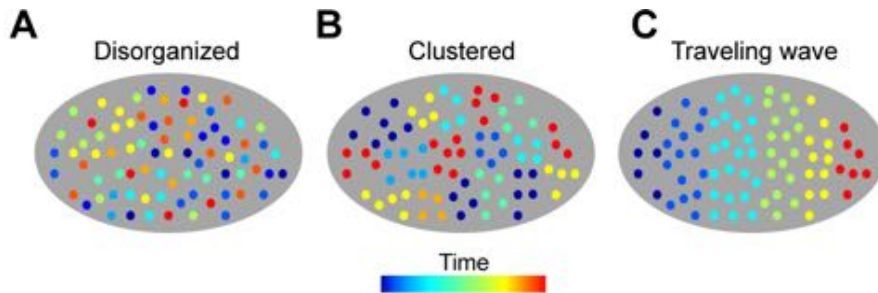


Figure 4-1: Three hypothetical models for the spatiotemporal organization of the song premotor code in HVC. The spatial organization of neural activity in HVC in singing birds is unknown. The geometry of neural activity could be described by three schematics that form a continuum: **a** a random, disorganized geometry (Harvey et al., 2012); **b** functional clustering (i.e. nearby cells code for similar elements) with a characteristic length-scale (Dombeck et al., 2009); and **c**, traveling waves (Rubino et al., 2006).

ganized in terms of three simple models that have been championed in various species (Dombeck et al., 2009; Harvey et al., 2012; Rubino et al., 2006) (Figure 4-1). The schematic models range from a random (Figure 1a) to a globally organized geometry (Figure 4-1c). Through calcium imaging and multi-electrode recordings in singing birds, we find that activity patterns of excitatory and inhibitory neurons are spatially clustered (resembling Figure 4-1b), and that two cell types, HVC_X projection neurons and HVC interneurons, fire in alternating phases of a local 30 Hz rhythm. Our results show that there is a fundamental mesoscopic length-scale and time-scale in the premotor cortical area HVC. Similar length-scales have recently been reported in calcium imaging of cortex in behaving rodents (Dombeck et al., 2009; Hira et al., 2013; Peters et al., 2014), and similar time-scales have been implicated in primate motor control (Bartolo et al., 2014; Murthy and Fetzi, 1992; Rubino et al., 2006) and in human speech (Giraud and Poeppel, 2012), suggesting that these may be fundamental properties of motor cortical micro-circuits across species.

4.2 Results

4.2.1 The representation of time in HVC projection neurons is spatially correlated over a 100 μm length scale

To examine the spatial organization of calcium activity in HVC we employed head-mounted fluorescence microscopes (Figure 4-2a) to image the genetically encoded calcium indicator GCaMP6s in freely behaving, singing adult male zebra finches (n=9 birds). The field of view for these recordings included a large fraction of the surface of HVC, but the lentiviral infections were sparse, limiting our results to observations of n=167 cells across these 9 birds (Figure 4-2b)

GCaMP6s provided robust calcium signals in single neurons. The observed calcium transients were time-locked and temporally sparse, consistent with the ultra-sparse, time-locked firing patterns reported previously in electrophysiological recordings (Hahnloser et al., 2002a) (Fig 2b-d), and also consistent with evidence that the high frequency bursts of projection neurons ride on top of calcium plateau potentials Long:2010db, but we can make no claim that the calcium and electrophysiological signals are equivalent. New technical developments will be needed to directly measure the correspondence between action potentials and calcium in GCaMP6s infected cells in freely moving birds.

Within this population of sparse firing projection neurons, we found strong spatial correlations in the calcium activity of neighboring neurons (Figure 4-2e-f). The median separation between co-active cells (calcium events times separated by less than 20 ms) was 143 μm (119-183 μm , 95% bootstrap confidence interval). Cells that fired at times separated by more than 50 ms showed no spatial correlations with one another ($p > .05$, bootstrap test; see Materials and Methods). These observations were based on the time between sparse calcium events rather than the similarity of the full time series (e.g. cross-correlation), and are therefore robust to contamination from

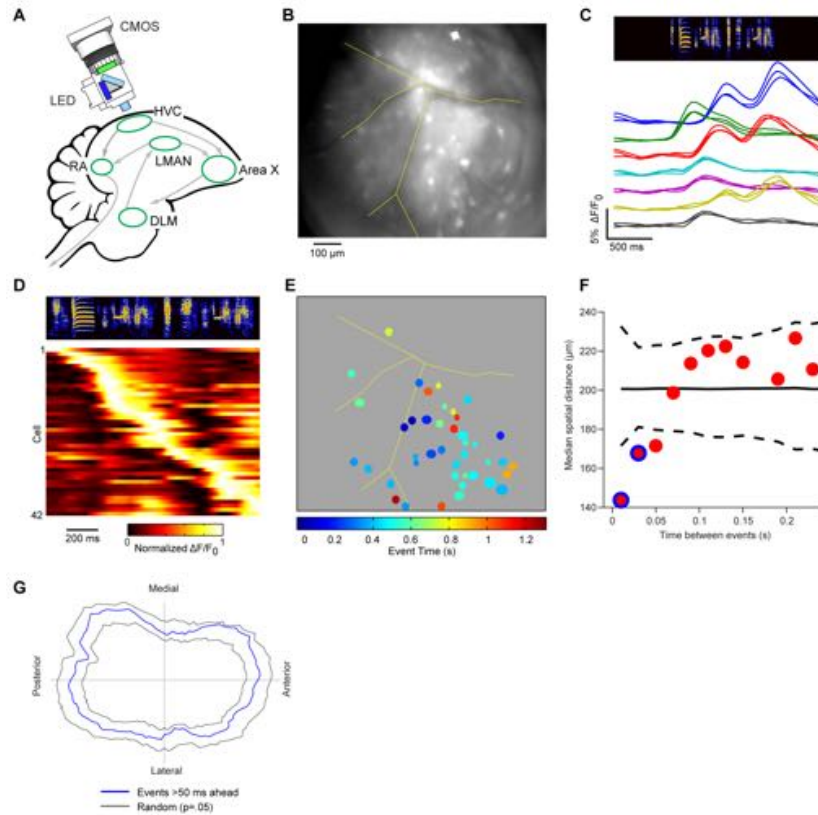


Figure 4-2: Calcium imaging of HVC projection neurons in awake, behaving birds.

a, Schematic of the zebra finch brain, showing a simplified song circuit and miniature fluorescence microscope (Ghosh et al., 2011). **b**, Field of view showing a mean fluorescence image of HVC (blood vessels are traced in yellow). **c**, Several example ROIs from 3 trials during a single day of imaging in the same bird, aligned to song. **d**, Peak-normalized trial-averaged activity from all song-related neurons from one animal sorted from earliest to latest peak. **e**, ROIs are colored to represent the time of cell firing within song, defined by the time of 50% rise in the onset of the calcium transient (see Materials and Methods). **f**, The median distance between ROIs is shown as a function of the time between their respective calcium events ($n=2230$ ROI pairs). This reveals that co-active ROIs (difference between event times <50 ms) are significantly closer to each other than if event times were randomly distributed across HVC ($p=0$ and $p=.038$ for the first two points respectively, bootstrap test, Bonferonni corrected). Dotted lines indicate the 95% confidence interval for the null model (see Materials and Methods), and the solid line indicates the median. Blue highlighting indicates $p<.05$. **g**, Histogram of relative distances between each ROI and all ROIs with calcium events >50 ms ahead ($n=1710$ ROI pairs). The distances are not significantly concentrated relative to the null model described in **f** ($p=.6924$, bootstrap test). The 95% confidence interval for the histogram of the null model is indicated by dashed lines. The asymmetry of the null model is a consequence of injecting virus at multiple sites distributed along the AP axis of HVC. (The population of recorded cells was presumably elongated in the AP axis as a result.)

out-of plane fluorescence or neuropil (see Materials and Methods).

Tests for a traveling wave structure in the data revealed no trend across animals along either the mediolateral or anterior-posterior axis ($p=.6924$, bootstrap test; Figure 4.2g). In contrast to a geometrically organized traveling wave (Figure 4.1c) (Rubino et al., 2006), the spatial correlations in HVC calcium activity appear to be characterized by a patchwork of domains (Figure 4.1b) organized in stereotyped sequences. These spatiotemporal 'domain sequences' do not contain any global geometric organization that we have yet detected. However, we did find a bias for co-active cells to be clustered along the anterior-posterior axis, similar to previous anatomical (Nottebohm et al., 1982; Stauffer et al., 2012) and physiological (Day et al., 2013) observations in anesthetized birds .

4.2.2 Inhibitory neuron activity is reflected in a 30 Hz local field potential and this field potential is also correlated over a 100 μm length scale

We next asked whether other measures of activity in HVC are correlated over a length scale comparable to the correlation length measured in the calcium activity. Simultaneous electrophysiological recordings of interneurons separated by defined spatial distances is not currently feasible in HVC, and the time-scale of the calcium indicator used here could not resolve the fast interneuron firing patterns previously described in HVC. Faced with these limitations, we turned to the local field potential (LFP), which can reflect synchronous ensemble neural activity over roughly 100 μm (Katzner et al., 2009; Xing et al., 2009), though the exact volume of activity represented by the LFP is under debate (Buzsáki et al., 2012). In some cases, the LFP also carries timing information about synchronous inhibition (Hasenstaub et al., 2005; Nir et al., 2007). To record both LFPs and multiple single-units, we developed an ultra-small carbon fiber electrode array capable of recording 16 channels of LFP and spiking data in singing birds (Guitchounts et al., 2013). We recorded 65 putative interneurons, 19

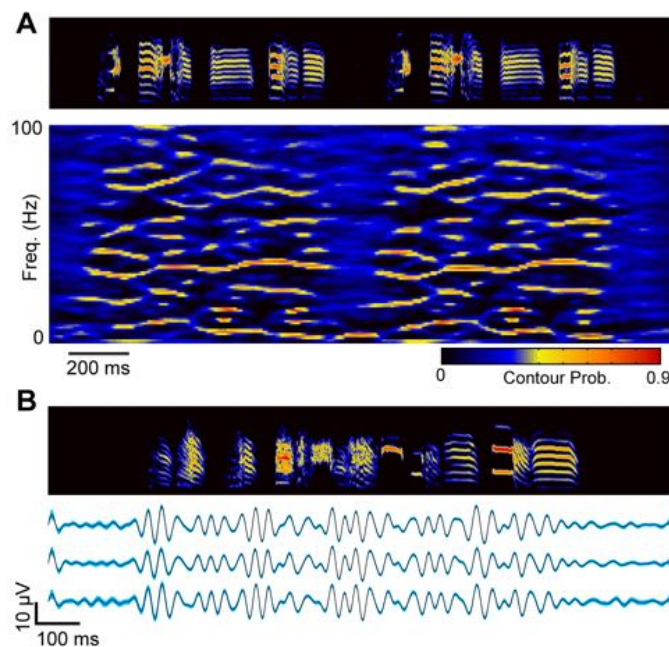


Figure 4-3: Stereotyped network activity during vocal production in songbird pre-motor cortex area HVC. **a**, Spectral density images (Markowitz et al., 2013b) reveal the time-frequency structure in the LFP during singing ($n=64$ song-aligned trials. No time warping was applied). In a spectral density image, color indicates the probability density of time-frequency structure (see Materials and Methods). Reliable time-frequency structure emerges as the bird begins to sing and disappears during the short gap between song motifs. **b**, Trial-averaged band-passed LFPs (25-35 Hz) from a single electrode on 3 consecutive days ($n=130$ trials for the top trace, and $n=200$ trials for the middle and bottom traces). Shading indicates 99% bootstrap confidence interval.

projection neurons, and 268 distinct LFP sites from HVC in adult male zebra finches ($n=27$ birds).

To date, the LFP has not been described in HVC in singing birds. We found that stereotyped LFP structure appears in the 10-100 Hz frequency range as song begins, and rapidly disappears during inter-motif gaps in sound and at the end of song (Figure 4-3a). In Figure 4-3b we show the stability of the trial-averaged LFP in the 25-35 Hz frequency band. This particular frequency band is the most phase-locked to song behavior (Figure 4-4a. In support of the importance of the 30 Hz frequency band in the LFP, a strong correlation was found between inhibitory neuron firing times and LFP phase at 30 Hz (Figure 4-4b; cross-spectrum analysis, $p=0$, bootstrap test).

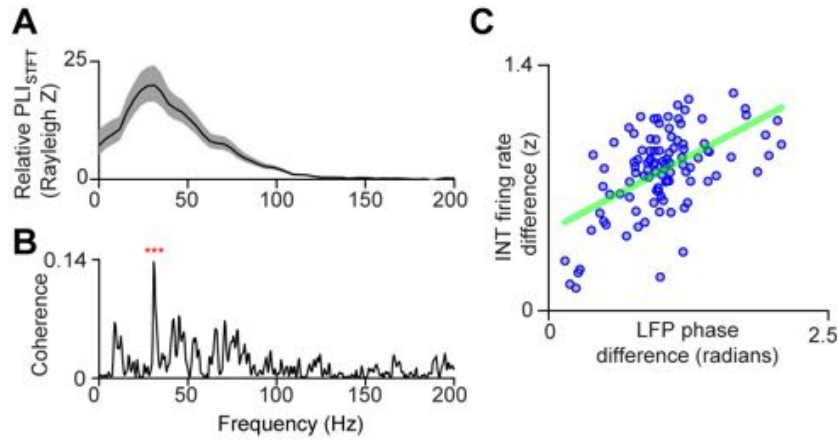


Figure 4-4: The phase of the 25-35 Hz LFP can be used to study the spatial structure of inhibition. **a**, The average change in phase locking for all birds (, see Materials and Methods) between song and awake quiescence indicates that the most stereotyped frequency band in the LFP is centered on 30 Hz. **b**, The magnitude squared coherence between LFPs and interneurons ($n=17$, see Materials and Methods) is highly significant at 30 Hz ($p=0$, bootstrap test, Bonferonni corrected). **c**, For a given pair of recording sites ($n=79$), phase shifts in the 25-35 Hz LFP predict interneuron firing pattern distance ($r=.48$, $p=8.5e-6$, Pearson correlation coefficient).

When two spatially separated electrodes in HVC show similar LFP patterns in this 25-35 Hz band, their local interneuron activity is also highly correlated (Figure 4-4c; $r=.48$, $p=8.5e-6$, Pearson correlation coefficient). Hence, the 25-35 Hz LFP can be used as a surrogate to study the spatial structure of interneuron firing.

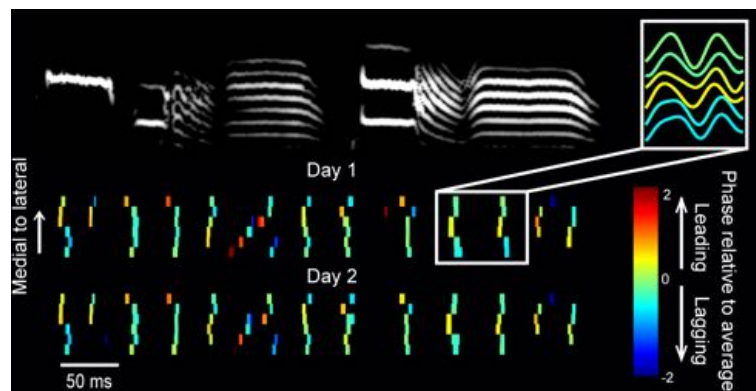


Figure 4-5: A stereotyped spatiotemporal 30 Hz pattern underlies premotor cortical activity during song. Trial-averaged 25-35 Hz LFP patterns on two consecutive days for six electrodes spaced by $175 \mu\text{m}$ along the mediolateral axis of HVC. Each tick mark indicates the timing of the local LFP peak, and the color indicates the phase relative to the average phase of all electrodes at that time point. (See breakout illustration top right).

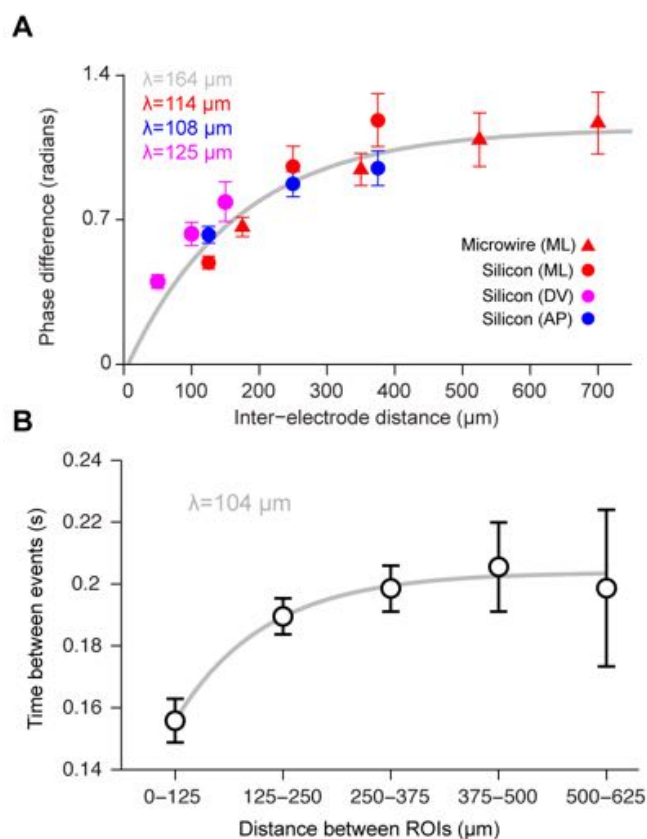


Figure 4-6: LFPs in the 25-35 Hz band and calcium events are correlated over a 100 μm length scale. **a**, Shown is the average phase difference between LFPs recorded on two separate electrodes as a function of distance along the dorsoventral (DV) ($n=70$), mediolateral (ML) ($n=57$), and anterior-posterior (AP) axes ($n=48$). This data was collected using four-shank silicon probes (Neuronexus) and commercial microwire probes (TDT). **b**, Average difference in event times between ROIs observed using calcium imaging. The equation was fit to data in both **a** and **b** using a least squares procedure. In **a**, the data from each of the axes was separately fit (matching colors), the gray line represents the fit using data combined from the AP and ML axes. Error bars represent SEM.

To characterize the length scale of the LFP correlation across the surface of HVC, we first implanted 4 four-electrode bundles of carbon fiber electrodes ($n=5$ birds), separated by approximately 200 μm in both the rostrocaudal and mediolateral axes of HVC. LFPs recorded from electrodes in the same bundle were highly similar, while those from different bundles revealed phases that were shifted by as much as 180° at some points in song ($p=5.2\text{e-}11$, $z=-6.57$, two-tailed Wilcoxon rank sum test). We also examined the spatial extent of LFP correlations using silicon and microwire arrays

with larger spacing and well-defined geometries. Figure 4.5 reveals the details of the spatiotemporal LFP pattern for one bird, measured at 6 points along the mediolateral axis of HVC. The syllable-specific microstructure in the LFP phase is not just noise; the structure can be seen to repeat precisely from one day to the next as the bird sings his stereotyped song. Averaging over all recording sites in all birds, we found that the 25-35 Hz LFP has a length scale of 108-125 μm along the dorsoventral, mediolateral, and anterior-posterior axes (exponential fit to the difference in phase as a function of distance, Figure 4.6a). This LFP correlation length is comparable to that reported in other LFP studies (Katzner et al., 2009; Xing et al., 2009) and calcium imaging studies of motor cortical activity (Dombeck et al., 2009; Hira et al., 2013; Peters et al., 2014), and approximately matches the correlation length of the calcium activity in projection neurons described above (exponential fit to the difference in calcium event times as a function of distance, Figure 4.6b). The LFP reflects a complex mixture of spiking and synaptic activity, but this analysis confirms through an independent modality that HVC activity is correlated over a 100 micron length scale during singing.

4.2.3 The HVC LFP reveals that on average, HVC_X and interneurons fire at different phases of a 25-35 Hz inhibitory rhythm

We next sought to examine the relationship between interneuron and projection neuron timing. Projection neuron calcium transients are correlated over a 100 μm length scale. Similarly, LFPs, which are highly correlated with the timing of inhibitory neurons, are also correlated over a 100 μm length scale (Figure 4.6a). Previous work has shown that dense, stereotyped interneuron firing patterns occur during song (Figure 4.7a) (Guitchounts et al., 2013; Kozhevnikov and Fee, 2007b), yet little is known about the relationship between projection neuron timing and patterned inhibition in singing birds. Figure 4.7b-d illustrates three distinct possibilities: (1) local interneurons are asynchronous and population activity is not deeply modulated during song (i.e. there

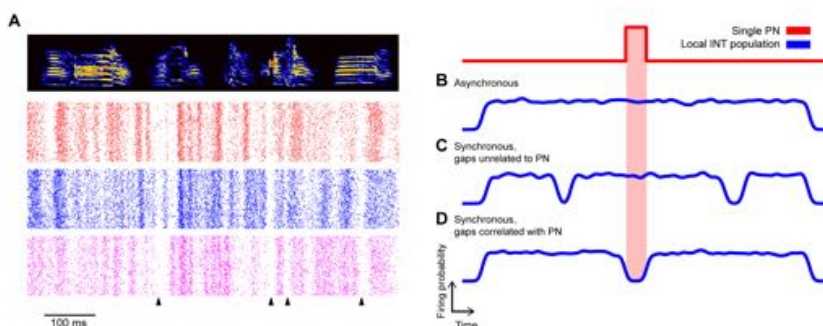


Figure 4-7: Inhibitory neuron firing patterns and three models for projection neuron/ inhibitory neuron relationships. **a**, Raster plots of three nearby ($<200 \mu\text{m}$) inhibitory neurons recorded during singing from the same bird. Synchronous pauses are marked with a black triangle. Shown to the right are three conceptual models of timing relationships between inhibitory and projection neuron activity. **b**, Interneuron firing patterns are uncorrelated in local ensembles, leading to a net local inhibition that is weakly modulated in time. **c**, Gaps in inhibitory neuron firing activity are locally correlated, leading to synchronous pauses in inhibition, but these pauses are unrelated to projection neuron activity. **d**, Gaps in inhibitory neuron firing activity are locally correlated, and contribute to defining projection neuron firing times.

is no spatial structure, Figure 4-7b), (2) local interneurons show correlated peaks or troughs in firing rates, but these modulations are unrelated to projection neuron activity (Figure 4-7c), or (3) local interneurons manifest correlated peaks or troughs in firing rates, and these modulations are involved in defining excitatory projection neuron firing times (Figure 4-7d). While it was not possible to distinguish between these models by direct measurement of connected ensembles of identified cells, we can do this indirectly by analyzing cell-type specific firing rules in HVC in relation to the LFP. In other species and contexts, spike-field correlations have revealed information about the relative timing of excitatory and inhibitory cells in local volumes of the brain (Csicsvari et al., 2003; Klausberger et al., 2003). Here, to study the relationship between different cell types and the LFP, we computed burst-triggered averages of the LFP. For interneurons, the burst-triggered average reveals a trough in the LFP at approximately zero latency (Figure 4-8a). In contrast, HVC_X neurons ($n=12$ cells, identified by the presence of two or more sparse bursts per song motif (Kozhevnikov and Fee, 2007b)) fire during peaks in the 25-35 Hz LFP (Figure 4-8b).

The time-lag between HVC_X neurons and interneurons is particularly visible by binning the LFP power by the LFP phase at each burst time (Figure 4.8c), revealing that HVC_X neurons fire in the peaks of the LFP while interneurons fire in the troughs. (The remaining projection neurons, $n=7$, fired only a single burst and could not be uniquely identified as either HVC_X or HVC_{RA} neurons, and were not included here.)

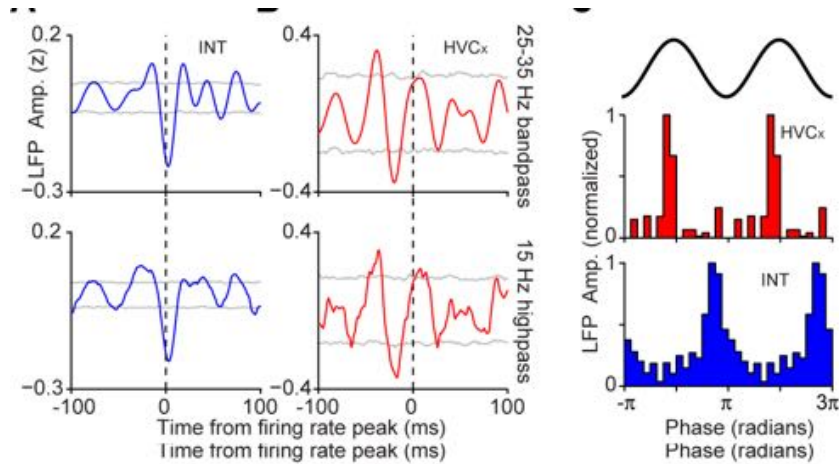


Figure 4.8: Spike-field analysis reveals an alternating 30 Hz rhythm in excitatory and inhibitory cells. **a**, Burst-triggered averages of 25-35 Hz (top row) and broadband LFPs (bottom) for interneurons ($n=65$, $p=0$, bootstrap test, see Materials and Methods) and **(b)** HVC_X neurons ($n=12$, $p=.006$, bootstrap test). Gray line indicates $p=.05$. **c**, Here, the power of the LFP is binned at the phase of the LFP for each burst time. A clear separation in preferred phase can be seen between HVC_X neurons and interneurons.

These observations indicate that the phase of the 30 Hz LFP is highly correlated with firing times in two cell classes in HVC, and suggests that HVC_X neurons, on average, fire during gaps in local inhibition. Anecdotally, within 16 channel carbon fiber bundles localized to volumes of $200 \mu\text{m}$, we have observed multiple interneurons that produce correlated firing patterns and occasional synchronous pauses in inhibition across cells (Figure 4.7a). More rigorously, we find that interneuron troughs that occur within song syllables are more synchronous than one would expect by chance (median correlation coefficient .3848, .2895-.4652 99% bootstrap confidence interval; $p=0$, bootstrap test, $n=71$ interneuron pairs, see Materials and Methods). This anal-

ysis was restricted to periods of time when the bird was actively producing sound. As a result, this analysis is not dominated by any modulations in firing rates related to the pattern of sound and silence that occurs within a song motif. Since the precise locations of the carbon fiber electrode tips were not known, we could not relate the degree of correlation to the spatial separation of interneurons. Synchronous pauses in inhibition could be responsible for the dips in the LFP, and could also provide a 'window of opportunity' for projection neurons to fire. At least one class of projection neuron, the HVC_X neuron, tends to fire out of phase with inhibition in a 30 Hz time scale.

4.3 Discussion

In singing birds, the relationship between time and space in premotor cortical area HVC is neither random, nor globally organized as a traveling wave, but clustered on a 100 μm length scale. A 30 Hz rhythm revealed in the LFP and correlated with inhibition influences the timing of at least one projection neuron sub-type (HVC_X), and this rhythm is coherent over a length-scale matching that of projection neuron calcium activity (Figure 4-6).

The mesoscopic pattern apparent in the LFP and calcium imaging has not been observable using prior methods of sampling single cell activity serially. The stereotyped behavior of the songbird has allowed for alignment of multiple renditions of song, and the construction of 'pseudo-ensembles' based on populations of cells that were not simultaneously recorded (Kozhevnikov and Fee, 2007b). This approach has proven to be powerful in the song system, and has led to rapid advances in understanding the song circuit, but it is an approach that cannot resolve microcircuit dynamics in small volumes of HVC, and cannot resolve spatial maps on the length scale described here. For lack of experimental data, prior dynamical models of HVC have not considered

local spatial correlations in time cell firing Long:2010db. Most models also ignore the contribution of inhibitory interneurons (Li and Greenside, 2006), or assume that inhibition guards against runaway excitation Long:2010db, but is otherwise uninvolved in shaping projection neuron firing times (though there are a few notable exceptions (Bertram et al., 2014; Gibb et al., 2009)).

In contrast, we find that the local microcircuit patterns in HVC in singing birds include correlations over space, and specific rules for the relative timing of distinct cell types. Our results reveal that HVC contains 'time domains' characterized by nearly synchronous projection neuron calcium activity. The 100 μm correlation length of the projection neurons also appears in the 30 Hz LFP, a signal that is highly correlated with an inhibitory rhythm expressed locally in HVC (Figure 4-4b-c, and Figure 4-8a). Within a local patch of HVC, the LFP and spike correlations indicate that, on average, HVC_X projection neurons fire in gaps of inhibition (Figure 4-8b). Taken together, these results indicate that functional domains within the HVC microcircuit may contain tightly correlated populations of excitatory and inhibitory cells that fire in opposition and generate a local 30 Hz rhythm in the LFP. These 'time domains' resemble the intermediate scale of geometric organization illustrated in Figure 4-1b, characterized by zones in which projection neurons fire at similar times with no coherent global organization.

Important questions remain about the microcircuit dynamics of HVC. Some are technical: the imaging results reported here rest on calcium imaging and the relationship between calcium transients and the underlying burst of action potentials in HVC projection neurons has not been measured during singing. At present, we can say that to first approximation the sparse calcium transients match the expected burst behavior of the excitatory cells in HVC. The head mounted microscope provided only single-color imaging, precluding cell type identification based on retrograde labeling

of downstream targets. Similarly, the rhythm between HVC_X cells and interneurons is inferred based on average timing relative to a common signal rather than direct measurements of connected pairs of cells. Our electrophysiology did not address the relationship between HVC_{RA} projection neurons and the inhibitory rhythm. The role each cell type in HVC plays in driving or responding to the local inhibitory rhythm, and the connectivity that gives rise to this rhythm, await more detailed description (Mooney, 2000; Mooney and Prather, 2005b). Moreover, neither the imaging nor the electrophysiology could directly measure the correlation length of the inhibitory neurons. This was extrapolated from the de-phasing length scale of the 30 Hz LFP, a measure clearly related to interneuron firing patterns, but still an imperfect surrogate.

Perhaps more interesting are the outstanding questions about the nature of a dynamical system governed by the properties observed here. Inhibitory neurons show patterned activity throughout song, whereas projection neurons fire sparsely, with calcium events that are correlated in time between neighbors. This suggests a microcircuit in which multiple pools of excitatory neurons drive local interneuron populations. Then, in a feedback loop, locally synchronized inhibition on a 30 ms time-scale also influences the firing times of multiple pools of excitatory neurons. Theoretical models of neural sequence generation have focussed on the direct interactions between excitatory neurons, but our results suggest that the indirect interactions between excitatory neurons, mediated through inhibition, may be important in defining which cells fire at a given time. In this view, the timing of a projection neuron's activity may be defined not only by its connections to other projection neurons, but also by rhythmic patterns of inhibition dominated by a 30 Hz time-scale. Considering the importance of excitatory-inhibitory coupling in defining the timing of at least one class of projection neuron in HVC, the space of possible models for HVC activity expands significantly.

HVC is one stage in a motor cortico-thalamic loop, and intriguingly the time-scale of activity propagation around this loop is similar to the 30 Hz rhythm described here (Ashmore et al., 2008). Future work is needed to examine whether this time-scale is intrinsic to cellular biophysics in HVC or arises from the time it takes excitatory pulses to propagate around the cortico-thalamic loop (Hamaguchi et al., 2012). Whether this 30 ms time-scale is related to acoustic transitions in song remain to be examined (Amador et al., 2013; Gardner et al., 2001) and will require more extensive sampling of excitatory and inhibitory cells over the entire volume of HVC.

Functional clustering in HVC has been described previously in the spontaneous activity of anesthetized songbirds (Graber et al., 2013), but we note a possible difference in the length scale reported in the spontaneous activity of anesthetized birds ($\lambda=263 \mu\text{m}$) and the length scale we observe in singing birds ($\lambda=104 \mu\text{m}$ for calcium imaging, $\lambda=108\text{-}125 \mu\text{m}$ for LFPs). The length scale of correlation during singing matches functional clustering on a $100 \mu\text{m}$ length scale that has been observed in mammalian motor cortex in the context of running, grooming (Dombeck et al., 2009), and simple forelimb movements (Hira et al., 2013; Peters et al., 2014). Additionally, the 30 Hz rhythm has parallels in primate motor control (Bartolo et al., 2014; Murthy and Fetz, 1992; Rubino et al., 2006) and speech perception and production in humans (Giraud et al., 2007). The recurrence of these length scales and rhythmic time-scales across species and contexts raises the possibility that spatial clustering and 30 Hz rhythmic alternation of excitatory and inhibitory cells may be a recurring feature of cortical motor microcircuits and critical for the sequential organization of behavior. If true, this would imply that relatively slow movements produced by primates and the fast vocal movements of a songbird are coarse-grained in a similar fundamental length-scale ($100 \mu\text{m}$) and time-scale (30 Hz) at the level of cortical control.

For studies of songbird vocal learning, these results suggest that future biophysical

models of motor sequence generation in HVC should move away from abstract models of neuronal chains defined only by the topology of excitatory cells, to consider new classes of models in which the interplay of spatially correlated patterns of excitation and inhibition form the basis of the stereotyped, robust sequences of neural activity driving birdsong.

4.4 Materials and Methods:

4.4.1 Subjects

All procedures were approved by the Institutional Animal Care and Use Committee of Boston University (protocol numbers 09-007 and 11-027). Electrophysiology data were collected from 27 adult male zebra finches (>120 DPH), and imaging data were collected from 9 adult male zebra finches. Birds were kept on a 14 hr light-dark cycle.

4.4.2 Viral labeling and calcium imaging.

Chronic calcium imaging was performed using a head mounted fluorescent microscope described in detail elsewhere (Ziv et al., 2013). To label HVC neurons with GCaMP6s calcium indicator, birds received four 250 nl injections of lentivirus packaged with GCaMP6s under an Rous sarcoma virus (RSV) promoter into HVC (Scott and Lois, 2005). To guide the injection of virus, the bounds of HVC were determined by bipolar antidromic stimulation of Area X (n=3) or through fluorescence imaging of a DiI retrograde tracer (n=6) (Scott et al., 2012). We collected 167 ROIs from 9 birds. Fluorescent traces for each animal were collected over a single day of imaging. For the calcium imaging, all songs were directed (i.e. a female was present).

Our unpublished results indicate that the RSV promoter in HVC does not label interneurons or glial cells. We have infected HVC with viruses with the RSV promoter driving transcription of several genes fused to GFP and performed double

immunocytochemistry against GFP and several markers for inhibitory interneurons (parvalbumin, somatostatin and calretinin). We did not observe co-localization of cells labeled with GFP with any of the interneuronal markers. In addition, we also observed that the morphology of the GFP-positive cells was consistent with that previously described for projection neurons in HVC, but not consistent with them being glial cells. This observation suggests that the RSV promoter labels projection neurons in HVC.

GFP immunolabeling and confocal imaging of the RSV-GCaMP6s construct used here revealed that in all cases where dendrites were visible they contained spines, consistent with the morphology of HVC_{RA} and HVC_X projection neurons. Thirty percent of all infected neurons were HVC_X projection neurons, based on retrograde labeling from Area X. Retrograde labeling was not performed from RA, but GFP immunolabeling revealed dense axon arbors in nucleus RA, indicating that a significant number of HVC_{RA} neurons were also labelled.

4.4.3 Electrophysiology.

Electrophysiological recordings were gathered from 27 birds, where 13 received implants from 16-channel carbon fiber bundles, 5 from four 4-channel carbon fiber bundles, 3 from tungsten microwire arrays (Tucker Davis Technologies, 16-channel 33 μm wire diameter), and 6 from four-shank silicon probes (n=4 along the mediolateral axis, n=2 along the anterior-posterior, Neuronexus model no. A4x4-3mm-50-125-703-CM16, 15 μm thickness). All arrays were implanted .7 mm anterior and 2.3 mm lateral of the midsagittal sinus at a depth of .4-.7 mm (head angle 30°) using previously described methods (Guitchounts et al., 2013). All songs used for electrophysiology were undirected (i.e. no female was present). In a subset of the implants the position of the implant (but not single units) was verified antidromically with bipolar stimulating electrodes placed in Area X (Hahnloser et al., 2002a). Excitatory

projection neurons and putative interneurons were identified based on their unique firing properties. In previous studies, antidromically identified HVC_X projection neurons were shown to produce two more sparse, time locked bursts per motif. Putative interneurons produced dense, tonically active, time-locked firing patterns (Hahnloser et al., 2002a). All neurons reported in this study fell into one of these two categories. Local field potentials (LFPs) were extracted from the same electrodes used to record single units after spike removal, and all statements about spike-field relationships were also confirmed with an analysis in which the LFPs were extracted from nearby electrodes, indicating that no cross-contamination between spikes and fields influenced these results.

4.4.4 Analysis of calcium imaging data

All offline analyses of the calcium imaging data were performed using custom software written in MATLAB. First, the raw imaging data was motion corrected using a previously published algorithm (Guizar-Sicairos et al., 2008). Regions of interest (ROIs) were then manually selected and converted into single cell fluorescence traces by taking the mean intensity of all pixels within a given ROI for each movie frame. Baseline fluorescence for each trace, F_0 , was computed by taking the 12th percentile in an 800 ms sliding window (400 ms at each time point). Changes in fluorescence were then defined as ΔF , where F is the fluorescence of a given ROI and t is a particular point in time. Event times were computed by first finding all local peaks in the trace that exceeded 1%. Events were excluded if ΔF did not remain above -1% for 300 ms after the detected peak. Then, the event time was defined as the 50% rise time between the peak and the preceding trough (Ziv et al., 2013). To provide an upper bound on the false positive rate given these criteria, which were chosen manually based on visual inspection of the data, we computed the likelihood of finding a peak when the birds were not singing ($p=.0156$ from 61.5 seconds of silence). All spatial correlations

were computed using these event times, derived from temporally sparse fluorescence transients in spatially sparse populations of infected neurons. The correlation length-scales calculated from these event times are robust to any out of focus fluorescence or neuropil signal that could add spurious correlations to the calcium time series.

We also note that the distance measurements used for the spatial correlations are exact only if the imaged cells are all in the same plane, which is not the case here. The depth of each cell could not be determined since the head mounted cameras were used in a fixed location. To estimate the error due to this ambiguity, we focused the microscope sharply on $1\ \mu\text{m}$ beads in agarose, and then defocused until the image of a bead expanded to the size of the largest selected ROI. This provides a conservative measurement of how far outside the imaging plane a cell can be while still meeting the selection criteria of an ROI. A more realistic measure was computed by defocusing from a spot the same size as the smallest ROI until it was distorted the same size as the largest ROI. By the conservative measure, cells that appear overlapping in the XY plane may be separated by a z-axis distance of $35\ \mu\text{m}$ ($22\ \mu\text{m}$ by the more realistic measure). However, the true ambiguity in the euclidian cell to cell distance, taking into account the ambiguity in the Z plane, will decrease as ROIs appear more distant in the XY plane. Cells observed to be $50\ \mu\text{m}$ apart will have an uncertainty of $11\ \mu\text{m}$ ($4.5\ \mu\text{m}$); at $100\ \mu\text{m}$, the uncertainty is only $6\ \mu\text{m}$ ($2.4\ \mu\text{m}$).

4.4.5 Bootstrap tests of imaging data.

The analyses shown in Figure 4-2f-g were all conducted using randomization tests. For Figure 4-2f, the median distance between ROIs was plotted against the time between their respective calcium events. To test for significance at each time bin, the same number of ROIs were randomly selected from the entire data set 10,000 times, and the median was computed for each randomization. The p-value is then the probability that a median value from the randomized distribution is lower than

the observed median. In Figure 4.2g, the same procedure was used to test the length of the resultant vector. That is, we tested whether the length of the resultant vector was longer than would be expected using the same randomization. Finally, we used a similar procedure to test for the significance of the standard deviation of the anterior-posterior and mediolateral distance distributions.

4.4.6 Analysis of electrophysiology data

Extracellular voltage traces were multiplexed and digitized at 25 kHz on the headstage and transferred to a PC over USB (Intan RHA2000). Song was recorded using a head-mounted microphone glued to the headstage (Guitchounts et al., 2013). All offline analyses were performed using custom software written in MATLAB (Mathworks). To isolate single units offline, the extracellular voltage traces were bandpass filtered between 600 and 11000 Hz (12th order Elliptic filter, .2 dB passband ripple, 40 dB stopband attenuation) and sorted using standard offline spike sorting techniques (Sahani, 1999; Tolia et al., 2007). For the LFPs, voltage traces were first median filtered (1 ms window) to remove spikes (Pesaran et al., 2008), then lowpass filtered with a 400 Hz corner frequency (4th order Butterworth filter) and downsampled to 1 kHz. To isolate the 25-35 Hz LFP, the downsampled LFP was bandpass filtered between 25-35 Hz (149 tap Kaiser window FIR filter, 40 dB stopband attenuation, .05 ripple). To minimize the impulse response, a more lenient 25-35 Hz bandpass was used for the spike-triggered analysis shown in Figure 8a-b (top) and for the comparison shown in Figure 4c (53 tap Kaiser window FIR filter, 20 dB stopband attenuation, .05 ripple). All event-triggered LFP analyses were also repeated using only a 15 Hz highpass filter (Figure 4.8a-b bottom; 84 tap Kaiser window FIR filter, 20 dB stopband attenuation, .5 ripple). All filtering was performed forwards and backwards to correct for phase distortion. Songs were aligned using previously described methods (Poole et al., 2012). Spike-field coherence was estimated using standard techniques

based on the cross-spectrum between the spike train and the LFP (Jarvis and Mitra, 2001) (1 Slepian taper, $NW=1$). To control for bias, the same number of trials were used for each neuron. Additionally, we only included high quality single units for measuring spike field coherence ($n=17$ interneurons, $SNR>8$ and $<.1\%$ ISIs $<1ms$, no sign of cluster contamination), all other measures included all interneurons.

4.4.7 Song alignments

Songs were aligned using previously described methods (Poole et al., 2012). In brief, trials were aligned to song by computing the Euclidean distance in spectral features between the data and a template song in a sliding window. Troughs in the distance were considered candidate matches, and the spectral features of all potential hits were plotted in 2 dimensions and the user performed a cluster cut.

4.4.8 Spike sorting.

First, positive and negative-going threshold crossings in the 600-11000 Hz bandpassed extracellular voltage traces were detected. The threshold was set to four times the bandpassed signal's robust standard deviation (Quiroga et al., 2004). Then, a 1.1 ms window centered on the negative peak was saved. The spikes were then re-aligned to the negative peak after upsampling by a factor of 8 using cubic splines. Features of the aligned spike windows were computed using robust principal components analysis (Sahani, 1999). A mixture of Gaussians model was fit to the spike features using the split and merge expectation maximization algorithm (Tolias et al., 2007), and the number of components in the mixture was determined by fitting models with 2-7 components and choosing the model with the minimum Bayes Information Criterion (Tolias et al., 2007). Unit quality was then assessed by signal-to-noise ratio and refractory period violations. For multi-unit analysis all threshold crossings above 3 standard deviations of the bandpassed signal were included.

4.4.9 Spectral density images

To compute the spectral density image for multiple aligned renditions of the same LFP, we used previously described methods (Markowitz et al., 2013b). First, a sparse, binary time-frequency representation of each rendition was generated using 'auditory contours' Lim:2012uh. Then, the binary representations were combined by summing across all renditions. In this representation, a false-color plot (Figure 4.3a bottom) represents the probability of a binary contour passing through a given point in the time-frequency plane.

4.4.10 Identifying bursts in projection neurons and interneurons for burst-triggered averaging.

To identify bursts of single-unit activity, we first estimated the smooth firing rate for each trial using a Gaussian kernel ($\sigma = 5$ ms). The smooth firing rate was averaged across trials, and we counted peaks above 100 Hz as burst events.

4.4.11 Computing firing pattern distance

We define the distance between two firing patterns as the mean absolute difference between their average firing rates. The firing rate was estimated by computing the smooth firing rate for each trial (Gaussian kernel, $\sigma = 5$ ms), and then averaging across trials.

4.4.12 Determining whether interneuron troughs are synchronous.

To determine whether interneuron troughs within syllables were more coherent than one would expect by chance, we conducted a bootstrap procedure. First, the firing rate was estimated within syllables by computing the smooth firing rate for each trial (Gaussian kernel, $\sigma = 5$ ms), averaging across trials, and z-scoring the average. All negative-going peaks less than -1.5 were counted as troughs. A binary vector was

created with 1s indicating the presence of a trough and 0s denoting all other time points (troughs in inter-syllable gaps were excluded), and each vector was smoothed with a Gaussian kernel $\sigma = 5$ ms. Then we computed the Pearson correlation coefficient between these smoothed vectors from all pairs of interneurons from the same bird (carbon implants only, $n=71$ pairs). The median correlation coefficient across all birds was then compared to a null distribution estimated using a bootstrap procedure in which the troughs were selected randomly from time points during syllables ($n=1000$ bootstraps).

4.4.13 Phase locking.

All phase locking indices (PLIs) were computed using previously described methods (Lachaux et al., 1999).

4.4.14 Statistical tests

No formal methods were used to pre-determine sample sizes, but the sample sizes used here are similar to those generally used in the field. All statistical comparisons were performed using either a one-tailed or two-tailed Wilcoxon ranksum test with four exceptions. First, the significance of the imaging results was determined using a bootstrap procedure (Figure 4.2f-g, described above). Second, the significance of spike-field coherence (Figure 4b) was determined using a bootstrap procedure by shifting the spike trains across all trials by a separate random number for each neuron (1000 bootstraps). Third, the significance of burst-triggered average LFPs was computed using a bootstrap procedure. Burst times were randomized for each neuron (1000 bootstraps), and the maximum absolute value was compared between the observed burst-triggered average and the bootstraps. Finally, the statistical significance of the trough timing was determined using a bootstrap where the troughs were randomly selected from time points during syllables (1000 bootstraps). Where

appropriate, we controlled for multiple comparisons using the Bonferroni correction.

Chapter 5

Unstable neurons underlie a stable learned behavior

This chapter is a reproduction of previously published work (Liberti et al., 2016), and contains critical contributions from *Dr. Jeffrey Markowitz*, who is an equal co-author on this manuscript.

5.1 Introduction

Questions about coding stability at the single neuron level have been notoriously difficult to address given the challenge of stably recording from single neurons using implanted electrodes. In the hippocampus, earlier methods tracking tens of cells emphasized stable neural tuning over a timescale of a week (Kentros et al., 2004; Thompson and Best, 1990), whereas recent studies tracking thousands of cells using new optical techniques (Chen et al., 2013; Ghosh et al., 2011) revealed that 75-85% of the cells change their tuning properties within the same timeframe (Ziv et al., 2013). In whisker motor cortex, individual neurons in mice trained on an object detection task were unstable, but the relationship between ensemble measurements and behavior remained stable (Huber et al., 2012b). These studies support the view that for stable behaviors, individual neurons involved can show substantial drift in their tuning properties. For motor systems, the stability of neural tuning is not necessarily expected; the convergence of vast numbers of neurons in motor cortex to relatively few muscles (Leonardo, 2005) suggests that a given muscular activation

pattern could be produced by many patterns of neural activity (Rokni et al., 2007). In parallel, single neurons in motor cortex are observed to switch tuning properties in a task-dependent manner (Ganguly and Carmena, 2009). In the case of the zebra finch, the precise timing and acoustic structure of song is preserved for years (Immelmann, 1969; Lombardino and Nottebohm, 2000b). Since the zebra finch sings a single song over the course of his lifetime, the question of neural coding stability is particularly well-defined. In the pre-motor cortical area that drives song, HVC (used as a proper name), three primary classes of neurons can be found: inhibitory interneurons, basal-ganglia projecting neurons (HVC_X), and motor projecting neurons (HVC_{RA}). A third projection neuron type HVC_{AV} sparsely populates the nucleus, but its activity has not yet been described in singing birds (Mooney et al., 2014). In HVC, each neuron type that has been observed produces a highly stereotyped pattern of action potentials, and these patterns are stable over the time intervals that they have been recorded: minutes to hours (Hahnloser et al., 2002a; Kozhevnikov and Fee, 2007b). For inhibitory interneurons, the pattern is dense, with spikes occurring throughout song (Guitchounts et al., 2013; Kozhevnikov and Fee, 2007b). In contrast, excitatory projection neurons that communicate with downstream targets fire ‘ultra sparsely’ (Hahnloser et al., 2002a).

During song production, HVC_X neurons and inhibitory interneurons exhibit cell-type specific phase-locking to the local field potential (LFP) at 30 Hz. HVC_X cells fire in the peak of the LFP and interneurons in the trough (Markowitz et al., 2015). Intracellular recordings in vitro and in vivo suggest that HVC_{RA} neurons also fire in gaps of inhibition (Kosche and Long, 2014). Moreover, both the LFP and projection neuron activity are clustered over a 100 μm length-scale. The phase locking relationship between excitatory and inhibitory neurons is only observed with multiple neuron recordings in the same small volume/phase shifts across HVC preclude a

global rhythmic relationship among cell types (Lynch et al., 2016). Taken together, these observations indicate that spatially coherent mesoscopic patterns of inhibition underlie HVC dynamics, where synchronous gaps in local interneuron population activity control projection neuron timing.

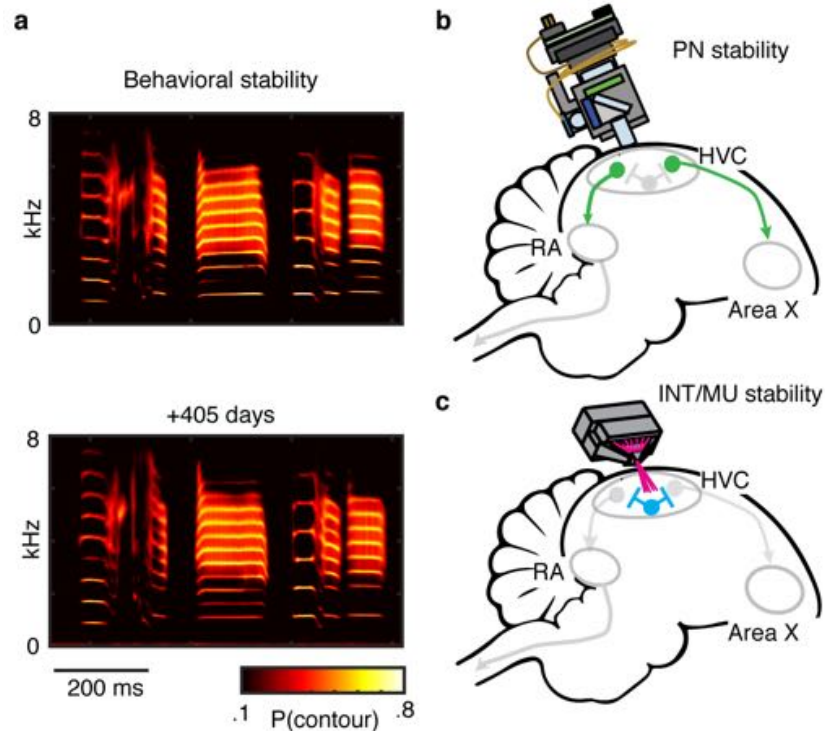


Figure 5-1: Approach to measuring the stability of neural firing patterns underlying a highly stable behavior. **A**, Spectral density images, for the songs sung by the same bird 405 days apart ($n=1412$ songs for Day 1 and $n=1480$ songs for Day 405). **B**, To measure projection neuron stability a head mounted fluorescence microscope was used to image GCaMP6 delivered by a cell-type specific virus. **C**, Stability of multi-unit and inhibitory interneuron activity patterns was measured using a custom carbon fiber electrode array.

The recent evidence in support of locally organized ensemble activity in HVC allows us to raise the following question: What explains the persistence of the song motor pattern, single neuron stability, or stability of ensemble dynamics. A recent experiment suggests that HVC ensemble dynamics measured through multi-unit activity (Crandall et al., 2007) is resilient to circuit perturbations (Otchy et al., 2015).

Here we extend the observation of HVC ensemble dynamics using new electrophysiological (Guitchounts et al., 2013) and imaging methods (Figure 5-1b,c) to address the stability of excitatory projection neurons.

5.2 Results

5.2.1 Song is supported by stable mesoscopic dynamics

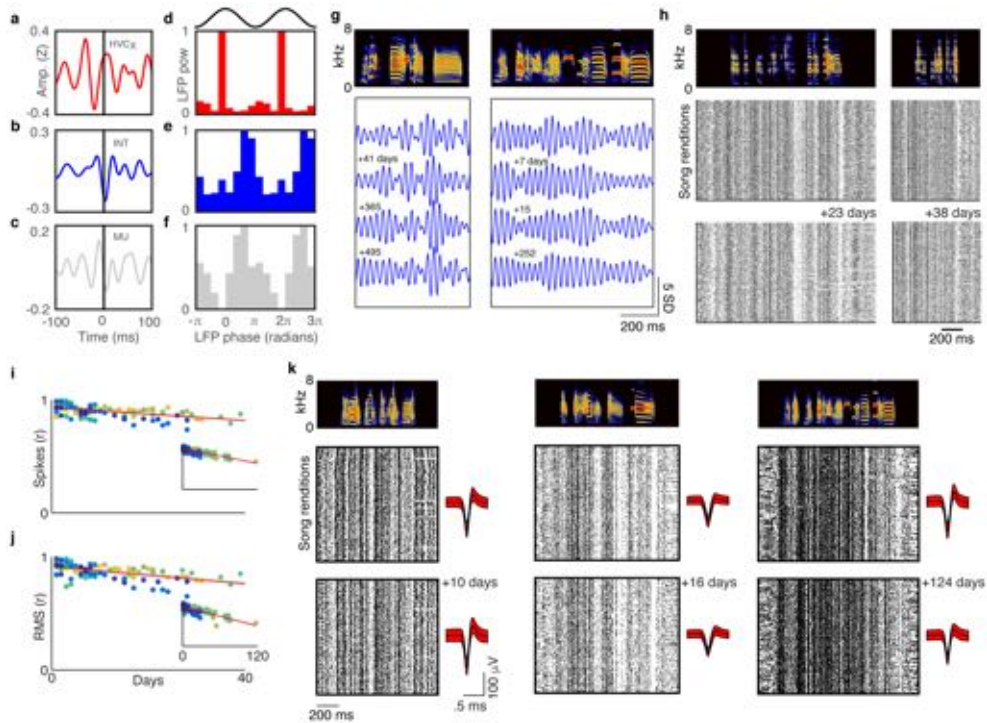


Figure 5-2: Activity patterns from multi-unit ensembles and single interneurons are highly stable. **A-C** Burst-triggered averages of 25-35 Hz Local Field Potentials (LFPs) for HVC_X neurons (**A**), interneurons (**B**), and multi-unit activity (**C**). **D-F**, The power of the LFP is binned at the phase of the LFP for each burst time. A clear separation in preferred phase can be seen between HVC_X neurons (**D**) and interneurons (**E**), but not between interneurons and multiunit activity (**F**). **G**, Trial-averaged band-passed LFPs (25-35 Hz) from a single electrode up to 495 days post-implant. **H**, Examples of multiunit stability over month-long timescales. **I-J**, Multi-unit patterns are stable over 40 days measured using patterns in firing rate (**I**) or root-mean-square (RMS) (**J**) values are stable over 40 days (inset, over 120 days). Colors indicate different birds, many are not visible due to overlap in the upper-left portion of the plot. **K**, Inhibitory interneuron firing patterns are stable for up to 124 days. Average waveforms are shown next to the corresponding raster (red shading indicates standard deviation).

Local field potentials (LFPs) can reflect the synchronous activity of neural ensem-

bles over a length-scale of approximately $100\ \mu\text{m}$ (Katzner et al., 2009) (though see (Buzsáki et al., 2012)). We recorded LFPs in the zebra finch pre-motor cortex (HVC) using both commercial and previously-described custom carbon fiber electrode arrays (Guitchounts et al., 2013; Markowitz et al., 2015). The phase of the 30 Hz LFP is coherent with the spiking of HVC interneurons and projection neurons (Figure 5·2a-f) (Markowitz et al., 2015); we found this local phase was precisely conserved over long timescales (Figure 5·2g). For most of our implants we were able to record the LFP for up to 30 days, and we found that over this timescale the 30 Hz phase exhibited a drift of less than 0.25 radians or approximately 15 degrees (modal change of .244 radians over 30 days, .232-.254, 95% bootstrap confidence interval, Extended Data Figure 5·1a). This stability in LFP phase indicates that ensemble activity, reflecting a combination of local spiking and presynaptic inputs, does not undergo a major reconfiguration.

5.2.2 Multi-unit ensembles and single inhibitory neurons are stable

While the LFP is thought to reflect ensemble activity from both excitatory and inhibitory neurons over a length-scale of approximately $100\ \mu\text{ms}$ (Markowitz et al., 2015), multi-unit recordings from large surface area electrodes can reflect the aggregate activity of tens of nearby neurons within a local volume. To monitor the stability of multi-unit activity, we implanted custom electrodes that were geometrically optimized for sampling from cells across the full depth of HVC. Figure 5·2h reveals a multi-unit raster with detailed temporal structure in the pattern of firing and silence, with an average precision of 4.587 ms (4.348-4.843 ms, 95% bootstrap confidence interval consistent with the stereotyped neural activity reported previously in HVC (Hahnloser et al., 2002a; Schmidt, 2003). The neural patterns shown in Figure 5·2h represent finely timed bursts of activity in an unknown number of cells (the number of simultaneously active cells recorded per electrode was more than could

be separated by spike-sorting). However, spike-field measurements indicate that the multi-unit signal is phase-locked to the trough of the 30 Hz component of the LFP, suggesting that multi-unit signals are dominated by inhibitory interneurons (Figure 5·2c,f) (Markowitz et al., 2015). Enhanced representation of inhibitory cell activity in multi-unit signals is to be expected since projection neurons fire 'ultra-sparsely' (Hahnloser et al., 2002a) and contribute many fewer spikes to a multi-unit signal than do interneurons. The suggestion drawn from multi-unit signals that patterns of inhibition are stable in HVC (Figure 5·2h-j) was directly confirmed in long term recordings from n=6 inhibitory neurons from single unit recordings (3 are shown in Figure 5·2k). As revealed in Figure 5·2k, inhibitory neuron firing patterns were remarkably stable, even over a time-scale of months. In addition to the stability of the 30 Hz LFP phase, this result indicates that ensemble inhibitory activity is highly stable.

We next asked if the stability of this multi-unit activity depended on sensory feedback. Previous work has shown that the statistical distribution of post-synaptic potentials (PSPs) in HVC projection neurons is invariant to altered auditory feedback induced through a tracheosyringeal (TS) nerve cut (Vallentin and Long, 2015), but over long timescales, the song of the adult zebra finch is known to be maintained through an auditory-feedback dependent process (Nordeen and Nordeen, 1992; Tschida and Mooney, 2012). We anticipated that the rate of drift in the ensemble pattern would increase with perturbations to auditory feedback. Following baseline recording, we severed the ipsilateral TS nerve that relays motor commands to the birds' vocal organ (n=5 birds). As in previous work (Vallentin and Long, 2015; Williams et al., 1992), this loss of muscular control results in acoustic distortions of song (Figure 5·3a-d, see Extended Data Figure 5·2 for all examples), degrading the learned sensory-motor correspondence between vocal commands and auditory feedback. Following this, we monitored the song and HVC multi-unit firing patterns for

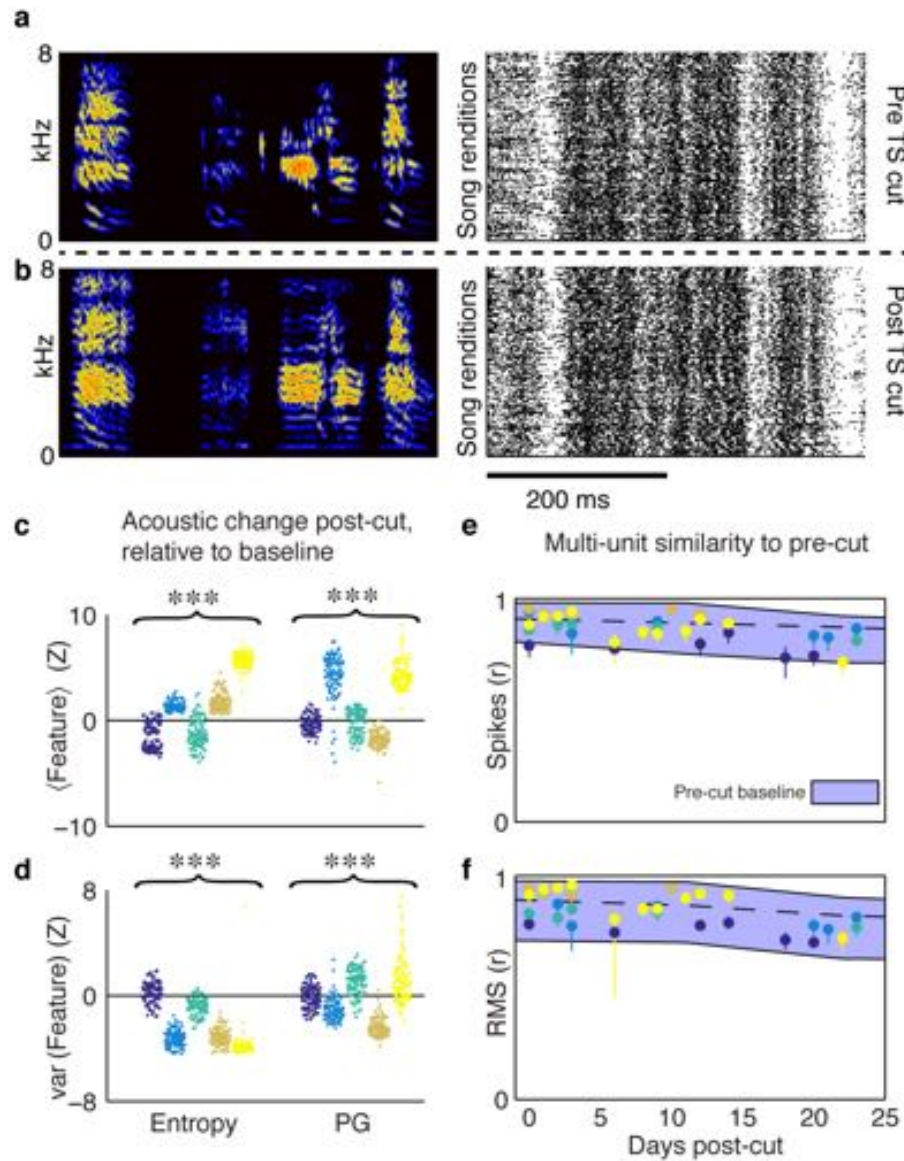


Figure 5-3: Drift in multi-unit firing patterns is not accelerated by unilateral TS nerve transection. **A-B**, Representative sonogram and multiunit raster before (A) and after the TS nerve cut (B). **C-D**, Change in acoustic features after nerve cut (***, $p < .001$, Wilcoxon ranksum test). **E-F**, Post-nerve cut, multi-unit activity does not drift more than expected in the pre-cut baseline condition. Shown is the average correlation at each time point relative to the last baseline day. Blue indicates the 95% confidence interval estimated from the baseline condition.

a period of approximately one month. Over this time interval, song did not recover, and the multi-unit spike patterns in the pre-motor cortical zone HVC remained stable

(Figure 5·3e-f). In all birds subject to the TS cut, shifts in the firing pattern occur at the same rate as for baseline recordings ($p > .05$ for all birds, one-tailed Wilcoxon ranksum test). We conclude from this that the ensemble pattern in HVC is robust to changes in auditory feedback that result from unilateral TS nerve removal. Here the song is altered, but the detailed ensemble firing patterns persist in HVC as though no change had occurred. This observation is synergistic with a recent study showing resilience of the HVC ensemble pattern after neural circuit perturbations (Otchy et al., 2015).

5.2.3 Individual projection neurons are unstable over days.

Our electrophysiology methods were unable to track individual excitatory projection cells over time-scales longer than a single day, perhaps due to limitations of recording from smaller cells (Guitchounts et al., 2013). To provide information about the stability of excitatory cells, we turned to optical imaging of genetically encoded calcium indicators (Figure 5·4a, Extended Data Figure 5·3) (Markowitz et al., 2015; Picardo et al., 2016b). This was accomplished through the use of miniature head-mounted microscopes (Ghosh et al., 2011) using both a commercial head-mounted microscope (Inscopix, $n=1$ bird), and a custom 3D-printed microscope (3 birds) designed to promote freedom of movement in singing birds (Figure 5·4 and Extended Data Figure 5·4).

This microscope provided the first measurements of calcium activity during undirected singing; the solo practice song of the zebra finch. Projection neuron calcium activity patterns were sparse and time-locked to singing (Figure 5·4a). The timing of most projection neurons was stable over several days (Figure 5·4a,b), however, across days, the probability that a given cell fires during singing can change dramatically (Figure 5·4c, Figure 5·7c). Across an interval of days, cells both drop out and newly appear in the song pattern (Figure 5·5b,c, and Figure 5·6-8). Stable and unstable

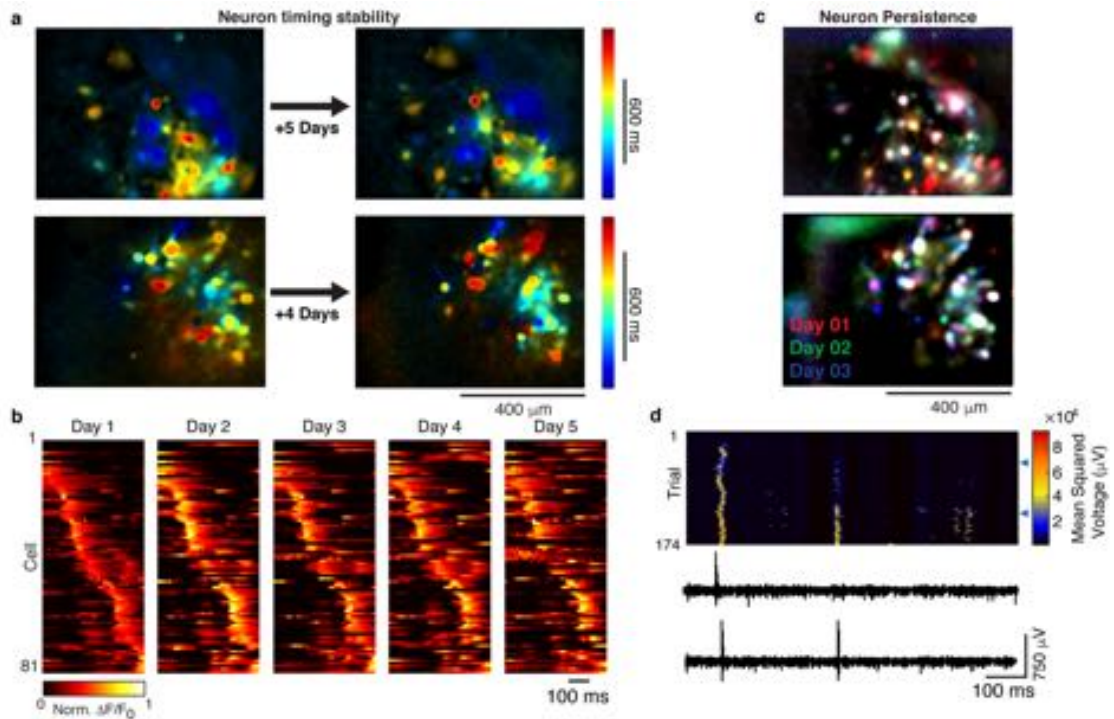


Figure 5-4: Projection neurons drift over a five day imaging session **A**, Images colored by the timing of max pixel intensity, for two birds, across 5 (*top*, Inscopix microscope) and 4 (*bottom*, custom microscope) day intervals. **B**, Trial-averaged activity from all song related neurons from one animal. Keeping the same sorting and amplitude color scale, the ROIs are plotted for 5 consecutive days. **C**, For the same animals shown in **A**, cells active on three consecutive days are combined in a single image where color now indicates neuron participation rather than timing. (red=day 1, blue=day 2, and green=day 3). Cells with persistent activity on all 3 days appear white in this representation (see Extended Data Fig 5). Cells that change amplitude or change firing probability appear colored. **D**, Electrophysiological recording of a projection neuron reveals a new song-locked burst that emerges over the course of a day. Blue triangles indicate trials plotted at the bottom of the figure.

ROIs can be found next to each other (Figure 5-6,7).

One class of projection neuron (HVC_X) can produce more than one time-locked burst of action potentials during singing. For cells with multiple timing peaks, amplitude or probability of activation could change independently for each peak. This was witnessed both in calcium imaging and in electrophysiological recordings. An example is the the 'fade-in' of a second burst time in a projection neuron recorded with a carbon fiber electrode over the course of a day (Figure 5-5d).

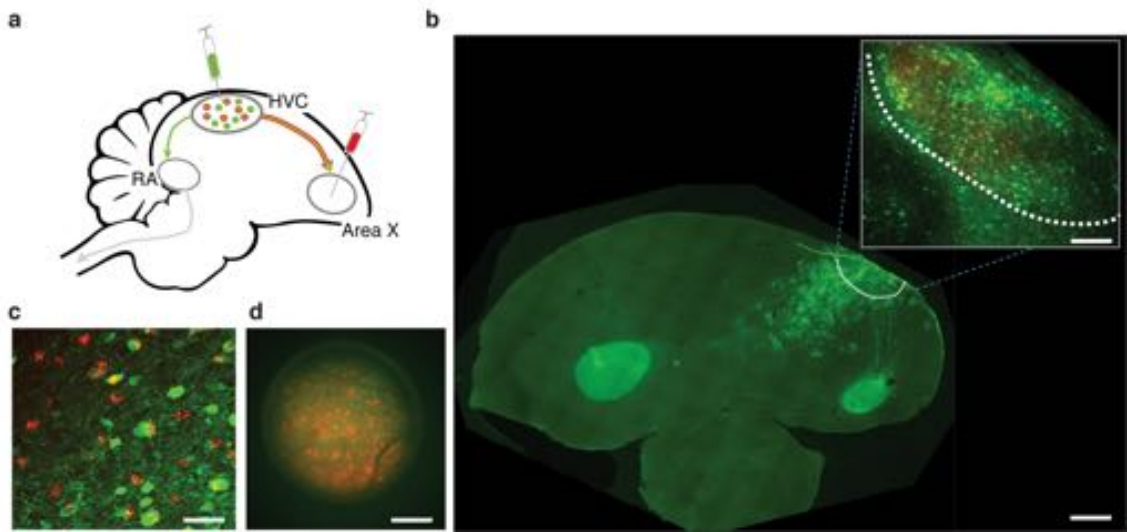


Figure 5-5: Histological verification of virus used for calcium imaging. **A**, Diagram of experimental strategy for calcium imaging. Virus expressing GFP was injected into HVC, and a retrograde dye (DiI) was injected into the downstream nucleus Area X. **B**, Lentivirus with an RSV promoter produces dense infections of HVC projection neurons. Scale bar indicates 400 μm . *Inset*, closeup of HVC, bounded by the white dotted line. Scale bar indicates 39 μm . **C**, Co-localization of GFP-expressing neurons in HVC and DiI backfill. Scale bar, 250 μm . **D**, View of HVC through a chronically implanted GRIN lens. Scale bar, 250 μm .

For the imaging data, we performed analysis on two ROI sets: one analysis included all ROIs found on any day of imaging ($\text{ROIs}_{\text{presentsomedays}}$). The second included only ROIs unambiguously present on all days of imaging ($\text{ROIs}_{\text{presentalldays}}$). By the fifth day of imaging, across all ROIs recorded, a large number of ROIs showed statistically significant changes in their mean song-aligned activity ($p < .01$ permutation test, 15/39 ROIs for bird 1, 27/38 for bird 2, 26/76 for bird 3 and 26/81 for bird 4, $\text{ROIs}_{\text{presentsomedays}}$; 11/18, 14/16, 11/22, 5/21, $\text{ROIs}_{\text{presentalldays}}$).

This shift in projection neuron activity across multiple days is inconsistent with their previously observed stability during brief, minutes-long recordings (Hahnloser et al., 2002a; Long et al., 2010). This contrast motivated us to examine whether the shift in projection neuron activity occurred primarily during the day or during the night. We found a much stronger shift in song-aligned projection neuron calcium traces after a night of sleep than over the course of a single day (Figure 5-8c, Figure

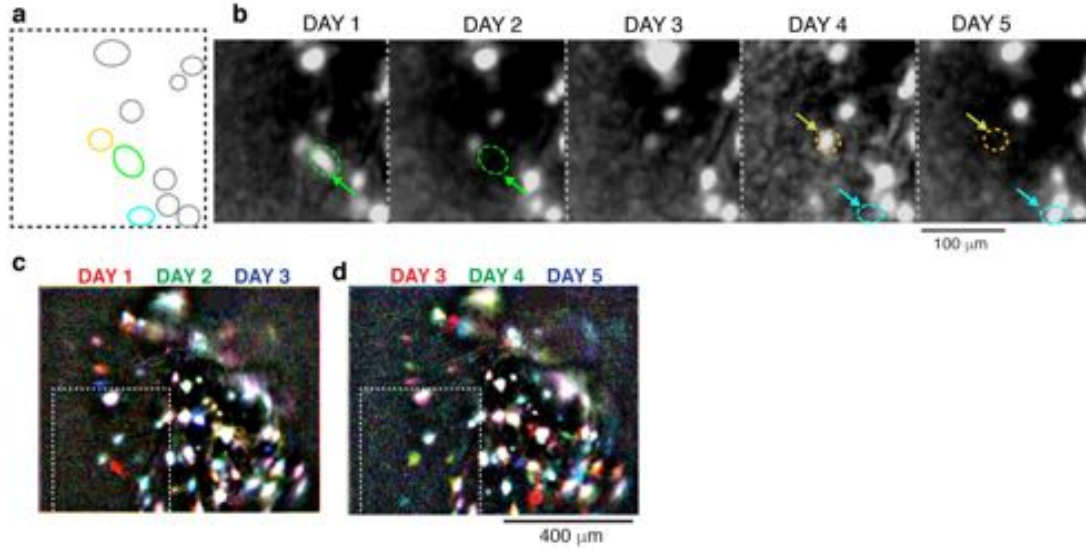


Figure 5-6: Examples of stable and unstable cells in the same imaging region. **A**, ROI masks for cells in region. **B**, Maximum projection of a $191 \times 203 \mu\text{m}$ subsection of the averaged, song-aligned calcium imaging movies, across 5 days. Arrows highlight two cells in this plane that either drop out (green, yellow) or in (blue) of the neural sequence across days. This region is a smaller section of the total imaging plane, shown in C and D. **C**, Three day maximum projection overlay from the data in Figure 5-4C. The maximum projection image is divided by a smoothed version of the same image, using a 100 pixel disk filter, to normalize across bright and dim ROIs. **D**, The last three days of the five day longitudinal study, using the same normalization as in C

5-6, $\text{ROIs}_{\text{presentsomedays}} p=2.0\text{e-}40$, $z=-13.26$, $\text{ROIs}_{\text{presentalldays}} p=1.6\text{e-}14$, -7.59 , one-tailed sign-rank test). Next, we checked if drift could be accounted for by time elapsed. First, we analyzed whether calcium activity decreased in consistency across the day, and, surprisingly found the opposite to be true: consistency increased over the course of each day (Figure 5-6e, $\text{ROIs}_{\text{presentsomedays}} r=.14$ $p=6.1\text{e-}9$, $\text{ROIs}_{\text{presentalldays}} r=.18$, $p=9.7\text{e-}7$). We also found that song-aligned calcium activity became more similar to the previous day's average over the course of the day (Fig 6f, $\text{ROIs}_{\text{presentsomedays}} r=.12$, $p=5.6\text{e-}3$, $\text{ROIs}_{\text{presentalldays}} r=.25$ $p=4.3\text{e-}9$, Spearman rank correlation). Thus, within each ROI, song-aligned calcium activity changes significantly over a night's sleep, and then recovers in both self-consistency and in similarity to previous day's average as the day progresses. Given these observations, the drift in calcium activity is not simply a function of time elapsed. Rather, these results suggest that sleep actively

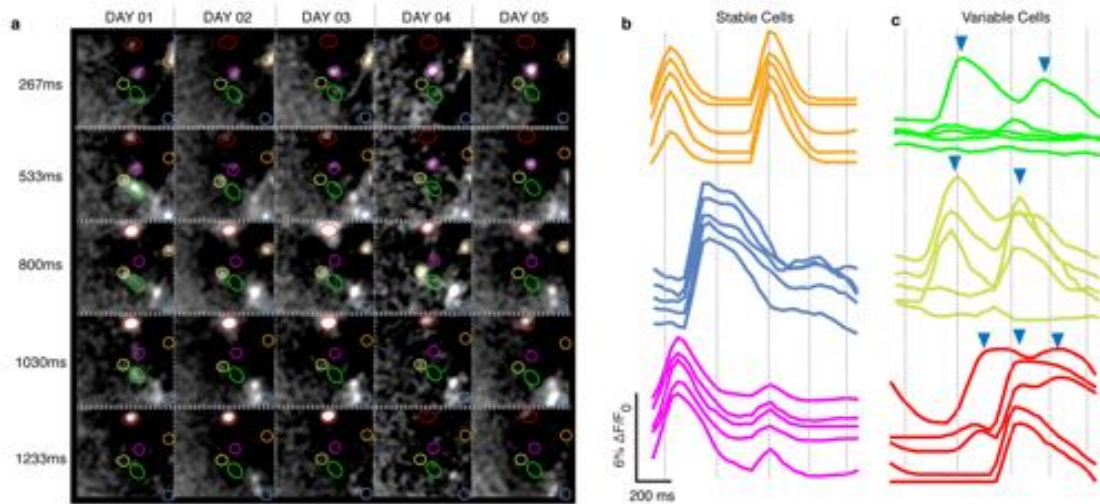


Figure 5.7: Examples of amplitude traces from multi-peaked traces. **A**, Single frame stills from the trial-averaged, song-aligned calcium imaging movies, across all 5 days of a longitudinal study. Columns are days, and rows are frame times relative to the start of song. **B**, Example of three cells with stable timing and amplitude. Five traces represent five days (ordered top to bottom from day 1 to day 5.). **C**, Traces of three cells that are unstable across days, with triangles indicating calcium peak times. Dashed lines indicate times corresponding to each frame from **A**.

destabilizes the representation of song in pre-motor cortex area HVC.

5.2.4 Close inspection reveals the fine microstructure of song also changes over intervals of sleep and increases in consistency throughout the day

The finding that projection neuron activity is unstable across days prompted us to carefully quantify the microstructure of song behavior using a highly-sensitive time-frequency analysis method developed for analysis of sparse signals such as zebra finch song. The persistence of zebra finch song structure has been documented over timescales of years (Immelmann, 1969; Lombardino and Nottebohm, 2000b) and as demonstrated in Figure 5.1, over the period of a year, the song appears to be remarkably stable (Figure 5.1A); however, the micro-structure of song is not precisely the same, even in this example. High resolution investigation of the microstructure reveals small but significant shifts over the interval of a week, and these changes occur

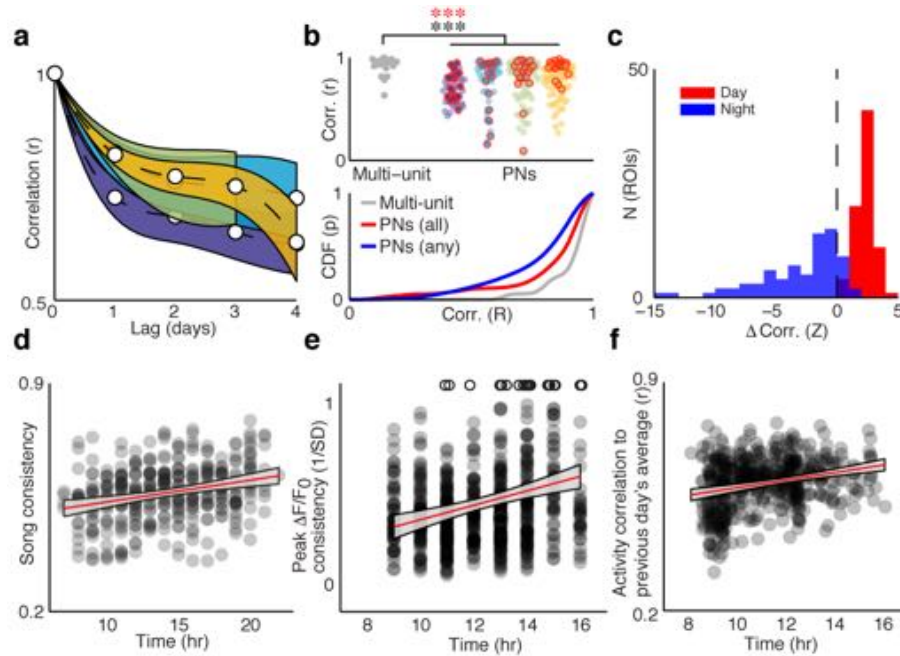


Figure 5-8: Projection neuron activity preferentially changes over periods of sleep.

A, Average correlation across all ROIs across different lags for each bird. Shaded region indicates 95% bootstrap confidence interval. **B**, *Top*, All correlations up to a lag of 4 days for multi-unit activity ($n=20$ birds, one sample per bird) and projection neuron activity ($p=2.7e-6$, $z=4.70$, one-tailed Wilcoxon ranksum test). Highlighted in red are cells that participate unambiguously across the entire time series ($p=9.8e-4$, $z=3.30$ one-tailed Wilcoxon ranksum test). *Bottom*, Cumulative density functions (CDF) of the data shown above: multiunit data in gray, projection neurons seen across all days in red, and seen clearly on any day are blue. **C**, The shift in correlation occurs primarily overnight (***, $ROI_{presentsomedays}$ $p=2.0e-40$, $z=-13.26$, $ROI_{presentalldays}$ $p=1.59e-14$, $z=-7.59$, sign-rank test; shown are data from $ROI_{presentalldays}$). **D**, Song consistency increases over the course of the day ($r=.28$, $p=6.6e-8$, Spearman rank correlation). **E**, Consistency in the timing of peak $\Delta F/F_0$ (measured over the whole song) increases with time of day ($ROI_{presentsomedays}$ $r=.14$ $p=6.1e-9$, $ROI_{presentalldays}$ $r=.18$ $p=9.7e-7$, Spearman rank correlation, shown are data from $ROI_{presentalldays}$). **F**, The trial by trial correlation between calcium activity and the previous day's average also increases with time of day ($ROI_{presentsomedays}$ $r=.12$ $p=5.6e-3$, $ROI_{presentalldays}$ $r=.25$ $p=4.3e-9$, Spearman rank correlation; shown are data from $ROI_{presentalldays}$). All time of day measurements in E and F use only data our custom microscopes detailed in Figure 5-4 ($n=3$ birds).

primarily over periods of sleep ($p=9.8e-4$ $w=55$, one-tailed Wilcoxon ranksum test). Moreover, similar to the drift in calcium activity, we found that song consistency also increases over the course of the day (Figure 5-8d, $r=.28$, $p=6.6e-8$, Spearman rank correlation).

5.3 Discussion

This study suggests that memories encoded in cortex, like many structural features of the human body, undergo renewal at the cellular level. Skin maintains its shape despite the turnover of cells (Weinstein and van Scott, 1965), and the intestinal surface has constant turnover but maintains its function (Creamer et al., 1961). For the stable song of the zebra finch, instability of the neural program qualitatively matches the microscopic day-to-day changes in song, but the song changes observed over intervals of sleep are small, while the pattern of neuronal participation changes more dramatically. What explains the persistence of the song motor pattern in spite of the unstable projection neurons underlying song?

A key finding is that on a mesoscopic scale, the premotor cortical area HVC is stable. Over the length scales recorded in LFPs and multi-unit recordings (approximately 100 microns), the neural basis of song is largely unchanged over a timescale of weeks to years. In contrast, individual projection neurons drift: not primarily in the timing of their activity, as much as in the burst probability and participation in the song pattern. One potential factor contributing to ensemble stability in HVC is the spatial correlation observed in the firing patterns of excitatory cells. Nearby excitatory neurons in HVC fire at similar times, and the length-scale of this correlation is roughly 100 microns (Markowitz et al., 2015). If individual projection neurons fade out of the ensemble, nearby projection neurons could be serving redundant roles.

A second related factor contributing to ensemble stability is the role of temporally patterned inhibition in HVC. In contrast to the instability of excitatory neuron coding, we have observed that inhibitory interneuron firing patterns are stable for weeks or months; for as long as it was technically possible to track the firing patterns of the cells. At the simplest level, excitatory neurons are postulated to be driven by a relatively sparse number of strong synapses, leading to the sparsity of their firing

patterns (Diesmann et al., 1999; Long et al., 2010). Support for this prediction has been found from serial electron microscopy reconstruction of HVC (Benezra et al., 2015). In contrast, many inhibitory interneurons in HVC have large dendritic arbors and are thought to be driven by a large number of presynaptic partners (Mooney and Prather, 2005b). The prediction from HVC connectivity is that changes in a small number of synapses could alter the excitatory neuron firing patterns dramatically, but influence the inhibitory interneurons very little. In light of this numerical argument, and the observed stability of single inhibitory neurons, HVC ensemble patterns could be stabilized if interneurons have a strong local influence over projection neuron activity. Indeed, this appears to be the case. The firing times of the excitatory projection neurons coincide with stereotyped pauses in local inhibition (Amador et al., 2013; Kosche and Long, 2014), and HVC_X excitatory neurons and inhibitory interneurons fire in opposite phases of the 30Hz LFP (Markowitz et al., 2015). The observations are true on the micro-scale, and due to phase shifts across the extent of HVC the rhythmic alternation is not observed globally (Lynch et al., 2016). Local blockade of inhibition releases the sparse firing cells from inhibition and they begin to fire at multiple times in song (Kosche and Long, 2014). Building on these observations, a recent modeling study suggested that inhibitory interneuron dynamics can increase the robustness of HVC dynamics to added noise (Cannon et al., 2015). These models can be made more precise in future experiments as additional cell type specific information is measured. The relative stability of different classes of HVC excitatory neurons (HVC_{RA} , HVC_X and HVC_{Av}) remain to be examined since our methods could not distinguish among the three cell types. In the songbird, stable patterned inhibition may be an important force in maintaining the dynamical pattern of song in spite of underlying drift in the projection neurons. For mammals, similar principles may apply. Aspects of the song corticothalamic loop resemble features of

mammalian motor cortex, including an important 30 Hz rhythm (Markowitz et al., 2015; Murthy and Fetz, 1992; Rubino et al., 2006) and spatial correlations in neuron firing patterns over 100 μm length scales (Dombeck et al., 2009; Hira et al., 2013; Peters et al., 2014). In mammalian motor cortex, it will be interesting to track drift in excitatory versus inhibitory neuron populations to see if patterned inhibition may provide a stabilizing force in motor skill persistence.

Finally, we address the importance of sleep in the rearrangement of the song code in HVC. The change in the song motor pattern occurs primarily over intervals of sleep, and we have shown that adult song behavior also undergoes microscopic shifts in sleep. Our results may be related to a previous observation: as juveniles learn to sing, their songs progress through a day of practice (Ohgushi et al., 2015), but degrade slightly over intervals of sleep (Deregnacourt et al., 2005). The depth of this overnight 'backsliding' is positively correlated with the eventual success of song learning, indicating that nighttime song rearrangements are important for learning. In addition, if song-aligned HVC activity is disrupted through lesioning Nif, an area upstream of HVC, the pattern recovers primarily overnight (Otchy et al., 2015). The present study provides a possible neural mechanism for these sleep effects, raising the possibility that new patterns of activity 'invented' over intervals of sleep provide important raw material for song learning and maintenance. Francis Crick proposed that noisy reactivation of neural circuits in sleep weakens the strongest pathways in the brain, promoting adaptive plasticity (Crick and Mitchison, 1983). Songbirds reactivate their song patterns in pre-motor cortex in a noisy or incomplete manner (Dave and Margoliash, 2000) and it is possible that this spontaneous activity drives the overnight shift in HVC excitatory neurons observed here. Future studies can track the firing patterns of excitatory neurons in HVC while silencing or overstimulating spontaneous activity in sleep, directly testing the hypothesis that noisy replay of song

in sleep promotes adaptive plasticity of the song motor program. It remains to be seen whether the overnight shift in HVC is random, or guided by vocal errors. Random shifts in the population could increase robustness of HVC dynamics by enforcing redundant representations of song. In contrast, shifts in the population that are influenced by vocal performance errors could be an active part of song learning and maintenance.

5.4 Methods

5.4.1 Subjects

All procedures were approved by the Institutional Animal Care and Use Committee of Boston University (protocol numbers 14-028 and 14-029). Electrophysiology data were collected from $n=27$ adult male zebra finches (>120 DPH), and imaging data were collected from $n=4$ adult male zebra finches. Birds were kept on a 14 h light-dark cycle.

5.4.2 Splaying microfiber electrodes

We used a previously-described minimally invasive carbon fiber electrode array (Guitchounts et al., 2013) in addition to commercially available arrays (TDT, Neuronexus) to chronically monitor both single units and LFPs. Extracellular voltages were amplified and digitized at either 25 or 30 KHz using the Intan acquisition system (RHA2000 and RHD2000).

5.4.3 Microfiber electrodes for multi-unit recordings

Bare (i.e. uncoated) microthread electrodes were prepared by extracting 5 and 11 μm diameter fibers from commercially available spools of carbon fiber (grade XAS, HTA, T300 or P25). Once extracted, sizing and other impurities were removed by heating fibers as previously described (Guitchounts et al., 2013). The bare fibers

were then attached to coated silver wire using a conductive silver paint (842-20G, MG Chemicals). Signals were sent from a custom headstage to a differential amplifier (AM Systems 1700, gain of 1000, low cut-off 300 Hz, high cut-off 5 kHz), and digitized with a National Instruments Acquisition board (PCIE 6323, 40 kHz sampling rate).

5.4.4 Tracheosyringeal nerve cut

The right TS nerve was removed using previously described methods (Williams et al., 1992). Briefly, birds were anesthetized with 4% isoflurane and maintained at 1-2% for the course of the surgery. Feathers were removed from the neck, and an incision was made over the trachea. The nerve was dissected from the surrounding tissue and the nerve was pulled from its roots on the syrinx, extracting a minimum of 1 cm of nerve length.

5.4.5 Calcium imaging

To image calcium activity in pre-motor cortical neurons during singing, we employed head-mounted fluorescence microscopes, both commercial and custom built. This method provides long-term recordings of calcium signals in HVC (Markowitz et al., 2015), and enables studies of motor stability and adaptive plasticity at the single neuron level. For viral delivery, we use the genetically encoded calcium indicators GCaMP6s (2 birds) and GCaMP6f (2 birds) delivered by lentiviruses (Markowitz et al., 2015; Wilson et al., 2008).

For 1 bird, we used a commercial microscope (Inscopix) to gather female-directed singing over 5 day periods. However, these microscopes could not be used with a rotary commutator and were too heavy to evoke reliable undirected or solo song. Motivated by these challenges, we developed a light-weight (1.7g), commutable, 3D printed single-photon fluorescent microscope that simultaneously records audio and video Figure 5-4, (n=3 birds). These microscopes enabled recording of hundreds of

songs per day (all birds sang at least 500 songs on at least one day of recording), and all songs were recorded from birds longitudinally in their home cage, without requiring adjustment or removal of the microscope during the imaging period. Birds were imaged for less than 20 minutes total on each imaging day, and LED activation and video acquisition were triggered on song using previous described methods. All analysis was restricted to a particular bout in song, either the first (1 bird) or the second (3 birds). Recordings were taken from 3 weeks to 3 months post-injection.

5.4.6 Virus Information

Addgene plasmids # 40753 (pGP-CMV-GCaMP6s) and # 40755 (pGP-CMV-GCaMP6f) (gift of the Douglas Kim Laboratory) were transformed into E.Coli bacteria by heat shock, and sequenced. The GCaMP6s and GCaMP6f fragments were PCR amplified with the addition of NotI/BamH1 to the 5' and 3' ends respectively. GCaMP6s was cloned into the pHAGE-CMV-eGFP vector to form pHAGE-CMV-GCaMP6s. The RSV promoter sequence was ordered from IDT as a gblock with the addition of SpeI/NotI to the 5' and 3' ends respectively and was cloned into pHAGE-CMV-GCaMP6s and pHAGE-CMV-eGFP to form pHAGE-RSV-GCaMP6s and pHAGE-RSV-eGFP. GCaMP6f was then cloned into pHAGE-RSV-GCaMP6s to form pHAGE-RSV-GCaMP6f. The viruses were packaged in HEK 293T cells and titered on FG293 cells, with titers ranging between $1.2 - 2.3 \times 10^{10}$ infectious particles/mL. Plasmid maps and sequences of all lentiviral vectors employed can be downloaded from the vectors page of the laboratory of Darrel Kotton:

The tropism of our virus for excitatory neurons was evaluated by counterstaining using known markers of inhibitory neurons, specifically Anti-Parvalbumin (Abcam Ab11427), Anti-Caretinin (SWANT 7697), and Anti-Calbindin (SWANT CB-300). All observable GCaMP6 labeled dendrites contained spines, consistent with the morphology of HVC_{RA} and HVC_X projection neurons. Out of 1000 counted neurons,

only 1 was double-labeled with inhibitory markers.

5.4.7 Song alignments

Songs were aligned using the Euclidean distance in spectral features between the data and a template song in a sliding window. Local minima in the Euclidean distance were considered candidates hits, which were then plotted in 2 dimensions for the user to perform a cluster cut. No time warping was applied to any data. These methods have been described in more detail previously (Poole et al., 2012).

5.4.8 Spectral density images

To plot the variability of multiple song renditions in time and frequency we used the spectral density image, which has been described previously (Markowitz et al., 2013b, 2015). A sparse binary time-frequency representation was generated using auditory contours (Aoi et al., 2015; Lim et al., 2013). These time-frequency binary images were combined by averaging across all renditions. In the resulting image, the value at each pixel is the probability that a time-frequency contour passes through it.

5.4.9 Acoustic change post-nerve cut

To analyze the change in song post-TS nerve cut, we computed amplitude modulation (AM), frequency modulation (FM), Wiener entropy, amplitude envelope, pitch and pitch goodness using Sound Analysis Pro for MATLAB (Tchernichovski et al., 2000). We then computed the average or standard deviation of each feature across the entire song and compared the last day of singing pre-cut with the first day of singing post-cut.

5.4.10 Acoustic change overnight

To quantify changes in song microstructure we used the similarity score derived from the spectral density image, which has been described previously (Markowitz et al., 2013b). For 5 consecutive days of singing from $n=10$ birds, we divided each day's worth of singing in half by trial number. To compute the change in song microstructure across the day, we computed the similarity scores between trials from the first half of the day and the spectral density image from the second half of the day. Then, to analyze the overnight change we computed similarity scores between trials from the first half of each day and the spectral density image from the previous evening. Finally the scores were averaged for each day prior to statistical comparison. Next, to estimate the change in song consistency over the course of the day (Figure 5-6d), we computed the spectral density image for each hour's worth of singing. Then, similarity scores were calculated for each trial and the corresponding spectral density image. Trends were estimated by taking the Spearman rank correlation between similarity score and the time of day.

5.4.11 Local field potentials and single units

Local field potentials (LFPs) were analyzed as described previously (Markowitz et al., 2015). Extracellular voltage traces were median filtered (1 ms window) to remove spikes and then lowpass filtered with 400 Hz corner frequency (4th order Butterworth filter) and downsampled to 1 kHz. Finally, the LFP was filtered with a 25-35 Hz bandpass (8th order Elliptic filter for Figure 5-1; 53 tap Kaiser window FIR filter, 20 dB stopband attenuation, .05 ripple for all analysis to minimize the impulse response). To compute the change in phase angle, we used the angle of the Hilbert-transformed narrowband LFP.

Single interneurons were analyzed as described previously (Markowitz et al., 2015).

Extracellular voltage traces were band-pass filtered from 600-11 kHz (12th order Elliptic filter, .2 dB passband ripple, 40 dB stopband attenuation) and sorted using standard offline spike sorting techniques (Sahani, 1999; Tolia et al., 2007).

5.4.12 Multi-unit electrophysiology

First, threshold crossings were taken from extracellular traces digitized at 40 kHz using a threshold of 2.5 robust standard deviations (Quiroga et al., 2004). Threshold crossings were then converted into firing rates on a single-trial basis by convolving with a Gaussian kernel (5 ms sigma). Alternatively, the root-mean-square (RMS) of the voltage signal was calculated in a 5 ms sliding window (box car). To estimate the stability of multi-unit activity, we averaged either the song-aligned firing rate or RMS across all trials in a given day. Then, the Pearson correlation coefficient was computed between the averages from either the first day of recording (or the last day pre-TS cut) and each subsequent day (Figure 5-2I-J). To compare the drift post-TS cut to the baseline condition, correlation values were binned in a 16 day sliding window (5 day overlap, the result was insensitive to binning parameters, data not shown), and a Wilcoxon ranksum test was performed between the binned pre- and post-nerve cut correlation values.

5.4.13 Calcium imaging

Time plots used in Figure 5-5a were created by averaging all song-aligned calcium imaging movies for a bird within a single day. The resulting 'average' movie was smoothed with a 15 μm gaussian filter, and each pixel was then colored by its center of mass in time. To create presence or absence images used in Figure 5-5c, Maximum projection images were created for all song-aligned calcium imaging movies, and the average pixel intensity of these maximum projections was taken for each day of imaging. Each day is mapped to the red, green, or blue. For Figure 5-6c, each maximum

projection image is divided by a smoothed version of the same image, using a $100\ \mu\text{m}$ pixel disk filter, to normalize across bright and dim ROIs.

Region Of Interest (ROI) based analysis was performed as described previously (Markowitz et al., 2015). In brief, raw imaging data was motion corrected using a previously published algorithm (Guizar-Sicairos et al., 2008). ROIs were manually selected and pixels were averaged for each frame within each ROI. ROI traces were converted to $\Delta F/F_0$ by estimating F_0 as the 12th percentile in an 800 ms sliding window. Drift was assessed by computing the Pearson correlation between trial-averaged calcium traces for each ROI at all possible lags. Correlation values from the same lag (e.g. between Day 1 and 2, Day 2 and 3) were averaged for each ROI (a similar procedure was used for the multi-unit data comparison in Figure 5-6b). To account for any uncertainty in the alignments due to the 30 Hz sampling rate of the camera, we then computed the maximum Pearson correlation in a 100 ms window. Drift was then quantified using a randomization permutation test with a $p=.01$ threshold. Specifically, we compared the correlations observed at each lag (1-4 days) with correlations from the same data with the group (i.e. day) labels scrambled. P-values were determined by estimating the probability that the observed correlation was greater than or equal to the correlation values from over 10,000 randomizations. We also repeated this analysis using the timing of peak $\Delta F/F_0$ on each day (not shown). Specifically, for each ROI the time of peak trial-averaged $\Delta F/F_0$ was computed on each day, and the peak times were compared between each day of imaging and the first. For each cell, if the peak timing changed by more than 100 ms or there were no peaks greater than .5% $\Delta F/F_0$, the cell was considered unstable.

In order to compare within-day to overnight changes in correlations, we split the data on each day in halves by trial number. The within-day correlations were measured by computing the correlations between trial-averages using the first and

second half of trials from a given day. The overnight correlations were measured by computing the correlations between trial-averages from the second half of one day with the first half of the subsequent day. The within-day and overnight correlations were averaged across days for each ROI, and subsequently a sign-rank test was used to compare the two groups of correlation values. For visualization (Figure 5-8c) the two populations were Z-scored using a bootstrap.

To test for significant interactions between time of day and the consistency of calcium activity, we first computed the amplitude of peak $\Delta F/F_0$ across the entire song for each ROI (this controlled for any change in song duration across the day). Next, we computed the standard deviation of the peak $\Delta F/F_0$ in 1 hour bins and computed the Spearman rank correlation between these values and time of day. To account for changes in peak $\Delta F/F_0$ due to photobleaching, we computed the partial rank correlation between the time of day and standard deviation of the peak $\Delta F/F_0$ after accounting for the correlation between time of day and mean of the peak $\Delta F/F_0$ (again using 1 hour bins). Then, to analyze the similarity of calcium activity to the previous day's average, we peak-normalized the $\Delta F/F_0$ values for each trial through dividing by the median peak $\Delta F/F_0$ determined using the 5 nearest trials (with respect to time of day). Then, we estimated the similarity between each trial and the previous day's average using the peak Pearson correlation coefficient (also over a 100 ms window). Finally, we computed the Spearman rank correlation between these similarity values and time of day.

5.4.14 Controls for stability of the imaging plane and photobleaching

Across days, ROIs were manually inspected and adjusted to insure that the center of mass for each ROI was at the center of the ROI mask, and that the mask did not overlap with neighboring cells. Masks that were moved more than the diameter of the mask at any point of the longitudinal study were excluded from analysis. To

ensure that our results were not impacted by the stability of the imaging plane, we then checked to see if there were any differences in the spatial profile of stable and unstable ROIs. More precisely, we computed the $\Delta F/F_0$ -weighted centroid within each ROI for each day. Then, we took the Euclidean distance between the $\Delta F/F_0$ weighted centroids for each ROI across adjacent days of imaging. Lastly, the distances for unstable and stable ROIs (determined using the permutation test described above) were compared and we found no significant effect (ROIs_{resentsomedays} $p=.17$ $z=.97$, ROIs_{resentalldays} $p=.99$ $z=-4.86$; one-tailed Wilcoxon ranksum test). Moreover, to control for any effects of bleaching in our time of day analysis we repeated the analysis of calcium activity consistency using the Spearman rank partial correlation accounting for correlation between time of day and average peak $\Delta F/F_0$. A significant effect was still observed (ROIs_{resentsomedays} $r=.12$ $p=7.8e-7$, ROIs_{resentalldays} $r=.17$ $p=2.8e-6$).

5.4.15 Statistical analysis

No formal methods were used to predetermine sample sizes, the sample sizes used here are similar to those used in the field. All statistical comparisons were performed using non-parametric tests (Wilcoxon ranksum, sign-rank, bootstrap, or randomization tests), with the exception of the correlation between variability computed within and between days (Pearson correlation, the result was verified with the Spearman rank correlation). Where appropriate, we controlled for multiple comparisons using the Holm-Bonferroni step-down procedure.

Chapter 6

Conclusions

6.1 Potential cellular mechanisms for reinforcement learning

The results described in *Chapter 4* & *Chapter 5* indicate that HVC operates on two distinct levels: a mesoscopic level that is highly stable and a microscopic level where individual neurons are free to change their participation. It is possible that HVC may represent learned behaviors in the stability of network patterns that persist in spite of the drifting participation of individual neurons. The instability of the neural program could reflect the microscopic day-to-day changes in song, but the difference in magnitude raises the question: If this drift is truly random, why does the song not undergo larger shifts? What explains the persistence of the song motor pattern in spite of the unstable projection neurons underlying song? What drives the instability of single neurons? Potentially, this fine scale plasticity we observe may reflect an underlying mechanism that is important for maintaining the stability of the motor behavior over time. It is possible that fine-scale drift may somehow actively serve as a mechanism for memory maintenance and motor stability.

Given that HVC receives inputs from the auditory system (Akutagawa and Konishi, 2010; (Vates and Nottebohm, 1995), it is an accommodating location for sensorimotor integration. HVC sits atop both the AFP and the VMP, evaluations of prediction-errors here would provide a convenient mechanism to mediate sensory-motor learning- and could provide the site of active maintenance of the song motor program throughout adulthood. However, many experimental findings have reduced

the likelihood of this being true. Deafened birds do not show significantly altered HVC responses during singing. In the canary, auditory responses cannot be elicited in HVC until seconds after song has ended. Intracellular recordings in HVC_X neurons have failed to find even sub-threshold changes in response to auditory stimuli that overlap with song production. (Hamaguchi et al., 2014). Together, this suggests that HVC projection neurons likely have no real-time mechanisms for updating the motor code in response to sensory feedback. However, our findings provide valuable insight- the changes we observe typically do not occur during the day, but discretely over periods of sleep. This suggests that there may be 'offline' mechanisms that update the motor plan. The premotor nucleus HVC might still serve as a structure where auditory feedback is ultimately integrated by constant fine-tuning of song motor commands, where individual cells adapt their firing times to regulate the behavioral output of song after the accumulation and evaluation performance errors across the course of a day. In this way, drifting excitatory neurons may be hedonistic; changing their activity based on a history of activity that led to 'good' performances, stabilized by mesoscopic inhibition (Zagha, Ge, and McCormick 2015). However, these changes may simply be the result of intrinsic homeostatic mechanisms that support network stability.

In this concluding section, I propose experiments that could address if changes in neural participation is important to memory stability or maintenance. A key question is whether drift at the single neuron level is influenced by an animal's perceived behavioral performance. These experiments hold the potential to reveal single neuron rules underlying sensory-motor learning and address long standing questions about the nature of memory stability for skilled movements.

6.2 Engineering limitations, and future directions.

6.2.1 Identifying the downstream projection targets of projection neurons during imaging

How do different projection neuron classes contribute to the stability of the motor program? One of the key limitations of the imaging studies performed in the previous chapters is that we could not segregate the downstream projection targets of the neurons we recorded from. While HVC projection neurons all appear to be discretely active at one or a few moments of song, these neurons are a heterogeneous population. They differ greatly in their morphology, and send their axons to opposite ends of the brain. HVC_X and HVC_{RA} possibly send independent computational contributions to the AFP and the VMP- if drift occurs in one population but not the other, it would have very different implications. It is clearly important to undertake experiments that can identify cell-type specific contributions to neural coding. There are two major challenges. First, we lack a precise way of genetically labeling distinct HVC projection neuron subpopulations. Secondly, we are limited in both the resolution and the bandwidth of our miniature head-mounted microscopes. I briefly describe how to address these challenges below.

Genetic Labeling of specific subpopulations

While the zebra finch genome has been sequenced and several transgenics have been generated (Scott and Lois, 2005), viral transduction in the songbird remains a challenge, and there are limited genetic tools currently available. While excitatory neurons in HVC can be preferentially targeted for genetic manipulation, it is still not possible to target specific sub-populations using cell type specific promoters. However, the nucleated structure of the songbird brain makes it ideal for intersectional strategies using retrograde labeling. Preliminary data suggests that it is possible to selectively

image from populations of cells of a known origin using retrograde viruses (Figure 6.1). New advances in blood based viruses and advanced retrograde viruses could spur advances as well.

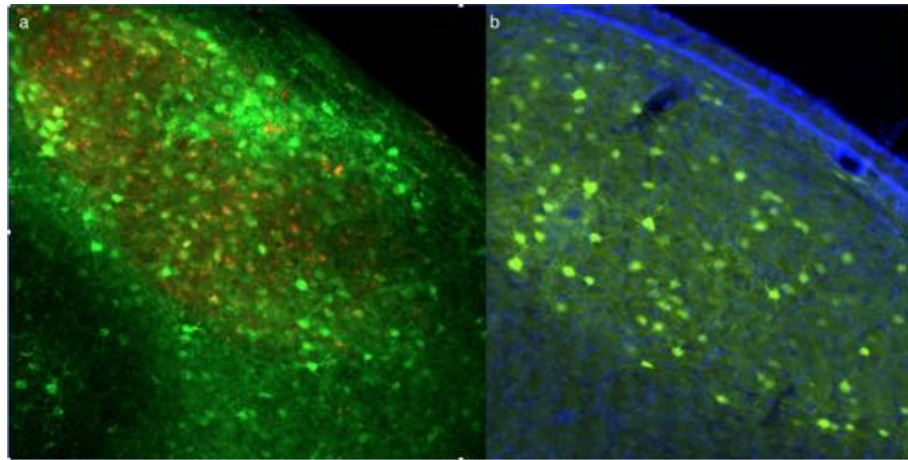


Figure 6.1: Cell-type specific genetic labeling *a, Left*, Lentiviral infection of a heterogeneous mixture of neurons in HVC *Right*, HVC-X labeling using retrograde viruses. Red = DiI, Green = GCaMP6, Blue = DAPI

Even without adopting new genetic tools, one could simultaneously image a retrogradely injected dye (for example, DiI; which has green excitation and red emission) and the GCaMP6, so that projection neuron subtypes in songbirds can be identified, based on the presence or absence of retrograde labeling from downstream nuclei. However, current miniscope designs can only image in one band/wavelength (In our case, we use blue excitation and green emission).

Multicolor Imaging

Single photon imaging with miniature head-mounted fluorescence microscopes has become an increasingly widespread and powerful method for recording neural activity in freely moving animals at the cellular-level. In our hands, it has been indispensable in describing the stability and spatial organization of single projection neurons- but in many ways, this technique still remain relatively limited. Due to bandwidth

constraints, it is not possible to use different fluorophores to segregate signals by cell type or projection target. One way to address this is to make modifications to incorporate two or more independent color channels. For example, microscopes could be designed with multi-peaked excitation LEDs and color cameras, and paired with the appropriate filter sets that accommodate multi-color imaging in blue/green and green/red. This strategy could increase the bandwidth of information gathered within the imaging plane- and allow the use of additional fluorescent proteins or tracers to disambiguate specific neuron types within an imaging field (Figure 6·2).

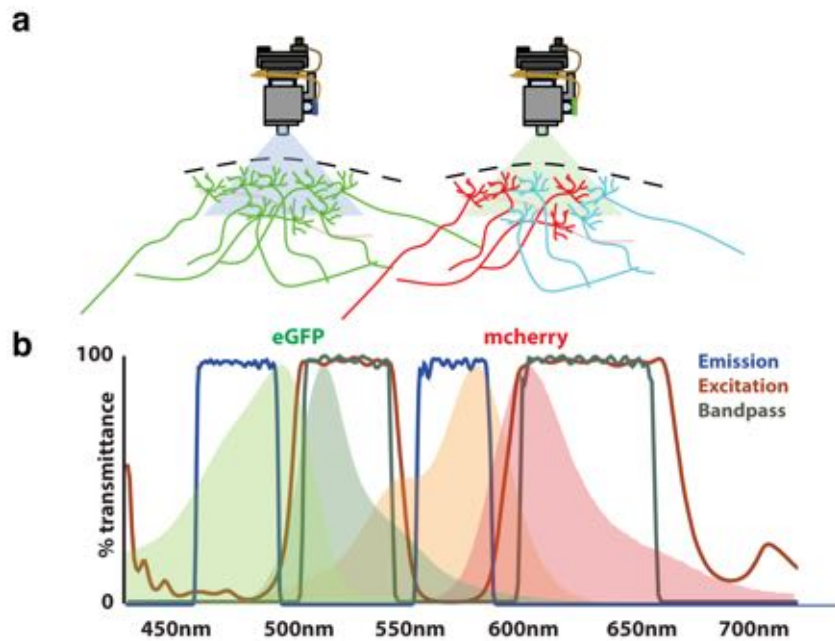


Figure 6·2: Multi-wavelength imaging in a 1P miniature microscope. A dual color microscopes can be used to disambiguate populations of neurons using a combination of genetically encoded reporters, or injected dyes. B Dual pass band filter sets are commercially available, and are well suited to this application.

6.2.2 Structured Illumination Microscopy

Another limitation to single photon imaging is that the poor z-axis resolution can make it difficult to disambiguate signals coming from overlapping cells. While com-

putational techniques have emerged that are designed to de-mix overlapping neurons, these techniques are still highly parameterized and need to be optimized per preparation (Picardo et al., 2016b). To date, the axial resolution of bench-top multi-photon microscopy is still vastly superior to that which can be achieved otherwise. In spite of these limitations, head-mounted microscopes are often the only way to optically observe neural populations during naturalistic behaviors. In the songbird, we have observed that head-fixation is an extremely stressful experience. Under this condition, birds rarely sing. To promote song production, birds are deprived of water and social exposure to females- and these are used as an incentive to sing. As a result, this paradigm likely impedes the study of 'undirected' singing in songbirds, a learning-intensive form of song practice in the zebra finch. We have evidence that excitatory neuron activity during the production of undirected song can differ from directed, and that network exploration during song 'practice' may be important. (Figure 6-3)

Compared to these laser scanning techniques, single photon imaging has relatively poor depth of field, z-axis sectioning, and tissue penetration. These issues can potentially be addressed using optical and computational techniques that improve signal quality by rejecting unwanted noise. The most promising of these techniques is Structured Illumination Microscopy (SIM). The increase in resolution from SIM is obtained by providing non-uniform excitation light patterns that feature sinusoidal intensity variations in one or more dimensions coupled with powerful image reconstruction techniques. This allows the rejection of out of focus light from outside of the image plane, at the expense of decreasing the effective frame-rate due to the computational step. Using micro LEDs and/or DMD coupled laser light deliver through coherent fiber bundles, SIM can be delivered through a custom designed miniature microscope. In preliminary experiments, a Digital Micro-mirror Device (Texas Instruments LightCrafter) is used to create three grid patterns that are shifted by a

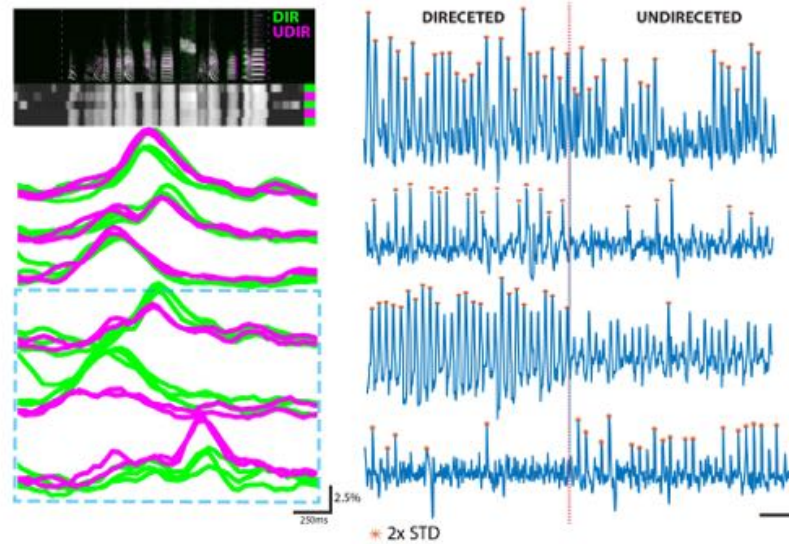


Figure 6-3: Context related changes in HVC projection neurons Head-fixing animals may preclude the study of differences in motor coding that relate to the behavioral context of the animal. We see some evidence that motor codes that underlie song that is 'directed' to a female may differ from when song is produced in isolation. *Left*, Several ROIs aligned to song, from interleaved trials of directed and undirected song. Blue dotted-line inset highlights several ROIs the show some differences related to behavioral context. (Green = Directed, Magenta = undirected) *Right*, fluorescent time series data from several ROIS, concatenated from many songs in each context. Differences in peak participation frequency in each context suggest that variabilities in coding may exist as a function of behavioral context.

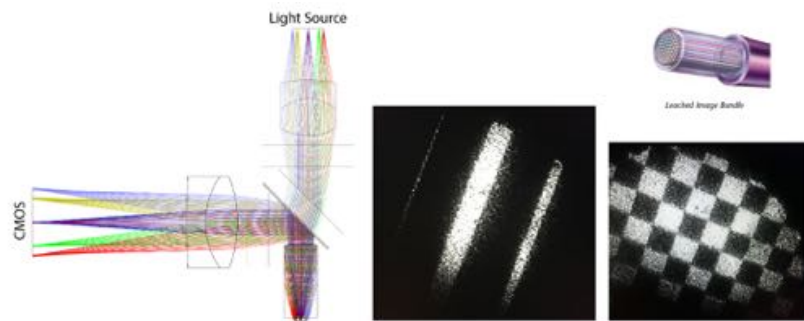


Figure 6-4: Structured Illumination Integration The integration of structured illumination into miniature microscopes would improve z-axis resolution, and could allow for arbitrary or patterned stimulation- creating all-optical interrogation, at the single cell level, in a freely behaving animal. *Left*, Light-path for SIM integration. *Right*, DMD patterns on the coherent fiber.

third of the period. After acquiring the three images, the sectioned image is computed and background fluorescence is rejected. Extensions to this strategy could afford new

ways to control neurons in freely behaving animals, using patterned illumination for optical control of spatially segregated neurons, and potentially could afford 3D volumetric imaging (Szabo et al. 2014). (Ventalon et al. 2015). Proof-of-concept work is currently occurring in an on-going collaboration with Bernardo Sabatini’s lab in Harvard Medical School. Figure 6-4 A fast cycle of ZEMAX optical path modeling and re-printing of the microscope housing has facilitated this collaborative engineering project, and has already demonstrated improvements in the rejection of out-of-plane fluorescence.

6.3 ROI based, near-Real-Time Brain Machine Interface (BMI) experiments

Motivated by the observation that the stability of song memory is rooted in the stability of network patterns that persist in spite of drifting individual neuron dynamics, we hope to explain the persistence of the song motor pattern in spite of unstable projection neurons. It is experimentally possible to directly test the hypothesis that changes in the firing patterns of single neurons in pre-motor cortex is beneficial to the motor program. I propose to explicitly test if the activity of these cells is shaped by the recent history of time-correlated reward or punishment signals, by using real-time Brain Machine Interfacing (BMI) experiments

What is the utility of drifting motor commands in the production of a stable motor behavior? Previous studies using extracellular and intracellular recordings indicate that during vocal production, HVC_X and HVC_{RA} neurons are insensitive to feedback perturbations, suggesting that performance evaluation does not occur in these cells during song. However, spine imaging studies have found that deafening shrinks and destabilizes dendritic spines on HVC_X neurons within 12-48 hr and that these changes precede and predict the severity of song degradation.

These results suggest that at least HVC_X cells are sensitive to auditory feedback on the same timescale that we have observed drifting firing patterns. The descending projections of HVC_X (and possibly HVC_{AV}) possibly relay an efference copy of the motor control sequences that could be evaluated alongside song performance in the basal ganglia. In this case, we might expect some mechanism would exist at the interface of auditory and motor generation systems to consolidate and reconcile this evaluation. Perhaps HVC serves as a structure where auditory feedback is ultimately integrated by constant fine-tuning of song motor commands. Individual cells slightly adapt their firing times to regulate the behavioral output of song after the accumulation and evaluation of performance errors, forming the mechanistic basis of memory maintenance and and motor stability.

This could be tested using a variant of CAF where the auditory feedback is made contingent on the activity of single HVC, LMAN, or Area X neuron (Figure 6·5). The aim would be to steer plasticity in cells that are already under observation. In this paradigm the bird can only escape aversive white noise playback by modulating the firing rate of either a single isolated unit, or group of units (Koralek et al. 2012) Potentially, this drift will be modulated by performance errors delivered through conditional auditory feedback and consolidated over periods of sleep. This drift could be blocked or reduced by selective blockage of the output structures of the basal ganglia.

This use of region of interest (ROI) based feedback could be contingent on one or several neurons, and could be contrasted by feedback contingent on specific times in song in order to tease apart the individual contributions of individual cell types in HVC- although this strategy could be deployed across the song system. It would be critical to characterize the response of different cell types within a day, as well as the responses across days, to this feedback contingency in order to tease apart the

cell-type specific mechanisms of motor adaptation. It would be equally interesting to monitor spontaneous activity during sleep, to observe what offline dynamics may be occurring that may predict changes or reinforce stable cells.

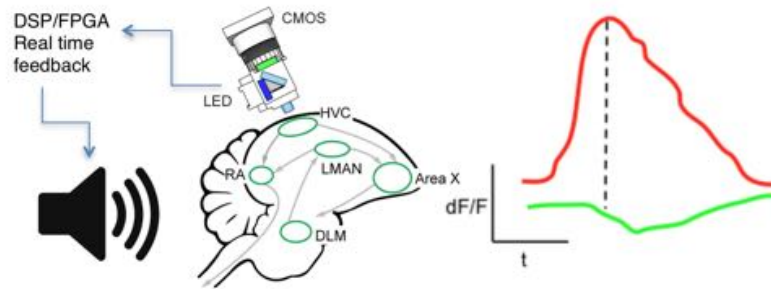


Figure 6-5: BMI experimental outline a, *Left*, To perform experiments that can identify neural activity patterns in real time, the video stream is constantly processed to extract fluorescence information in pixels within predefined regions of interest. Rules (such as the ratio of activity in two ensembles) are applied to this near-real-time ROI analysis to trigger external events through a programmable digital output, such as white noise contingent on the pattern of activity recorded through the head-mounted microscope. In this mode, the camera provides a near-real-time brain machine interface allowing songbirds to control sounds directly through the measured calcium signals in the brain. Other BMI possibilities include closed-loop stimulation experiments that seek to electrically or optically disrupt patterns of activity in real-time.

6.4 The role of sleep

The role of sleep, as it relates to motor systems, is not well-known. Chapter 5 describes changes in neural patterns that discretely occur overnight, raising the possibility that new patterns of activity 'invented' over intervals of sleep provide important raw material for song learning and maintenance. Francis Crick proposed that noisy reactivation of neural circuits in sleep weakens the strongest pathways in the brain, promoting adaptive plasticity. This leads us to ask: Do new strategies of motor exploration occur over periods of sleep, seeking the source of variability and motor-sequence exploration both during periods of song and during 'offline' periods? It is possible that spontaneous activity patterns in sleep allow injection of randomness that

allows for fine grade motor exploration through subtle changes in synaptic weights. This may be visualized in large scale changes seen across days, which may be guided or biased within a day by small populations of auditory and/or context sensitive neurons.

Sleep is widely believed to be important for learning and memory, particularly in the process of consolidation, a process that transforms short-term memories into more stable longer-term memory representations (Kuriyama, Stickgold, and Walker 2004) (Walker and Stickgold 2004) (Stickgold and Walker 2004) (Diekelmann and Born 2010). Many theoretical models propose that consolidation would benefit from sleep, suggesting that sleep prevents the possible interference of daily activities on memory consolidation (Diekelmann and Born 2010). However, experimental evidence explicitly demonstrates that this process is sparse.

Learning new tasks is often associated with stabilization of ensemble activity patterns across days as well as increased task-related neuronal activity. The emerging picture is that neuronal representations are explored by behavioral relevance during the day, but that long-term integration of beneficial changes occurs during offline periods. This 'consolidation' may transform short-term optimal network configurations, that may be inherently fragile and plastic, into more stable and long-term stored memory representations without the interference of online processes (Ferry and McGaugh 2000; Izquierdo and McGaugh 2000) (Eisenberg and Dudai 2004) (Diekelmann and Born 2010) (Diekelmann and Born 2010)

During sensorimotor learning of song in juvenile zebra finches, improvements in structure from the previous day deteriorated over a night of sleep, but through the following day improved and surpassed the previous day's quality. Interestingly, juveniles that showed the most 'backsliding' as a result of sleep, achieved the best tutor song imitation as adults (Dergnaucourt et al. 2005). In addition, spontaneous

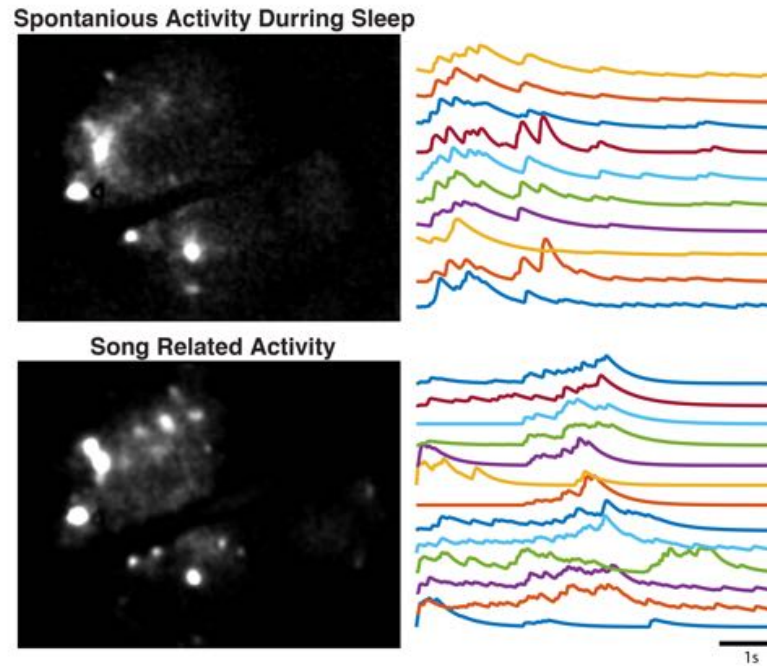


Figure 6-6: Spontaneous activity during sleep *Top*, activity during sleep, contrasted with *Bottom*, song related neural activity.

activity during sleep in juvenile male zebra finches was correlated to the quality of tutor song imitation (Gobes, Zandbergen, and Bolhuis 2010). In the adult, spontaneous neuronal activation during sleep in several parts of the song system appear to resemble motor activation patterns that drive song (Dave and Margoliash 2000). Moreover, there is evidence that noisy or partially incomplete spontaneous sequence reactivations of the motor plan occur during sleep in the motor region RA (Shank and Margoliash 2009). Taken as a whole, there is strong evidence that sleep plays a role in memory formation in songbirds and that this may be critical for the learning and maintenance of song. (Gobes, Zandbergen, and Bolhuis 2010; Gobes and Bolhuis 2008)

Future experiments are needed to robustly answer how sleep is involved in motor stability and planning. I believe that the zebra finch will presents itself as an ideal model system for addressing these questions; the extreme precision and long-term sta-

bility of song structure offers a unique opportunity to observe how motor memories are maintained at the network level- and how these networks are shaped by different states, like sleep. Using cell-type specific genetically encoded calcium indicators and custom miniature head-mounted microscopes, neural populations can be observed over weeks and months. Combining the use of near real-time behavioral and neurally guided feedback mentioned in the previous section, along with round-the-clock monitoring of neural activity will help elucidate the learning roles of individual cells that lead to stable memories.

List of References

- Ahrens, M. B., Li, J. M., Orger, M. B., Robson, D. N., Schier, A. F., Engert, F., and Portugues, R. (2012). Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature*, 485(7399):471–477.
- Ali, F., Otchy, T. M., Pehlevan, C., Fantana, A. L., Burak, Y., and Ölveczky, B. P. (2013). The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron*, 80(2):494–506.
- Amador, A., Perl, Y. S., Mindlin, G. B., and Margoliash, D. (2013). Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature*.
- Andalman, A. S. and Fee, M. S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Science U. S. A.*, 106(30):12518–12523.
- Andrew, A. M. (1998). Reinforcement Learning:983Richard s. sutton, andrew g. barto. reinforcement learning: An introduction. cambridge, MA: Adaptive computation and machine learning series MIT press (bradford book) 1998. xviii 322 pp, ISBN: 0262193981 hardback 31.95. *Kybernetes*, 27(9):1093–1096.
- Aoi, M., Lepage, K., Lim, Y., Eden, U. T., and Gardner, T. J. (2015). An Approach to Time-Frequency Analysis With Ridges of the Continuous Chirplet Transform. *IEEE Transactions on Signal Processing*.
- Aronov, D., Andalman, A. S., and Fee, M. S. (2008). A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science*, 320(5876):630–634.
- Aronov, D. and Fee, M. S. (2012). Natural changes in brain temperature underlie variations in song tempo during a mating behavior. *PLoS One*, 7(10):e47856.
- Ashmore, R. C., Renk, J. A., and Schmidt, M. F. (2008). Bottom-up activation of the vocal motor forebrain by the respiratory brainstem. *Journal of Neuroscience*.
- Azim, E., Jiang, J., Alstermark, B., and Jessell, T. M. (2014). Skilled reaching relies on a v2a propriospinal internal copy circuit. *Nature*, 508(7496):357–363.
- Barth, A. L. and Poulet, J. F. A. (2012). Experimental evidence for sparse firing in the neocortex. *Trends in Neuroscience.*, 35(6):345–355.

- Bartolo, R., Prado, L., and Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *Journal of Neuroscience*.
- Benezra, S., Kornfeld, J., Narayan, R. T., Oberlaender, M., Denk, W., and Long, M. A. (2015). Connectomic analysis of local HVC circuitry in the zebra finch. In *Society for Neuroscience*.
- Bertram, R., Daou, A., Hyson, R. L., Johnson, F., and Wu, W. (2014). Two neural streams, one voice: Pathways for theme and variation in the songbird brain. *Neuroscience*.
- Brainard, M. S. and Doupe, A. J. (2001). Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. *Journal of Neuroscience*, 21(7):2501–2517.
- Buzsáki, G., Anastassiou, C. A., and Koch, C. (2012). The origin of extracellular fields and currents - EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*.
- Cannon, J., Kopell, N., Gardner, T., and Markowitz, J. (2015). Neural Sequence Generation Using Spatiotemporal Patterns of Inhibition. *PLoS Computational Biology*.
- Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., and Kim, D. S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*.
- Colonnese, M. T., Stallman, E. L., and Berridge, K. C. (1996). Ontogeny of action syntax in altricial and precocial rodents: Grooming sequences of rat and guinea pig pups. *Behaviour*, 133(15):1165–1195.
- Crandall, S. R., Aoki, N., and Nick, T. A. (2007). Developmental modulation of the temporal relationship between brain and behavior. *Journal of Neurophysiology*.
- Creamer, B., Shorter, R. G., and Bamforth, J. (1961). The turnover and shedding of epithelial cells Part I The turnover in the gastro-intestinal tract. *Gut*.
- Crick, F. and Mitchison, G. (1983). The function of dream sleep. *Nature*.
- Crowe, D. A., Zarco, W., Bartolo, R., and Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *Journal of Neuroscience*.
- Csicsvari, J., Jamieson, B., Wise, K. D., and Buzsáki, G. (2003). Mechanisms of gamma oscillations in the hippocampus of the behaving rat. *Neuron*.

- Daou, A., Ross, M. T., Johnson, F., Hyson, R. L., and Bertram, R. (2013). Electrophysiological characterization and computational models of HVC neurons in the zebra finch. *Journal of Neurophysiology*.
- Dave, A. S. and Margoliash, D. (2000). Song replay during sleep and computational rules for sensorimotor vocal learning. *Science*.
- Daw, N. D., O’Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–879.
- Day, N. F. and Nick, T. A. (2013). Rhythmic cortical neurons increase their oscillations and sculpt basal ganglia signaling during motor learning. *Developmental neurobiology*.
- Day, N. F., Terleski, K. L., Nykamp, D. Q., and Nick, T. A. (2013). Directed functional connectivity matures with motor learning in a cortical pattern generator. *Journal of Neurophysiology*.
- Deregnaucourt, S., Mitra, P. P., Feher, O., Pytte, C., and Tchernichovski, O. (2005). How sleep affects the developmental learning of bird song. *Nature*.
- Diedrichsen, J. and Kornysheva, K. (2015). Motor skill learning between selection and execution. *Trends in Cognitive Science*, 19(4):227–233.
- Diesmann, M., Gewaltig, M. O., and Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature*.
- Dombeck, D. A., Graziano, M. S., and Tank, D. W. (2009). Functional clustering of neurons in motor cortex determined by cellular resolution imaging in awake behaving mice. *Journal of Neuroscience*.
- Eshel, N., Bukwich, M., Rao, V., Hemmelder, V., Tian, J., and Uchida, N. (2015). Arithmetic and local circuitry underlying dopamine prediction errors. *Nature*, 525(7568):243–246.
- Fee, M. S. and Goldberg, J. H. (2011a). A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience*.
- Fee, M. S. and Goldberg, J. H. (2011b). A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience*, 198:152–170.
- Fee, M. S. and Long, M. A. (2011). New methods for localizing and manipulating neuronal dynamics in behaving animals. *Current Opinion in Neurobiology*, 21(5):693–700.

- Funabiki, Y. and Funabiki, K. (2009). Factors limiting song acquisition in adult zebra finches. *Developmental Neurobiology*, 69(11):752–759.
- Ganguly, K. and Carmena, J. M. (2009). Emergence of a stable cortical map for neuroprosthetic control. *PLoS Biology*.
- Gardner, T., Cecchi, G., Magnasco, M., Laje, R., and Mindlin, G. (2001). Simple Motor Gestures for Birdsongs. *Physical Review Letters*.
- Georgopoulos, A. P., Merchant, H., Naselaris, T., and Amirkian, B. (2007). Mapping of the preferred direction in the motor cortex. *Proceedings of the National Academy of Sciences*.
- Ghosh, K. K., Burns, L. D., Cocker, E. D., Nimmerjahn, A., Ziv, Y., Gamal, A. E., and Schnitzer, M. J. (2011). Miniaturized integration of a fluorescence microscope. *Nature methods*.
- Gibb, L., Gentner, T. Q., and Abarbanel, H. D. I. (2009). Inhibition and recurrent excitation in a computational model of sparse bursting in song nucleus HVC. *Journal of Neurophysiology*.
- Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. J., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*.
- Giraud, A.-L. and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations.
- Glaze, C. M. and Troyer, T. W. (2006). Temporal structure in zebra finch song: implications for motor coding. *Journal of Neuroscience*, 26(3):991–1005.
- Goldberg, J. H., Adler, A., Bergman, H., and Fee, M. S. (2010). Singing-Related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: Comparison to the primate internal and external pallidal segments. *Journal of Neuroscience*, 30(20):7088–7098.
- Goold, C. P. and Nicoll, R. A. (2010). Single-cell optogenetic excitation drives homeostatic synaptic depression. *Neuron*, 68(3):512–528.
- Graber, M. H., Helmchen, F., and Hahnloser, R. H. R. (2013). Activity in a premotor cortical nucleus of zebra finches is locally organized and exhibits auditory selectivity in neurons but not in glia. *PLoS one*.
- Guitchounts, G., Markowitz, J. E., Liberti, W. A., and Gardner, T. J. (2013). A carbon-fiber electrode array for long-term neural recording. *Journal of Neural Engineering*.

- Guizar-Sicairos, M., Thurman, S. T., and Fienup, J. R. (2008). Efficient subpixel image registration algorithms. *Optics letters*.
- Hahnloser, R. H. R., Kozhevnikov, A. A., and Fee, M. S. (2002a). An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*.
- Hahnloser, R. H. R., Kozhevnikov, A. A., and Fee, M. S. (2002b). An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*, 419(6902):65–70.
- Hamaguchi, K., Tschida, K. A., Yoon, I., Donald, B. R., and Mooney, R. (2014). Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. *Elife*, 3.
- Hamaguchi, K., Yoon, I., Donald, B. R., and Mooney, R. (2012). Efficient coding of vocal transitions revealed by sparse synaptic activity in the singing bird. In *Society for Neuroscience*.
- Harvey, C. D., Coen, P., and Tank, D. W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature*.
- Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., and McCormick, D. A. (2005). Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron*.
- Hengen, K. B., Lambo, M. E., Van Hooser, S. D., Katz, D. B., and Turrigiano, G. G. (2013). Firing rate homeostasis in visual cortex of freely behaving rodents. *Neuron*, 80(2):335–342.
- Hengen, K. B., Torrado Pacheco, A., McGregor, J. N., Van Hooser, S. D., and Turrigiano, G. G. (2016). Neuronal firing rate homeostasis is inhibited by sleep and promoted by wake. *Cell*, 165(1):180–191.
- Hessler, N. A. and Doupe, A. J. (1999). Singing-related neural activity in a dorsal forebrain-basal ganglia circuit of adult zebra finches. *Journal of Neuroscience*, 19(23):10461–10481.
- Hira, R., Ohkubo, F., Ozawa, K., Isomura, Y., Kitamura, K., Kano, M., Kasai, H., and Matsuzaki, M. (2013). Spatiotemporal dynamics of functional clusters of neurons in the mouse motor cortex during a voluntary movement. *Journal of Neuroscience*.
- Hirase, H., Leinekugel, X., Czurkó, A., Csicsvari, J., and Buzsáki, G. (2001). Firing rates of hippocampal neurons are preserved during subsequent sleep episodes and modified by novel awake experience. *Proceedings of the National Academy of Science U. S. A.*, 98(16):9386–9390.

- Holtmaat, A. and Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nature Reviews Neuroscience*, 10(10):759–759.
- Holy, T. E. and Guo, Z. (2005). Ultrasonic songs of male mice. *PLoS Biology*, 3(12):e386.
- Hromádka, T., Deweese, M. R., and Zador, A. M. (2008). Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS Biology*, 6(1):e16.
- Huber, D., Gutnisky, D. A., Peron, S., O’Connor, D. H., Wiegert, J. S., Tian, L., Oertner, T. G., Looger, L. L., and Svoboda, K. (2012a). Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature*, 484(7395):473–478.
- Huber, D., Gutnisky, D. A., Peron, S., O’Connor, D. H., Wiegert, J. S., Tian, L., Oertner, T. G., Looger, L. L., and Svoboda, K. (2012b). Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature*.
- Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. *Bird vocalizations*.
- Jarvis, M. R. and Mitra, P. P. (2001). Sampling properties of the spectrum and coherency of sequences of action potentials. *Neural Computation*.
- Kaelbling, L. P. (1996). Introduction. *Machine Learning*, 22(1-3):7–9.
- Kao, M. H. (2006). Lesions of an avian basal ganglia circuit prevent Context-Dependent changes to song variability. *Journal of Neurophysiology*, 96(3):1441–1455.
- Karni, A., Meyer, G., Rey-Hipolito, C., Jezard, P., Adams, M. M., Turner, R., and Ungerleider, L. G. (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proceedings of the National Academy of Science U. S. A.*, 95(3):861–868.
- Katzner, S., Nauhaus, I., Benucci, A., Bonin, V., Ringach, D. L., and Carandini, M. (2009). Local origin of field potentials in visual cortex. *Neuron*.
- Kentros, C. G., Agnihotri, N. T., Streater, S., Hawkins, R. D., and Kandel, E. R. (2004). Increased attention to spatial context increases both place field stability and spatial memory. *Neuron*.
- Klausberger, T., Magill, P. J., Márton, L. F., Roberts, J. D. B., Cobden, P. M., Buzsáki, G., and Somogyi, P. (2003). Brain-state- and cell-type-specific firing of hippocampal interneurons in vivo. *Nature*.

- Kojima, S. and Doupe, A. J. (2011). Social performance reveals unexpected vocal competency in young songbirds. *Proceedings of the National Academy of Science U. S. A.*, 108(4):1687–1692.
- Koralek, A. C., Jin, X., Long, 2nd, J. D., Costa, R. M., and Carmena, J. M. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. *Nature*, 483(7389):331–335.
- Kornfeld, J., Benezra, S. E., Narayanan, R. T., Svara, F., Egger, R., Oberlaender, M., Denk, W., and Long, M. A. (2017). EM connectomics reveals axonal target variation in a sequence-generating network. *Elife*, 6.
- Kosche, G. and Long, M. A. (2014). Inhibitory coupling in a cortical premotor pattern-generating network. In *Society for Neuroscience*.
- Kozhevnikov, A. A. and Fee, M. S. (2007a). Singing-related activity of identified HVC neurons in the zebra finch. *Journal of Neurophysiology*, 97(6):4271–4283.
- Kozhevnikov, A. A. and Fee, M. S. (2007b). Singing-Related Activity of Identified HVC Neurons in the Zebra Finch. *Journal of Neurophysiology*.
- Krützfeldt, N. O. E. and Wild, J. M. (2004). Definition and connections of the entopallium in the zebra finch (*taeniopygia guttata*). *Journal of Comparative Neurology* ., 468(3):452–465.
- Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human brain mapping*.
- Lefort, S., Tómm, C., Floyd Sarria, J.-C., and Petersen, C. C. H. (2009). The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron*, 61(2):301–316.
- Leonardo, A. (2005). Degenerate coding in neural systems. *Journal of comparative physiology*.
- Leonardo, A. and Fee, M. S. (2005). Ensemble coding of vocal control in birdsong. *Journal of Neuroscience*, 25(3):652–661.
- Li, M. and Greenside, H. (2006). Stable propagation of a burst through a one-dimensional homogeneous excitatory chain model of songbird nucleus HVC. *Physical review. E, Statistical, nonlinear, and soft matter physics*.
- Lim, Y., Shinn-Cunningham, B., and Gardner, T. J. (2013). Stable Time-Frequency Contours for Sparse Signal Representation. In *Proceedings of the 21st European Signal Processing Conference (EUSIPCO)*.

- Lombardino, A. J. and Nottebohm, F. (2000a). Age at deafening affects the stability of learned song in adult male zebra finches. *Journal of Neuroscience*, 20(13):5054–5064.
- Lombardino, A. J. and Nottebohm, F. (2000b). Age at deafening affects the stability of learned song in adult male zebra finches. *Journal of Neuroscience*.
- Long, M. A. and Fee, M. S. (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature*, 456(7219):189–194.
- Long, M. A., Jin, D. Z., and Fee, M. S. (2010). Support for a synaptic chain model of neuronal sequence generation. *Nature*.
- Lynch, G. F., Okubo, T. S., Hanuschkin, A., Hahnloser, R. H. R., and Fee, M. S. (2016). Rhythmic Continuous-Time Coding in the Songbird Analog of Vocal Motor Cortex. *Neuron*.
- MacDonald, C. J., Lepage, K. Q., Eden, U. T., and Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontinuous events. *Neuron*.
- Margolis, D. J., Lütcke, H., Schulz, K., Haiss, F., Weber, B., Kügler, S., Hasan, M. T., and Helmchen, F. (2012). Reorganization of cortical population activity imaged throughout long-term sensory deprivation. *Nature Neuroscience*, 15(11):1539–1546.
- Markowitz, J. E., Ivie, E., Kligler, L., and Gardner, T. J. (2013a). Long-range order in canary song. *PLoS Computational Biology*, 9(5):e1003052.
- Markowitz, J. E., Ivie, E., Kligler, L., and Gardner, T. J. (2013b). Long-range order in canary song. *PLoS Computational Biology*.
- Markowitz, J. E., Liberti, W. A., Guitchounts, G., Velho, T., Lois, C., and Gardner, T. J. (2015). Mesoscopic patterns of neural activity support songbird cortical sequences. *PLoS Biology*.
- McCasland, J. S. (1987). Neuronal control of bird song production. *Journal of Neuroscience*, 7(1):23–39.
- Merchant, H., Pérez, O., Zarco, W., and Gámez, J. (2013). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *Journal of Neuroscience*.
- Mooney, R. (2000). Different subthreshold mechanisms underlie song selectivity in identified HVC neurons of the zebra finch. *Journal of Neuroscience*.

- Mooney, R., Hisey, E., and Tanaka, M. (2014). A novel cell type mediates motor to auditory interactions necessary to feedback-dependent song plasticity. In *Society for Neuroscience*.
- Mooney, R. and Prather, J. F. (2005a). The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. *Journal of Neuroscience*, 25(8):1952–1964.
- Mooney, R. and Prather, J. F. (2005b). The HVC Microcircuit: The Synaptic Basis for Interactions between Song Motor and Vocal Plasticity Pathways. *Journal of Neuroscience*.
- Mrsic-Flogel, T. D., Hofer, S. B., Ohki, K., Reid, R. C., Bonhoeffer, T., and Hübener, M. (2007). Homeostatic regulation of eye-specific responses in visual cortex during ocular dominance plasticity. *Neuron*, 54(6):961–972.
- Murthy, V. N. and Fetz, E. E. (1992). Coherent 25- to 35-Hz oscillations in the sensorimotor cortex of awake behaving monkeys. *Proceedings of the National Academy of Sciences of the United States of America*.
- Naselaris, T., Merchant, H., Amirikian, B., and Georgopoulos, A. P. (2006). Large-scale organization of preferred directions in the motor cortex. II. Analysis of local distributions. *Journal of Neurophysiology*.
- Nir, Y., Fisch, L., Mukamel, R., Gelbard-Sagiv, H., Arieli, A., Fried, I., and Malach, R. (2007). Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Current biology : CB*.
- Niv, Y. (2009). Reinforcement learning in the brain. *J. Math. Psychol.*, 53(3):139–154.
- Nixdorf-Bergweiler, B. E., Wallhäusser-Franke, E., and DeVoogd, T. J. (1995). Regressive development in neuronal structure during song learning in birds. *Journal of Neurobiology*, 27(2):204–215.
- Nordeen, K. W. and Nordeen, E. J. (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral and neural biology*.
- Nottebohm, F. and Arnold, A. P. (1976). Sexual dimorphism in vocal control areas of the songbird brain. *Science*, 194(4261):211–213.
- Nottebohm, F., Kelley, D. B., and Paton, J. A. (1982). Connections of vocal control nuclei in the canary telencephalon. *Journal of Comparative Neurology*.
- Nottebohm, F., Stokes, T. M., and Leonard, C. M. (1976). Central control of song in the canary, *serinus canarius*. *Journal of Comparative Neurology* ., 165(4):457–486.

- Ohgushi, E., Mori, C., and Wada, K. (2015). Diurnal oscillation of vocal development associated with clustered singing by juvenile songbirds. *The Journal of experimental biology*.
- Olveczky, B. P., Andalman, A. S., and Fee, M. S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biology*, 3(5):e153.
- Otchy, T. M., Wolff, S. B. E., Rhee, J. Y., Pehlevan, C., Kawai, R., Kempf, A., Gobes, S. M. H., and Olveczky, B. P. (2015). Acute off-target effects of neural circuit manipulations. *Nature*.
- Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*.
- Pesaran, B., Nelson, M. J., and Andersen, R. A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature*.
- Peters, A. J., Chen, S. X., and Komiyama, T. (2014). Emergence of reproducible spatiotemporal activity during motor learning. *Nature*.
- Picardo, M. A., Merel, J., Katlowitz, K. A., Vallentin, D., Okobi, D. E., Benezra, S. E., Clary, R. C., Pnevmatikakis, E. A., Paninski, L., and Long, M. A. (2016a). Population-Level representation of a temporal sequence underlying song production in the zebra finch. *Neuron*, 90(4):866–876.
- Picardo, M. A., Merel, J., Katlowitz, K. A., Vallentin, D., Okobi, D. E., Benezra, S. E., Clary, R. C., Pnevmatikakis, E. A., Paninski, L., and Long, M. A. (2016b). Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch. *Neuron*.
- Poole, B., Markowitz, J. E., and Gardner, T. J. (2012). The song must go on: resilience of the songbird vocal motor pathway. *PloS one*.
- Pytte, C., Yu, Y.-L., Wildstein, S., George, S., and Kirn, J. R. (2011). Adult neuron addition to the zebra finch song motor pathway correlates with the rate and extent of recovery from botox-induced paralysis of the vocal muscles. *Journal of Neuroscience*, 31(47):16958–16968.
- Quiroga, R. Q., Nadasdy, Z., and Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Computation*.
- Rokni, U., Richardson, A. G., Bizzi, E., and Seung, H. S. (2007). Motor learning with unstable neural representations. *Neuron*.
- Rubino, D., Robbins, K. A., and Hatsopoulos, N. G. (2006). Propagating waves mediate information transfer in the motor cortex. *Nature neuroscience*.

- Sahani, M. (1999). Latent variable models for neural data analysis.
- Schmidt, M. (2003). Pattern of Interhemispheric Synchronization in HVC During Singing Correlates With Key Transitions in the Song Pattern. *Journal of Neurophysiology*.
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.
- Scott, B. B., Gardner, T., Ji, N., Fee, M. S., and Lois, C. (2012). Wandering neuronal migration in the postnatal vertebrate forebrain. *Journal of Neuroscience*.
- Scott, B. B. and Lois, C. (2005). Generation of tissue-specific transgenic birds with lentiviral vectors. *Proceedings of the National Academy of Sciences of the United States of America*.
- Shmuelof, L. and Krakauer, J. W. (2014). Recent insights into perceptual and motor skill learning. *Frontiers in Human Neuroscience*, 8:683.
- Shmuelof, L., Yang, J., Caffo, B., Mazzoni, P., and Krakauer, J. W. (2014). The neural correlates of learned motor acuity. *Journal of Neurophysiology*, 112(4):971–980.
- Sossinka, R. and Böhner, J. (1980). Song types in the zebra Finch *Poephila guttata castanotis*. *Zeitschrift für Tierpsychologie*, 53(2):123–132.
- Stauffer, T. R., Elliott, K. C., Ross, M. T., Basista, M. J., Hyson, R. L., and Johnson, F. (2012). Axial organization of a brain region that sequences a learned pattern of behavior. *Journal of Neuroscience*.
- Sturdy, C. B., Wild, J. M., and Mooney, R. (2003). Respiratory and telencephalic modulation of vocal motor neurons in the zebra finch. *Journal of Neuroscience*, 23(3):1072–1086.
- Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press.
- Tchernichovski, O., Nottebohm, F., Ho, C., Pesaran, B., and Mitra, P. (2000). A procedure for an automated measurement of song similarity. *Animal behaviour*.
- Thompson, L. T. and Best, P. J. (1990). Long-term stability of the place-field activity of single units recorded from the dorsal hippocampus of freely behaving rats. *Brain research*.
- Tolias, A. S., Ecker, A. S., Siapas, A. G., Hoenselaar, A., Keliris, G. A., and Logothetis, N. K. (2007). Recording chronically from the same neurons in awake, behaving primates. *Journal of Neurophysiology*.

- Tschida, K. A. and Mooney, R. (2012). Deafening drives cell-type-specific changes to dendritic spines in a sensorimotor nucleus important to learned vocalizations. *Neuron*.
- Tumer, E. C. and Brainard, M. S. (2007). Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature*, 450(7173):1240–1244.
- Turrigiano, G. G. (2008). The self-tuning neuron: synaptic scaling of excitatory synapses. *Cell*, 135(3):422–435.
- Vallentin, D. and Long, M. A. (2015). Motor origin of precise synaptic inputs onto forebrain neurons driving a skilled behavior. *Journal of Neuroscience*.
- Vates, G. E. and Nottebohm, F. (1995). Feedback circuitry within a song-learning pathway. *Proceedings of the National Academy of Science U. S. A.*, 92(11):5139–5143.
- Vicario, D. S. and Nottebohm, F. (1988). Organization of the zebra finch song control system: I. representation of syringeal muscles in the hypoglossal nucleus. *Journal of Comparative Neurology* ., 271(3):346–354.
- Vu, E. T., Schmidt, M. F., and Mazurek, M. E. (1998). Interhemispheric coordination of premotor neural activity during singing in adult zebra finches. *Journal of Neuroscience*, 18(21):9088–9098.
- Wang, C. Z. H., Herbst, J. A., Keller, G. B., and Hahnloser, R. H. R. (2008). Rapid interhemispheric switching during vocal production in a songbird. *PLoS Biology*, 6(10):e250.
- Weinstein, G. D. and van Scott, E. J. (1965). Autoradiographic Analysis of Turnover Times of Normal and Psoriatic Epidermis1. *Journal of Investigative Dermatology*.
- Williams, H. (2004). Birdsong and singing behavior. *Annual New York Academy of Science*, 1016(1):1–30.
- Williams, H., Crane, L. A., Hale, T. K., Esposito, M. A., and Nottebohm, F. (1992). Right-side dominance for song control in the zebra finch. *Journal of Neurobiology*.
- Wilson, A. A., Kwok, L. W., Hovav, A.-H., Ohle, S. J., Little, F. F., Fine, A., and Kotton, D. N. (2008). Sustained expression of alpha1-antitrypsin after transplantation of manipulated hematopoietic stem cells. *American journal of respiratory cell and molecular biology*.
- Wiltschko, A. B., Johnson, M. J., Iurilli, G., Peterson, R. E., Katon, J. M., Pashkovski, S. L., Abaira, V. E., Adams, R. P., and Datta, S. R. (2015). Mapping Sub-Second structure in mouse behavior. *Neuron*, 88(6):1121–1135.

- Woolley, S. M. N., Fremouw, T. E., Hsu, A., and Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nature Neuroscience*, 8(10):1371–1379.
- Xing, D., Yeh, C.-I., and Shapley, R. M. (2009). Spatial spread of the local field potential and its laminar variation in visual cortex. *Journal of Neuroscience*.
- Yassin, L., Benedetti, B. L., Jouhanneau, J.-S., Wen, J. A., Poulet, J. F. A., and Barth, A. L. (2010). An embedded subnetwork of highly active neurons in the neocortex. *Neuron*, 68(6):1043–1050.
- Ziv, Y., Burns, L. D., Cocker, E. D., Hamel, E. O., Ghosh, K. K., Kitch, L. J., El Gamal, A., and Schnitzer, M. J. (2013). Long-term dynamics of CA1 hippocampal place codes. *Nature neuroscience*.

Curriculum Vitae

