Diss. ETH NR. 21248

Correlated Population Dynamics in the Songbird Auditory Forebrain: Insights into the Origins of Song Selectivity

A dissertation submitted to

ETH Zurich

for the degree of

Doctor of Sciences

presented by

JANIE MICHELLE ONDRACEK
MSc. Neural Systems and Computation, University of Zurich

Born on 12.04.1982
citizen of
Schaffhausen, SH

accepted on the recommendation of

Richard H. R. Hahnloser
Richard Mooney

2013

Abstract

The brain distributes sensory information over large neuronal networks. To understand how neural responses are shaped by these networks, correlated dynamics across neuron pairs have been researched extensively for different animal models and sensory modalities.

In order to examine the effect that neural correlations may have on the processing of complex acoustic stimuli, we recorded from ensembles of neurons in the songbird auditory cortex in response to both natural and synthetic auditory stimuli. We calculated spike count correlations and mean-subtracted, spike train cross-covariances from the spiking responses and examined how correlation dynamics might be related to the emergence of stimulus selectivity in the songbird auditory forebrain.

The songbird is a useful animal model for the study of auditory processing, because the songbirds learning of its own song is tightly linked to the auditory processing and perception of birdsong vocalizations. Young birds raised in acoustic isolation produce abnormal isolate songs [Marler, 1970], and auditory feedback is essential for song learning [Nottebohm, 1970; Konishi, 1985].

We began by examining correlations between single pairs of neurons and small ensembles. We found that nearby neurons recorded on the same electrode rarely shared significant spike count correlations, although a subset of nearby neuron pairs were strongly anti-correlated in response to playbacks of the bird's own song (BOS). In an ensemble of neurons that responded to sound onsets, we observed diverse patterns of spike time correlations, despite the fact that the neurons responded to similar spectro-temporal features of the auditory stimuli.

We then pooled mean correlations over birds to examine whether different correlation dynamics emerge in response to playbacks of distinct classes of auditory stimuli. For a population of 106 auditory forebrain neurons, we examined correlations in response to auditory stimuli including BOS, reversed BOS (REV), conspecific songs (CON), and synthetic white noise (WN stimuli). We then compared these auditory-evoked correlations to spontaneous firing during silence.

We analyzed correlations in spike counts for both large (1500 ms) and small (200 ms) time windows. In both cases, we observed almost no significant correlation effect in the

population responses. While some neurons pairs shared significant trial-to-trial variability in response to certain stimuli, this effect was largely averaged out over the population.

In contrast, when we examined correlations in the timing of spikes in smaller 10 ms windows, we observed that neural spike-time correlations exist spontaneously during silence and are strongly reduced by sensory stimulation. This effect was also true for a smaller population of 62 primary auditory neurons in the primary auditory cortex.

During stimulus-playbacks, spike-time correlation were positive mostly due to shared intrinsic correlations. In response to playbacks of the BOS, spike-time correlations at short time lags were significantly reduced compared to the other stimulus classes, which agrees with sparse coding strategies that shape synaptic networks to reduce neural responses in proportion to the prominence of a stimulus, which in our case is highest for the BOS.

In order to determine whether the correlations dynamics observed in our population of auditory neurons could reflect a general pattern of decorrelation that accompanies stimulus-selective neural responses, we simulated 11 auditory networks of 1000 neurons each using the efficient coding algorithms implemented in Blaettler and Hahnloser [2011]. We found that as network selectivity increased to favor the BOS, correlations in spike timing in response to the BOS stimulus were decorrelated, matching the effect we observed in vivo. Thus, the avian auditory cortex may shape its networks to selectively suppress response redundancy for behaviorally relevant stimuli, including self-generated ones such as the BOS.

Zusammenfassung

Im Gehirn werden sensorische Reize über viele neuronale Netzwerke hinweg verteilt. Eine Großzahl an Studien in diversen Tiermodellen und Sinnesmodalitäten hat die korrelierten Dynamiken von Neuronenpaaren untersucht um zu verstehen, wie neuronale Aktivität von diesen Netzwerken moduliert und geformt wird.

Um zu untersuchen welchen Einfluss neuronale Korrelationen auf die Verarbeitung von komplexen akustischen Reizen haben, haben wir elektrophysiologische Ableitungen im auditorischen Kortex des Singvogels durchgeführt, während zeitgleich natürliche und künstliche Stimuli abgespielt worden sind. Wir haben die Korrelationen der Anzahl der Aktionspotentiale sowie die Kreuzkovarianz (nach Mittelwertsubtraktion) der Aktionspotentiale berechnet um zu untersuchen welchen Einfluss korrelierte Aktivität auf die Entstehung von Stimulus-Selektivität im auditorischen Vorderhirn des Singvogels haben könnte.

Der Singvogel ist ein nützliches Tiermodell für die Untersuchung der auditorische Verarbeitung im Gehirn, da das Lernen des eigenen Gesangs eng verknüpft ist mit der auditorischen Verarbeitung und Wahrnehmung der vom Vogel produzierten Vokalisierung. Junge Vögel welche in akustischer Isolation aufwachsen, produzieren abnormalen “isolierten” Gesang [Marler, 1970]; auditorisches Feedback ist eine notwendige Vorraussetzung für das Erlernen des Gesangs[Nottebohm, 1970; Konishi, 1985].

Zunächst haben wir die Korrelationen zwischen einzelnen Neuronenpaaren und kleinen Neuronenpopulationen untersucht. Dabei haben wir herausgefunden, dass benachbarte Neuronen welche auf der gleichen Elektrode aufgenommen worden sind, nur sehr vereinzelt signifikante Korrelationen in der Anzahl der Aktionspotentiale aufweisen. Eine Teilmenge dieser Neuronen zeigt jedoch eine starke Antikorrelation der Aktivität sobald der eigene Gesang des Vogels abgespielt wird (bird’s own song - BOS). In einer anderen Subpopulation welche fast ausschliesslich nur zu Begin des auditorischen Reizes feuert, haben wir sehr unterschiedliche Korrelationenmuster in der Anzahl der Aktionspotentiale vorgefunden und das obwohl alle Neuronen auf ähnliche zeit-frequenz Charakteristiken der auditorischen Reize reagiert haben.

Anschliessend haben wir die gemittelten Korrelationen (gemittelt über mehrere Tiere) betrachtet, um zu untersuchen, ob unterschiedliche Korrelationsdynamiken aufgrund

der unterschiedlichen Reizkategorien entstehen. In einer Population von 106 auditorischen Neuronen des Vorderhirns, haben wir die Korrelationen der Aktivitäten als Antwort auf die unterschiedlichen Reize BOS, umgehrter eigener Gesang (reverse BOS; REV), konspezifischer Gesang (CON) und künstliches Weisses Rauschen (WN) Untersucht. Anschliessend haben wir die Korrelationen welche durch auditorische Reize hervorgerufen werden, mit denen in Phasen von spontaner Aktivitäten (kein Abspielen von Reizen, völlige Stille) verglichen.

Wir haben die Korrelationen in der Anzahl der Aktionspotentiale sowohl für grosse Zeitfenster (1500 ms) als auch für kleine Zeitfenster (200 ms) berechnet. In beiden Fällen haben wir fast ausschlielich nicht signikante Korrelationen auf Populationsebene vorgefunden. Während die Antworten einiger Neuronenpaare auf bestimmte Reize zu signifikanten Korrelationen geführt haben, war dieser Effekt nicht auf Populationsebene ersichtlich, da sich Korrelationen im Mittel aufgehoben haben.

Im Gegensatz dazu haben wir positive Korrelationen zwischen den Zeitpunkten der Aktionspotentiale (in Zeitfenstern von 10 ms) während Phasen von spontaner Aktivität (kein Abspielen von Reizen, völlige Stille) gefunden, welche jedoch stark vermindert werden sobald auditorische Reize abgespielt werden. Den selben Effekt haben wir auch in einer kleineren Subpopulation von 62 auditorischen Neuronen im primären auditorischen Kortex vorgefunden.

Als Antwort auf das Abspielen des eigenen Gesanges (BOS) haben die oben erwähnt Korrelationen sogar mehr abgenommen. Dieses Ergebnis steht im Einklang mit sogenannten “sparse coding” Theorien welche davon ausgehen, dass synaptische Netzwerke moduliert werden um neuronale Antworten proportional zu der Signifikanz des jeweiligen Stimulus zu reduzieren. Diese ist in unserem Fall für BOS am höchsten.

Wir haben 11 auditorische Netzwerke (bestehend aus 1000 Neuronen) mit einem “efficient coding” Algorithmus trainiert (Blaettler und Hahnloser [2011]) um zu untersuchen ob sich unsere Korrelationsergebnisse auch in simulierten neuronalen Netzen ergeben. Diese Simulationen haben gezeigt, dass neuronale Aktivitäten als Antwort auf BOS dekorreliert werden, sobald sich die Selektivität des Netzwerks in Richtung BOS bewegt. Dieses Ergebnis steht im Einklang mit unseren experimentellen Beobachtungen. Daher ist es vorstellbar, dass der auditorische Kortex des Singvogels seine neuronalen Netze gezielt anpasst um selektiv auf verhaltensrelevant Stimuli zu antworten (einschließlich BOS).

Acknowledgements

I would like to thank Richard Hahnloser for financially supporting me over the years, for giving me the freedom to learn and practice new experimental techniques, and for giving me the opportunity to to attend many different conferences over the years that helped to shape my interest in population coding.

I would like to thank Richard Mooney for agreeing to be on my thesis committee, for reading and critiquing my thesis, and for traveling to Switzerland for my examination. Furthermore, I greatly appreciated the expert tutelage of Richard and his post-doc Kosuke Hamaguchi, as they taught me their intracellular recording methods during a visit to Duke University in Dec. 2011.

I would like to thank both Andreas Kotowicz and Joshua Herbst for their help and assistance to the lab during my PhD. Andreas was heroic in his effort to teach me how to program in Matlab, to provide the lab with general software analysis routines, and in recovering our code repositories and wikis that were lost during the INI IT crash of 2010. Josh's commitment to updating the Labview data acquisition software - a process that took more than a year and involved many revisions, patient explanations, and tutorials - was an enormous gift to the lab that made my job as an experimentalist much, much easier.

I probably could not have made it through this work without the camaraderie of the songbird group - thanks to you all for the fun and great times, whether it was hiking up a mountain in Bregel, skiing in St. Moritz, watching a football game over pizza, or carving pumpkins and dressing up like a zebra finch for Halloween. These memories will be forever connected to my life and studies in Zurich.

Contents

List of Figures

Abbreviations

Anatomy

Av	**A**valanche nucleus
CM	**C**audal **M**esopallium
GP	**G**lobus **P**allidus
HVC	used as a proper name
ICo	**I**ntercollicular **Co**mplex
LaM	**La**mina **m**esopallialis
LFS	**L**amina **F**rontalis **S**uperior
LMAN	**L**ateral **M**agnocellular nucleus of the **A**nterior **N**idopallium
MMAN	**M**edial **M**agnocellular nucleus of **A**nterior **N**idopallium (MMAN)
NCM	**N**idopallium **C**audal **M**edial
NIf	**N**ucleus **I**nter**f**ace of the nidopallium
Ov	**Ov**oidalis
RA	**R**obust nucleus of the **A**rcopalliu1m
Uva	**Uva**eform nucleus
X	Area **X**

Auditory Stimuli

BOS	**B**ird's **O**wn **S**ong
CON	**CON**specific Song
ML-Noise	**M**odulation **L**imited **Noise**
REV	**Rev**ersed Bird's Own Song
RO-BOS	**R**eversed **O**rder - **B**ird's **O**wn **S**ong
WN	**W**hite **N**oise
WN-S | **WN-E**	**W**hite **N**oise - **S**tart of file | **E**nd of file

Analysis

CCV	Spike Train **C**ross **Co**variance
CV	**C**oefficient of **V**ariation
PSTH	**P**eri-**S**timulus **T**ime **H**istogram
R_{SC}	**S**pike **C**ount Correlation
STRF	**S**pectro - **T**emporal **R**eceptive **F**ield
$STRF_{SI}$	**STRF S**imilarity **I**ndex

Other

dph	**d**ays **p**ost **h**atch
GABA	**G**amma - **A**mino**B**utyric **A**cid
PSP	**P**ost **S**ynaptic **P**otential
EPSP	**E**xcitatory **P**ost **S**ynaptic **P**otential
IPSP	**I**nhibitory **P**ost **S**ynaptic **P**otential
SPL	**S**ound **P**ressure **L**evel
S/N	**S**ignal-to-**N**oise

Part I

Introduction

Chapter 1

Project Motivation

How does the brain process the complex signals that we encounter in the natural environment? We have some understanding of how physical stimuli are transduced by our sense organs into patterns of neural activity, and how such responses could be processed, transformed, and propagated through the brain. However, we have only a vague understanding of the neural circuits and the computations necessary to extract an abstract sensory perception from a series of action potentials.

Conceptually, sensory perception has often been discussed in the context of a feed-forward, hierarchical network [Riesenhuber and Poggio, 1999; Ungerleider and Haxby, 1994], where neurons in lower sensory brain regions encode information related to fundamental features of the stimulus. Because the neural activity is usually most strongly modulated by a specific feature of the stimulus, these neurons are often simply termed "feature-detectors" [Barlow, 1953]. As neural activity propagates hierarchically through subsequent brain areas, simple stimulus representations are transformed into high-level representations, and the concept of a feature-detector gives way to stimulus-selectivity, e.g., the concept that neurons in higher brain areas may be selective to a composite stimulus, such as a particular birdsong [Margoliash, 1983] or a particular face [Quiroga et al., 2005].

Although individual neurons may exhibit stimulus selectivity, neural activity is known to be variable for identical repetitions of a stimulus [de Ruyter van Steveninck et al., 1997; Diba et al., 2004; Montemurro et al., 2007; Padmanabhan and Urban, 2010; Tovée et al., 1993]. Because of this variability, the brain must be capable of averaging the individual neural responses in order to compute a population code that provides an accurate estimate of the stimulus [Averbeck et al., 2006].

Further adding to the complexity of neural computations is the fact that variability is often correlated across neurons. Correlations in spike count variability - often called "intrinsic noise" - have been reported for neurons in the visual cortex of macaque [Kohn and Smith, 2005; Womelsdorf et al., 2012], the extrastriate cortex of rhesus monkeys [Huang and Lisberger, 2009], the auditory cortex of mice [Rothschild et al., 2010], and the somatosensory cortex of rats [Greenberg et al., 2008]. Although seemingly ubiquitous, the computational role of neural correlation is not well understood. However, one fact remains clear: because the activity of pairs of neurons is correlated, the activity of a population of neurons cannot be fully characterized by the measurement of its individual responses [Ohiorhenuan et al., 2010].

In order to examine the effect that correlations in spike rates and spike timing may have on the processing of complex acoustic stimuli - and how they might be related to the emergence of stimulus selectivity - we calculated spike count correlations and mean-subtracted, spike train cross-covariances from the spiking responses of ensembles of neurons in the auditory cortex of a songbird, the zebra finch (*Taeniopygia guttata*).

The songbird is a useful model animal for the study of auditory processing, because the songbird's learning of its own song is tightly linked to the auditory processing and perception of birdsong vocalizations. Young birds raised in acoustic isolation produce abnormal isolate songs [Marler, 1970a], and auditory feedback is essential for song learning [Konishi, 1985; Nottebohm, 1970].

The songbird auditory cortex is analogous to the auditory cortex of mammals [Dugas-Ford et al., 2012; Karten, 1991; Wang et al., 2010b], and consists of the areas Field L, the caudal mesopallium (CM), and the nidopallium caudal medial (NCM). Neurons in this brain region are heterogeneous [Fortune and Margoliash, 1992], highly interconnected [Vates et al., 1996], and have complex responses to synthetic and natural auditory stimuli [Amin et al., 2004; Grace et al., 2003].

Although receptive field estimation has done much to provide clues as to which features of auditory stimuli these neurons are tuned [Nagel and Doupe, 2006; Theunissen et al., 2004b; Woolley et al., 2009], the underlying coding strategies used by networks of neurons involved in auditory processing and stimulus selectivity are poorly understood. Furthermore, the role of correlated spiking activity and how it may relate to the representation of complex acoustic signals and stimulus selectivity has not been explored in the songbird auditory cortex.

In the following thesis, we will first introduce key concepts relevant to the understanding of neural computation, population coding, and neural correlations. Then we will introduce auditory processing in the songbird, discussing the strategies that are used in

song learning as well as the relevant auditory brain anatomy. A methods section documents the experiments and analysis performed, and the results sections summarize our experimental findings. The final synthesis section discusses our experimental findings in the larger context of neural correlation and population coding.

Chapter 2

Intrinsic Noise, Correlation, and Information

1 Introduction

The quest to crack the neural code has driven neuroscience research for almost a century. Although the neural code is extremely complex, the answer seems to involve the way that external information is relayed to the brain, and how the brain communicates with itself and with the sense organs of the body. But how does the brain, a mass of fat and tissue, actually communicate such information? The following section summarizes some elementary principles that are essential for the the understanding of neural computation.

1.1 Neuronal Specializations Support Information Transfer

The functional element of the brain is the neuron. Neurons are electrically excitable cells that comprise the networks of our brains, transforming and relaying information through a combination of electrical and chemical signals. Some estimate that the human brain has as many as 100 billion neurons [Williams and Herrup, 1988], each one capable of receiving input signals from and sending output signals to other neurons.

Neurons have unique morphological specializations that allow for the propagation of signals from one area in the brain to another. The dendrites are the branching structures that make synaptic contact with and receive input from other neurons. Some neurons, such as the cortical pyramidal neurons (Fig. 2.1), can receive thousands of synaptic inputs, whereas the purkinje neurons of the cerebellum (Fig. 2.1) can receive more than

100,000 [Fox and Barnard, 1957]. If the dendrites are the signal receiving structure of neurons, the axon, in contrast, is the structure that transmits the output of the neuron. The axon is long can travel distances far across the brain; for example, in the mouse brain, it has been estimated that cortical neurons typically send out about 40 mm of axon [Dayan and Abbott, 2001].

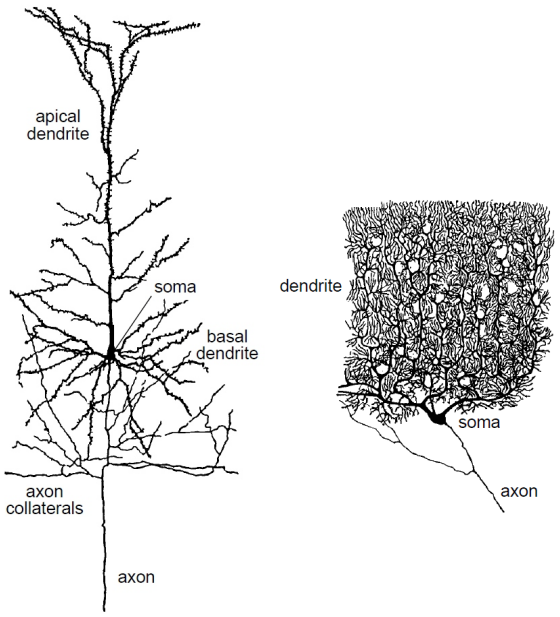

FIGURE 2.1: **Morphological specializations of neurons.** Left, illustration of a cortical pyramidal neuron, characterized by its morphological specializations including spine-laden apical and basal dendritic arbors and multiple axonal collaterals. Right, illustration of a cerebellar purkinje neuron, characterized by its morphological specializations including the dense, tree-like dendritic arbor and and simple axonal structure. Figure adapted from [Dayan and Abbott, 2001].

In addition to their morphological specializations, neurons also have diverse physiological specializations, and perhaps the most important of these specializations are the ion channels that allow sodium, potassium, calcium, and chloride ions to move in and out of the cell. The flux of these ions defines the electrical potential between the interior of the neuron and the extracellular medium in which it sits. The difference in electrical potential between the outside and the inside of the neuron, referred to as the membrane potential, is the electrical signal of relevance in the nervous system.

The neuron's membrane potential depends on the electrochemical equilibrium of the currents flowing in and out of the neuron. Although the membrane potential is maintained at equilibrium through the activity of a number of active and passive processes, it is sensitive to perturbations. When the membrane potential is sufficiently depolarized, e.g., through the opening of sodium channels that allows sodium current to pass into

the neuron, a positive feedback loop is generated along the axon of the neuron that may result in the initiation of an action potential.

An action potential, also often referred to as "spike" due to its characteristic shape, is a stereotypical fluctuation in the electrical potential across the cell membrane. Action potentials generated by sodium currents are about 100 mV in magnitude and last about 1 ms [Hodgkin and Huxley, 1939, 1952].

Because the axon is lined with sodium channels, action potentials can be actively regenerated along the axon, making them the only form of membrane potential fluctuation that can propagate over large distances. The action potential travels down the length of the axon and terminates at synapses, where voltage transients lead to the opening of ion channels and produce an influx of calcium ions that causes the release of chemical neurotransmitters (e.g., glutamate, GABA, dopamine) into the synaptic cleft.

Neurotransmitters bind to receptors at the post-synaptic side of the synapse and cause ion channels on the post-synaptic neuron to open. The opening and closing of ion channels cause post-synaptic potentials (PSPs) that perturb the resting membrane potential of the post-synaptic neuron. Depending on the combination of ion channels that open, these potentials might be excitatory post-synaptic potentials (EPSPs) that depolarize the membrane potential (often resulting in an action potential), or inhibitory post-synaptic potentials (IPSPs) that hyperpolarize the post-synaptic neuron (and suppresses the initiation of an action potential).

Since information about the exact shape and magnitude of the action potential cannot be transferred across neurons, only the number and timing of action potentials can encode the signals and computations transmitted by neurons across brain areas. Understanding how information is computed and transformed through these networks of neurons is key to understanding this complicated and remarkable neural code.

2 Neural Operating Regimes and Biological Noise

In the previous section we briefly discussed how neurons produce action potentials and argued that the number and timing of action potentials is the source of information transfer across the brain. However, biological systems are noisy [Fatt and Katz, 1950], and the trains of action potential produced by neural circuits are no exception. How does this biological noise affect the the ability of the neurons to extract information from noisy patterns of action potentials?

2.1 Biological Variability: Intrinsic Noise

Neural activity patterns are known to show variability across identical repetitions of a stimulus [de Ruyter van Steveninck et al., 1997; Dean, 1981; Montemurro et al., 2007; Padmanabhan and Urban, 2010; Schiller et al., 1976; Tolhurst et al., 1983; Tovée et al., 1993], and neurons in the songbird auditory forebrain are no exception. Figure 2.2 displays the response of two songbird auditory forebrain neurons to repeated playbacks of the bird's own song (BOS) motif through a loudspeaker. Both neurons respond to certain spectro-temporal features of the auditory stimulus, and yet their responses are noisy: certain stimulus features might elicit a neural response in one repetition, but not in another (Fig. 2.2; yellow arrows).

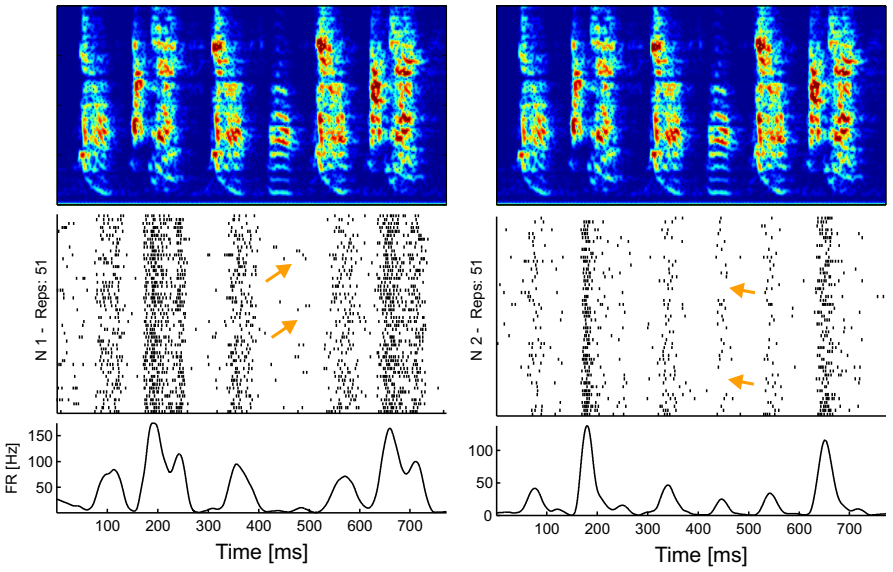

FIGURE 2.2: **Variable responses to repeated presentations of the bird's own song** [g18r2 01] Panels show the responses of two auditory forebrain neurons in response to auditory playback of the bird's own song (spectrograms, top). The response of the neuron is depicted as a spike raster plot, where each line indicates a spike. The neuron on the left inconsistently emits spikes in response to a stimulus feature (yellow arrows), whereas the neuron on the right occasionally does not emit spikes (yellow arrows). The average firing rate curve is depicted beneath the raster plots and smoothed using the matlab "lowess" function using a span of 45 ms. Spectrograms are depicted in the range of 0 to 8 kHz.

Noisy responses are thought to result from the variability associated with truly random intrinsic events, such as the stochastic opening and closing of ion channels and the inherent variability of synaptic transmission [Diba et al., 2004; London et al., 2010], coupled to non-linear thresholding involved in spike-generation. This variability is called *intrinsic noise*, since it is the result of variability intrinsic to the neuron that cannot be eliminated.

Among the electrophysiological studies that have examined sources of intrinsic noise, Mainen and Sejnowski [1995] obtained somatic whole-cell recordings from rat cortical slices and stimulated neurons by injecting constant and fluctuating current. Whereas constant current injection produced temporally noisy spike trains with reliable spike counts (Fig. 2.3A), fluctuating current produced spike trains with precise and stable timing and spike counts (Fig. 2.3B). These results suggest that low intrinsic noise is present in the mechanisms of spike generation [Mainen and Sejnowski, 1995] and that cortical neurons are capable of producing spikes with high temporal precision of less that 1 ms.

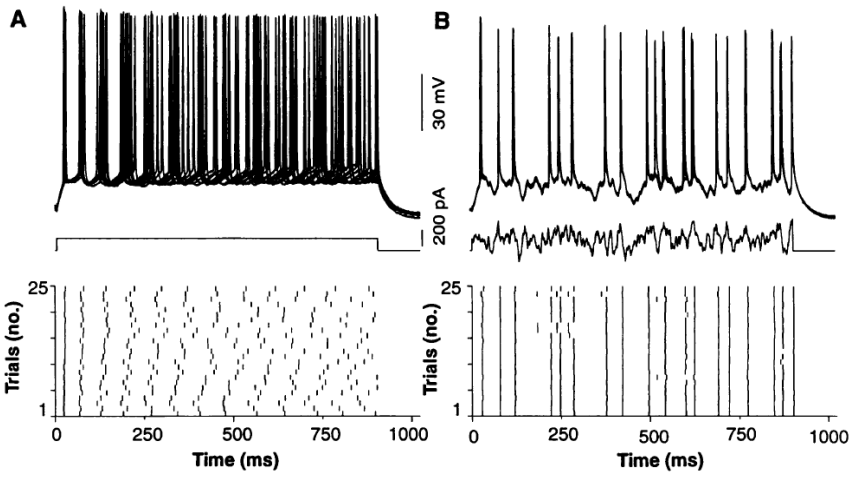

FIGURE 2.3: **Reliability of spiking patterns evoked by constant and fluctuating current. A)** A direct current pulse (150 pA, 900 ms; middle plot) evoked trains of action potentials. Intracellular spiking responses are shown superimposed upon each other (top) and as a rasterplot (bottom). Whereas the first spike was tightly locked to the onset of the current pulse, the timing of the last spike was highly variable. **B)** The same cell in **A** was stimulated with fluctuating current. Unlike the direct current pulse stimulation, fluctuating current evokes responses that are both temporally precise and highly reliable. Figure from [Mainen and Sejnowski, 1995].

2.2 Integrators or Coincidence Detectors?

How does the presence of intrinsic noise affect the real-time process of extracting information from patterns of action potentials? Discussions of neuronal noise typically focus on how neurons can overcome or compensate for this noise and still process and transmit information [Faisal et al., 2008].

Although the fine microstructure detail of spike trains is often noisy and unreproducible, the mean firing rate of a neuron (typically averaged over a fraction of a second or more)

often varies predictably with stimulus parameters and is reproducible. Such observations have given rise to the notion of a *rate code*: the theory that neurons encode information from the mean number of spikes produced in response to a stimulus.

The rate coding idea has existed since Adrian and Zotterman [1926] first observed that peripheral touch receptors respond to increasing pressure stimulation with an increase in the amount of neuronal firing. Firing rate metrics were used extensively in the following decades to describe the response properties of neurons in somatosensory cortex [Mountcastle, 1957] and visual cortex [Hubel and Wiesel, 1959], and are still commonly used today (e.g., peri-stimulus time histogram (PSTHs) and normalized comparisons of firing rates (z-scores)). The firing rate is widely accepted as the primary variable relating neuronal responses to stimuli, and the resulting assumption is that the fine temporal detail of the spike patterns are the result random intrinsic noise and carry no useful information.

Such views have resulted in the theory that neurons act as *integrate-and-fire* devices. This idea has dominated cortical physiology and most neural-network models for years [König et al., 1996]. Acting as integrate-and-fire devices, neurons sum synaptic potentials over integration time periods that are constrained by the membrane time constant of the neuron.

One criticism of this idea, many have argued, is that the integration period required to distinguish subtle stimulus-modulated effects is on the order of several hundreds of milliseconds [Masuda and Aihara, 2007], and in many situations, this integration period is too long to explain rapid information processing in the brain [Gautrais and Thorpe, 1998], which can be less than 150 ms in human visual processing [Thorpe et al., 1996].

An alternative theory, in which the the precise timing of spikes are the important signal, can be traced to MacKay and McCulloch [1952]. In perhaps one of the first reports to examine temporal patterns of spike times, Segundo et al. [1963] showed that the precise timing of a pre-synaptic barrage of input influenced the production of a post-synaptic response in Aplysia.

Since then, the idea that the fine temporal structure of the spike train encodes information about the features of the stimulus has gained interest [Bialek et al., 1991; Rieke et al., 1997; Shadlen and Newsome, 1994; Softky, 1995], largely because neural codes that utilize the precise timing of the spike can make more efficient use of the capacity of neural connections than those that simply rely on the average firing rate [Mainen and Sejnowski, 1995].

How could neurons detect temporal patterns in spike times? Instead of acting as integrate-and-fire devices, neurons could function as “coincident detectors” by detecting temporal

coincidence of synaptic inputs. This would require the detection of temporally simultaneous synaptic potentials on the scale of milliseconds [König et al., 1996].

A crucial difference between the concept of an integrate-and-fire neuron and a coincidence detector is apparent when one considers the conversion of incoming synaptic potentials to the actual spiking output of the receiving neuron. If neurons act as integrators, then most - if not all - incoming synaptic potentials contribute to the generation of the action potentials. One the other hand, if neurons act as coincidence detectors, most of the incoming synaptic potentials contribute very little to the generation of output signals [König et al., 1996].

Differences also exist in terms of the processing dynamics: if integration is the key, then the precise timing of afferent signals is irrelevant because the output response of the neuron is not affected by temporal patterns in the input (all information is encoded in average firing rates) and no information can be carried by the precise timing of action potentials. In contrast, for neurons that act as coincidence detectors, the precise temporal structure of the afferent activity is important, and the generated output reflects temporal patterns in subsets of the inputs [König et al., 1996].

Experimental evidence suggests that some neurons do indeed act as coincident detectors rather than integrators. In the locust olfactory system, Perez-Orive et al. [2002, 2004] showed that the post-synaptic Kenyon cells act as coincident detectors that are sensitive to synchronized inputs that they receive from the antennae lobes. Additionally, neurons in the the nucleus laminaris of the the barn owl detect delays in the ipsilateral and contralateral afferent input from the nucleus magnocellularis to localize sound in the azimuth [Carr and Konishi, 1990]. Finally, convergent evidence in the rodent whisker system suggests that layer 4 neurons act as coincident detectors because the time window during which EPSPs can summate to drive spiking responses can be as short as 1 ms [Cruikshank et al., 2007; Gabernet et al., 2005], suggesting that spike timing plays an important role in information transfer from the thalamus to cortex [Usrey, 2002].

2.3 Multiple Operating Regimes

Considering that neural coding is most likely extremely complicated, it has also been suggested that neurons could use multiple strategies to extract information from noisy spike trains. Masuda and Aihara [2007] have suggested possible ways in which neurons could use multiple coding strategies.

In the first example, what they refer to as the “simultaneous use of multiple codes”, they hypothesize that a neuron or an ensemble of neurons is able to use more than

one code simultaneously. Because firing rates and the detection of coincident events can be independently modulated by different stimuli, they could theoretically operate as distinct codes. Examples of the simultaneous use of multiple codes abound. Recordings from salamander and rabbit retinal neurons suggests that information on time is encoded in spike timing, whereas information on object identity is encoded in spike counts [Berry et al., 1997].

Another example of a dual-coding strategy would be the spatial segregation of neural codes, such that one part of the brain uses one strategy, and another part of the brain uses another [Lu et al., 2001]. In the olfactory system of the locust, oderant identity is encoded in the olfactory receptor neurons by their mean firing rates, and these receptors send their input to the projection neurons of the antennae lobe, which respond to odor identity by modulating their mean firing rate. Additionally, however, particular epochs of the spiking responses of individual projection neurons are synchronized with the spike patterns of other responsive projection neurons (depending on the odor identity), and the resulting spatial-temporal patterns carry information about the odor identity that could not be obtained simply by examining the mean firing patterns of the neurons [Theunissen, 2003].

Similarly, in a simulation of a multilayer feed-forward network, Reyes [2003] found that synchrony developed in successive layers of the network, even when the initial input was uncorrelated. Firing rate was represented by a classical rate code in the initial layers of the network, but switched to a synchrony-based code in the deeper layers, suggesting that synchrony was critically involved in the propagation of rate signals across layers.

In summary, it seems reasonable to expect that neurons use diverse strategies to communicate information throughout the brain, and furthermore, that these computational strategies are related to the way information flow is organized throughout the brain. Whereas low-level neurons may encode basic stimulus features through modulation of firing rates, higher-order neurons may use temporal information extracted from the synchronous activation of neurons (e.g., the binding problem). We will see in the next section that much of the information processing accomplished by the brain involves organized and hierarchical processing stages that transform simple stimulus features into complex and composite stimulus perceptions.

3 Hierarchical Processing, Selectivity, and Sparseness

3.1 Sensory Processing: Feedforward, Hierarchical Networks

Conceptually, sensory perception is often discussed in the context of a feed-forward, hierarchical network, where complex computations are broken down into cascades of simpler operations that can be implemented by individual neurons [Riesenhuber, 2012]. As information passes though successive stages of processing, low-level information about the stimulus is transformed into more useful, higher-level representations. Neuronal response latencies and average receptive-field size increase as one proceeds from one area to the next, and the response properties of the neurons become more complex and non-linear [Ungerleider and Haxby, 1994].

The visual system provides an excellent example of hierarchical processing. In the visual system, the ventral visual stream for object recognition can be characterized by its simple-to-complex processing that utilizes a hierarchy of brain areas. The primary visual cortex V1 sits at the bottom of this hierarchy and receives its main feed-forward input from neurons in the lateral geniculate nucleus of the thalamus. Neurons in V1 function as simple spatio-temporal filters [Ungerleider and Haxby, 1994] and detect features of the stimulus such as oriented edges. As one moves to the secondary visual cortex, V2, neurons respond to simple edge combinations that form angles [Ito and Komatsu, 2004] and to illusory contours of figures [von der Heydt et al., 1984]. In V4, neural responses become much more selective, and some neurons only respond to a stimulus if it stands out from its background due to a difference in color or form [Desimone and Schein, 1987]. Finally, in the inferior temporal cortex, neurons respond selectively to global features of a stimulus (e.g., shape), and a small subset are specialized for faces [Desimone et al., 1984]. As one can see, visual processing involves several processing stages, from the low-level feature extraction in primary visual areas to the complex processing related to perceptual interpretation in higher areas [Lamme et al., 1998].

It is important to note that although feed-forward, hierarchical networks are observed across sensory modalities and animal models, anatomical studies have shown that feed-forward projections are often reciprocated by feed-back projections [Salin and Bullier, 1995]. Such recurrent networks, where higher-order neurons project back to lower-order neurons, could mediate “top-down” aspects of sensory processing (such as the role of selective attention; [Ungerleider and Haxby, 1994]). Although it is convincing that many sensory modalities extract information from stimulus features using feed-forward hierarchical processing, it is important to remember that the story is much more

complicated, and the effect of feed-back connections on the computations performed by these networks is not well-understood.

3.2 Disentangling Selectivity and Sparseness

The concept of a selective neural response and its roles in neural computation is a fascinating subject that has been explored from many different perspectives. In describing the exact nature of neural selectivity, the terms "selectivity" and "sparseness" are often used side by side.

A stimulus-selective response is typically thought of as the increased activity of a neuron in response to one stimulus versus another. In the songbird field, stimulus-selective responses are often quantified by comparing the normalized evoked firing rates of one stimulus to another (i.e., the d-prime score). The logic, then, is that a neuron will fire more in response to the "preferred" stimulus to which it is selectively tuned, and less to the comparison stimulus. In this sense of the word, selectivity refers to increased activity in response to a particular stimulus in comparison to another stimulus.

The term "sparse" is also used to describe highly selective neural responses. In this case, sparse refers to the term *lifetime sparseness*, which is defined as the fraction of the stimuli that are presented to those that elicit a significant response [Wolf et al., 2010]. In this case, a neuron may fire rarely in response to a stimulus, but when the neuron does respond, its response is very large [Willmore and Tolhurst, 2001]. Lifetime sparseness is related to stimulus selectivity and is a property of single neurons.

Another use of the term sparseness refers to the fraction of neurons within a population that respond to a single stimulus, known as *population sparseness* [Wolf et al., 2010]. In this case, a sparse neural response for a population of neurons is thus characterized by a small percentage of responsive neurons within a large population of silent neurons [Wolf et al., 2010]. Barlow [1972] observed that in many sensory nervous systems, neurons at later stages of processing are generally less active than those at earlier stages.

Population sparseness is often quantified by calculating the kurtosis of a response distribution [Vinje and Gallant, 2000]. As the distribution of responses becomes more sparse, the proportion of moderate responses decrease and the proportion of both large and small responses increases, which is reflected by an increase in the kurtosis of the distribution [Vinje and Gallant, 2000].

3.3 Sparse and Efficient Codes

Many theories exist about the nature of neural codes, and sparse and dense codes are two such theories that lie at opposite ends of the spectrum. Dense codes usually involve the activation of a large proportion of neurons at all times, in which each neuron contributes a small amount to the representation of the stimulus [Willmore and King, 2009]. Sparse codes, on the other hand, can transmit information with minimum redundancy and relatively few spikes, making them both metabolically and informationally more efficient than dense codes [Vinje and Gallant, 2000]. Sparse coding models are considered to be efficient because, as a stimulus engages only a small fraction of neurons, and each neuron produces a highly selective response [Hromádka et al., 2008], action potentials remain relatively rare events. The minimization of the number of action potentials is metabolically efficient, considering that action potential production is a major part of cortical energy consumption [Attwell and Laughlin, 2001].

Evidence for sparse coding by populations of neurons has been demonstrated for several sensory modalities including vision [Olshausen and Field, 1996; Vinje and Gallant, 2000], audition [Hromádka et al., 2008; Lewicki, 2002], somatosensation [Brecht and Sakmann, 2002], and olfaction [Poo and Isaacson, 2009]. Mormann et al. [2008] have shown that even in human medial temporal lobe, visual information is processed hierarchically and exhibits both lifetime and population sparseness [Waydo et al., 2006].

3.4 How Powerful Are Solitary Spikes?

Much of the theory behind sparse and efficient codes relies on the assumption that single spikes impact neural computations, but is that really the case?

London et al. [2010] used intracellular and extracellular recordings to show that by adding a single spike to the output normally produced by a neuron, a detectable increase in firing rate was observed in the local network, suggesting that single spikes can indeed shape local neural computations.

Furthermore, in vivo studies in rodents have demonstrated that single spikes can have a powerful effect on behavior and perception. Huber et al. [2008] trained mice to report photostimulation of channelrhodopsin-2 expressing pyramidal cells in the somatosensory cortex and found that the activation of approximately 60 neurons was sufficient to drive reliable detection of photostimulation.

In a similar study, Houweling and Brecht [2008] showed that short trains of about 14 action potentials induced by electrical microsimulation of a single neuron in the barrel

cortex induced a behavioral response in a small but significant fraction of trials. Both sets of evidence indicate the powerful effects that single neurons can have on behavioral output.

Overall, we see that that distributed, sparse activity of single neurons within a population can have a strong impact on neural computations and may be an efficient strategy in the coding of complex perceptions. In the following section, we will examine in greater detail how populations of neurons interact with each other, and how correlated population activity can impact neural computations.

4 Population Codes: Spike Train Independence

Neurons are embedded in circuits. Although single neurons can display remarkable properties (e.g., stimulus-selectivity), it is the composite interactions of populations of neurons that lead to complex behavior and perception. If we assume that the responses of one neuron do not co-fluctuate with the responses of other neurons, that is, that neural responses are independent and individual neurons do not share correlated intrinsic noise, we understand quite a lot about how populations of noisy neurons could encode information [Pouget et al., 2000; Seung and Sompolinsky, 1993] and how networks could extract information by decoding such population codes [Deneve et al., 1999]).

However, copious evidence suggests that spike trains from different neurons are not independent and that intrinsic noise is correlated across neurons. Correlated intrinsic noise has been demonstrated in the visual cortex of macaque [Kohn and Smith, 2005; Womelsdorf et al., 2012], the extrastriate cortex of rhesus monkeys [Huang and Lisberger, 2009], the auditory cortex of mice [Rothschild et al., 2010], and the somatosensory cortex of rats [Greenberg et al., 2008]. Because the activity of pairs of neurons is correlated, the activity of a network of neurons cannot be fully understood through measurements of its individual responses [Ohiorhenuan et al., 2010].

The strength of correlated neural activity between neurons can depend on the parameter such as the similarity of preferred stimuli [Kohn and Smith, 2005; Zohary et al., 1994], the distance between the neurons [Constantinidis and Goldman-Rakic, 2002; Ohiorhenuan et al., 2010; Smith and Kohn, 2008], and attention [Roelfsema et al., 2004; Tiesinga et al., 2005]. Furthermore, correlations in the membrane potential fluctuations, which reflect the input of the neurons, can also be modulated by external stimuli [Lampl et al., 1999; Yu and Ferster, 2010].

4.1 Correlations of Spike Counts

Experimentally, correlated trial-to-trial variability, or intrinsic noise correlations, are often assessed using Pearson's correlation coefficient and calculated from spike counts measured over durations that last from hundreds of milliseconds to several seconds [Cohen and Kohn, 2011]. Although this measure informs us about how neurons may co-vary their firing rates together, it does not inform us about how the timing of action may be correlated across neurons.

Several studies have used spike count correlations to examine correlations in populations of neurons, and it has become something of a standard tool for the investigation correlations in populations of neurons.

Direction-selective neurons in middle temporal visual area (MT or V5) have been thought to provide the signals necessary to discriminate the direction of movement in the random dot patterns, and Zohary et al. [1994] obtained data from 100 neuron pairs in MT from three rhesus monkeys as they viewed random dot patterns presented on a video monitor. A single electrode was used to record from nearby pairs of neurons. The directional tuning of the pair of neurons was assessed as the monkeys maintained fixation during presentation of the visual stimulus withing the receptive fields of both neurons. The monkeys were then asked to report the direction of coherent motion of the random dot patterns during a stimulus discrimination task. The authors combined data across stimulus conditions and over the two behavioral tasks (fixation and discrimination), because a chi square test found that in 89% of experiments, the resulting spike count correlations were stimulus-independent and that there was no difference between the two behavioral tasks (paired t-test, $p > 0.75$; [Zohary et al., 1994]). The authors found that the mean correlation coefficient for 100 pairs of neurons was 0.12 and significantly greater than zero (t-test, $p < 0.0001$), and concluded that adjacent MT neurons covary weakly in response to the visual stimuli. They also found that the strength of the correlation coefficient significantly depended on the preferred direction of the two neurons in the neuron pair, with the mean correlation coefficient being significantly less for neurons pairs that differed in their directional tuning by more than 90 degrees [Zohary et al., 1994].

Another study examined the role of correlations in nearby neurons pairs in the visual cortex of macaque monkeys. Ecker et al. [2010] used chronically implanted tetrodes to record from ensembles of neurons in the visual cortex of awake macaque monkeys during static or drifting grating visual stimulation. The size of the gratings were large enough to cover the receptive fields of all neurons recorded. The authors found that correlation coefficients for nearby neurons pairs recorded by the same tetrode were exceedingly low (0.005 ± 0.004), and furthermore, that neurons with similar preferred

orientations also had very weak correlation coefficients (0.028 ± 0.010) [Ecker et al., 2010]. Such weak correlations between spatially nearby pairs and similarly tuned neurons stands in contrast to previously reported correlations coefficients on the order of 0.1 to 0.3, and highlight the confounds that differences in experimental techniques, including anesthesia, electrode recording techniques, and correlation window size, may have on the estimates of trial-to-trial variability.

In addition to investigations of noise correlations during visual processing, perceptual learning has also been shown to influence spike count correlations. Gu et al. [2011] trained macaque monkeys to either passively fixate on a stimulus ("naive" animals) or to perform a heading discrimination task ("trained" animals). The authors recorded neural responses in the dorsal medial superior temporal area (MSTd) and found that in trained animals, spike count correlations were significantly weaker when compared to naive animals [Gu et al., 2011].

Similarly, attention has also been linked to a reduction in spike count correlations. Mitchell et al. [2009] recorded from area V4 in macaque monkeys trained to either fixate on a stimulus that was located within the receptive fields of a neuron pair or on a stimulus that was located outside of the receptive fields of the neuron pair. When attention was directed to a stimulus inside a neuron's receptive field, spike count correlations were reduced. The authors concluded that the attention-dependent reduction of correlated activity improves the signal-to-noise ratio of pooled neural signals substantially more than attention-dependent increases in firing rate [Mitchell et al., 2009].

4.1.1 Experimental Factors that Affect Spike Count Correlations

Overall, Pearson's correlation coefficient (R_{SC}) is an analytical method often used to probe pairwise covariation in firing rate within populations of neurons. Although the resulting measure, known interchangeably as spike count correlation, noise correlation, correlation coefficient, or trial-to-trial variability, provides a value that seems to measure the strength of covariation between two neurons, it is affected by a large range of experimental factors that must be taken into consideration when comparing results across experimental paradigms.

For example, when the Pearson's correlation coefficient is used to analyze data containing few spikes, the resulting R_{SC} tends to be small [Cohen and Kohn, 2011]. The number of spike that a neuron produces is determined by both the neuron's underlying firing rate and the time window over which the correlation coefficient is calculated, and both low firing rates and short response durations can lead to low R_{SC} values. Cohen and Kohn [2011] performed a simulation examining the effect of firing rate on R_{SC} and

found that its magnitude depended more on the minimum response of the the neurons than their geometric mean rate. For example, a neuron pair in which one neuron has a mean firing rate of 0.01 Hz and the other 100 Hz has the same geometric mean firing rate as a neuron pair in which both neurons fire at 1 Hz. However, the measured R_{SC} for the first pairs is much reduced compared to the second pair [Cohen and Kohn, 2011]. This effect is illustrated in Figure 2.4A.

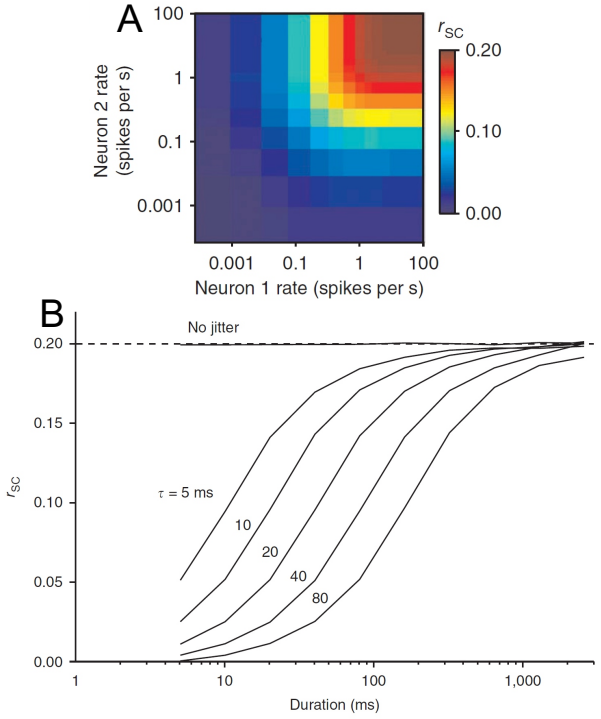

FIGURE 2.4: **R_{SC} is affected by firing rate and response duration. A)** R_{SC} is small for neurons with low firing rates. Confusion matrix depicts the effect of firing rate on measured R_{SC} values, where warm colors represent large correlation coefficients, and cool colors represent zero correlation. Vertical and horizontal bands are present, indicating that R_{SC} depends more strongly on the minimum firing rate in the pair than the geometric mean. A dependence on the geometrical mean would be indicated by diagonal stripes from the top left to bottom right. **B)** Curves depict measured R_{SC} as a function of response duration ranging from 5 to 2000 ms. The number next to each curve represents the s.d. of the Gaussian jitter in ms. For spike train with jitter at large time scales, longer response durations are required to accurately estimate the full strength of the correlation. Figure adapted from [Cohen and Kohn, 2011].

Furthermore, spike count correlations are systematically underestimated if the response duration over which the correlation coefficient is calculated is shorter than the jitter in the timing of the coincident spikes. Cohen and Kohn [2011] simulated spiking responses for pairs of neurons and controlled the timescale of correlation by adding a small number of common spikes to otherwise independent (Poisson-distributed) spike trains. They then jittered spike times using a range of Gaussian distributions whose standard deviations (s.d.) varied from 5 to 80 ms. By measuring R_{SC} as function of time window duration for the jittered spike trains, they found that for spike train where

coincident spikes were jittered at large time scales (e.g. 80 ms s.d.), longer response durations were required to capture the full strength of the correlations (Fig. 2.4B).

In addition to firing rate and response duration affecting R_{SC}, spike sorting errors may also confound measurements of R_{SC}. For experimental techniques that make use of low-impedance electrodes or tetrodes to record from several units simultaneously, care must be taken so as not to mistake multi-unit responses for single units. Combining several different units into one averages out each neuron's independent variability, and the result is that correlations calculated on clusters of multi-unit activity are larger than between pairings of the individual constituent neurons. Finally, internal factors such as arousal, attention, motivation, time-dependent variations in anesthesia level may also impose slow fluctuations in the brain state of the experimental animal and lead to confounds in the measurement of R_{SC} [Cohen and Kohn, 2011].

4.2 Correlations of Spike Times

The spike train cross-correlogram [Perkel et al., 1967a] is a statistical method that can reveal probabilistic patterns in spike timing that may exist between neuron pairs. Variations of this method have been used extensively to detect correlated patterns in the spiking output of pairs of neurons [Bair et al., 2001; Eggermont, 2006; Eggermont et al., 1983; Huang and Lisberger, 2009; Kimpo et al., 2003; Kohn and Smith, 2005].

The theory of this statistical measure depends on the basic result that if two spike trains are independent (i.e., the occurrence of a spike from one neuron is not correlated with the subsequent activity of the another neuron, and vice versa), the cross-correlation function between those spike trains will be flat. This assumes that spikes in train A occur at random moments with respect to train B. When the cross-correlation function is not flat, we can conclude that some functional correlation exists between the two neurons [Moore et al., 1970].

As such, this method can be used to detect synaptic connections among pairs of neurons; however, some care must be taken in interpreting the statistics. In particular, it is not always clear whether a given cross-correlation is better explained by a synaptic connection between the two neurons (potentially mediated by other neurons), or is alternatively shared by a source of input to the two neurons [Perkel et al., 1967a].

Indeed, there are a number of distinct classes physiological phenomena which might induce a correlation between the spike trains of two neurons [Brody, 1999]. Both neurons might be synaptically connected, with one neuron producing EPSPs or IPSPs in the other, or they may receive common input from pre-synaptic sources. Therefore,

an observed dependence between the two spike trains can arise from one (or both) of two sources: 1) a functional interaction, or any mechanism by which the firing of one neuron influences the firing by the other neuron, and 2) common input, or any mechanism that simultaneously modulates the firing patterns of both neurons. Furthermore, it is also possible to attribute independence to spike trains that are in fact dependent, if, for example, the dependence may be so weak that its effects are indistinguishable from "noise" [Perkel et al., 1967a].

In the following sections, we will introduce and discuss types of dependences observed for different classes of functional correlations and how they can be detected from cross-correlation functions.

4.2.1 Excitatory Synaptic Coupling

The most elementary interaction is that exhibited by two neurons A and B connected by an excitatory synapse, such that an axon collateral of one of the neurons forms a synapse with the other (Fig. 2.5A). In this case, every firing of neuron A is followed (after the conduction and transmission delay) by a EPSP in neuron B, leading to an enhanced firing of spikes by neuron B. A cross-correlation function between two such monosynaptically connected neurons thus shows a central peak near but offset to the origin (e.g. at time lag of around 10 ms; Fig. 2.5B). The shape of the central peak can be influence by other factors, such as bursts in the pre and post-synaptic neurons, and fluctuations in the conduction delay [Moore et al., 1970].

If one of the two neurons displays rhythmicity in its firing, then these secondary effects will also be evident in the cross-correlation function as secondary peaks at long time lags (Fig. 2.5B). This is due to the fact that to the extent that the post-synaptic neuron is a "follower" of the pre-synaptic neuron, the cross-correlation will reflect the characteristics of the driver neuron's autocorrelation [Moore et al., 1970].

4.2.2 Inhibitory Synaptic Coupling

In the same vein, two neurons can also be connected to each other by an inhibitory synapse (Fig. 2.5C). In this case, following an impulse in the pre-synaptic neuron, there will be a reduced probability of firing in the post-synaptic neuron. The primary effect of inhibition on the cross-correlation is a trough near the origin (Fig. 2.5D). Of course, for this interaction to be visible in the cross-correlation, there must be sufficient back ground firing of the post synaptic neuron to observed a suppression in firing [Moore et al., 1970].

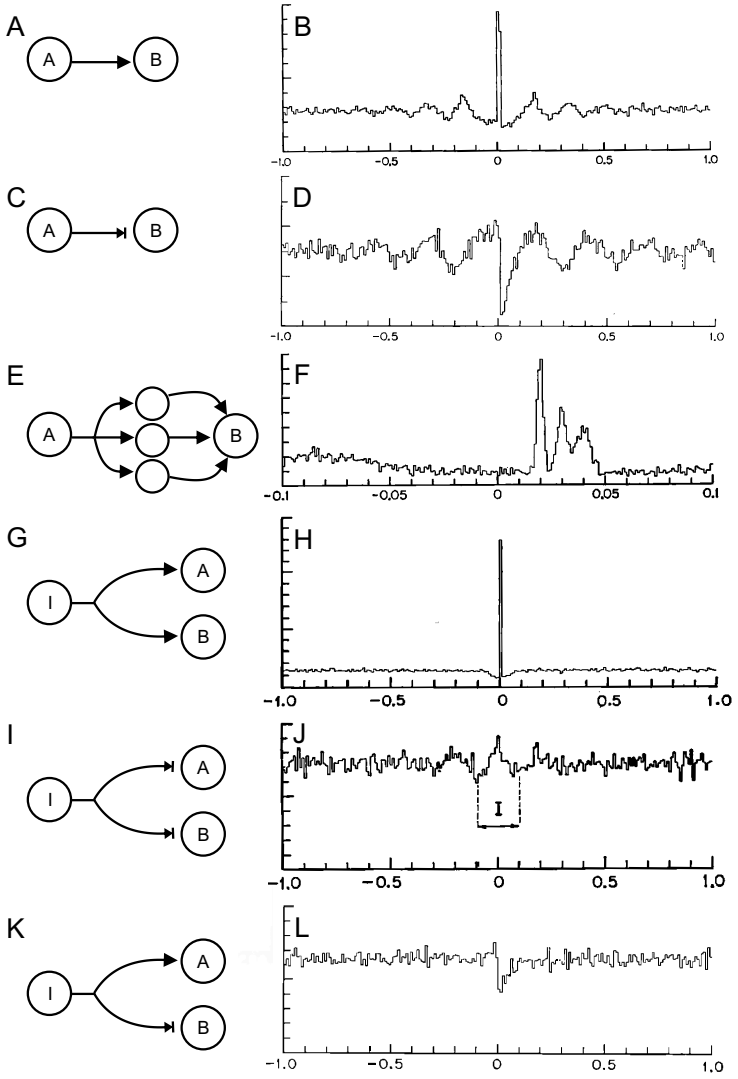

FIGURE 2.5: **Patterns of spike time dependencies. A, B)** Monosynaptic excitatory coupling (**A**) shows a central peak offset to origin in the cross-correlation (**B**). Secondary peaks in the cross-correlation reflect rhythmic firing of the leading neuron. **C, D)** Monosynaptic inhibitory coupling (**C**) shows a central trough offset to origin in the cross-correlation (**D**). Secondary peaks in the cross-correlation reflect rhythmic firing of the leading neuron. **E, F)** Indirect excitatory coupling (**E**) shows peaks offset to origin by a time determined by the latencies of the involved neurons in the cross-correlation (**F**). See text for details. **G, H)** Shared excitatory input (**G**) shows a central peak at the origin in the cross-correlation (**H**). **I, J)** Shared inhibitory input (**I**) shows a paradoxical central trough at the origin in the cross-correlation (**J**), see text for details. Shared inhibitory input is revealed as paracentral troughs offset to the origin (indicated by I). **K, L)** Shared reciprocal input (**K**) shows a central trough at the origin in the cross-correlation (**L**). Filled arrows indicate excitatory connections, barred arrows indicate inhibitory connections. Figure adapted from [Moore et al., 1970].

4.2.3 Indirect Excitatory Coupling

In addition to direct influences that neurons may have on each other, it is also possible that neurons may influence each other indirectly though the activity of other neurons.

Figure 2.5E depicts one such example, where the the output of Neuron A produces excitatory responses in three interneurons after a delay of 10 ms. In turn, the output of the interneurons drive Neuron B after delays of 10, 20, and 30 ms. Dependencies between Neuron A and B are revealed as peaks offset to the origin at longer time lags that incorporate both the latency between Neuron A and the interneurons in addition to the latency between the interneurons and Neuron B (Fig. 2.5F).

4.2.4 Shared Synaptic Input

Shared synaptic input is quite common in neural networks, and such common input will induce correlations in the discharges of those neurons that are influence by the common source. A common driver will introduce complexities not found in direct connections, and pose additional problems for cross-correlations between the post-synaptic neuron pair [Moore et al., 1970], and common sources of input are generally more difficult to detect than direct or indirect connections [Perkel et al., 1967a]

In the case of shared excitation, common excitatory input received by both neurons, as depicted in Figure 2.5G, is reflected in a tendency toward synchronous firing that generates a peak at the origin (Fig. 2.5H). If, however, the shared excitatory input is received by two neurons after different delay times, the peak will be shifted by an amount equal to the difference in conduction times [Moore et al., 1970]. The secondary effects of the cross-correlation function will correspond to feature of the autocorrelations of both neurons A and B.

In the case of shared inhibition (Fig. 2.5I), common inhibitory input tends to synchronize post synaptic neuron periods of non-firing, leading to the paradoxical result that if their periods of non-firing overlap, then their periods of uninhibited firing must also overlap, and they tend to fire in a correlated way. This uninhibited correlated firing causes a central broad peak (Fig. 2.5J) that is usually not large and may have considerable spread [Moore et al., 1970]. The time of occurrence of shared inhibitory input are unlikely to be in the region of 0 time lag, and therefore appear as two paracentral troughs (Fig. 2.5J; indicated by I).

In the case of shared reciprocal input, where neuron A receives and EPSP and neuron B receives and IPSP from a common source (Fig. 2.5K), the effect will be observed as an anti-synchrony of firing in the post-synaptic neurons and will appear as a trough in the cross-correlation at time lag 0 (Fig. 2.5L). Caution should be taken not to confuse this effect with direct inhibition [Moore et al., 1970].

If rate changes are shared by two otherwise independent neurons, the cross-correlation will typically display a symmetrical elevation above its "null" level in the neighborhood of the origin [Perkel et al., 1967a].

4.2.5 Detecting Stimulus Effects on the Cross-Correlation

The cross-correlation can also be used to compare the effects of stimulation on the spiking output of the neurons pairs. These effects on the cross-correlation may arise through 1) changes in firing rates of one or both neurons 2) though direct or indirect synaptic input to both neurons from a common source that responds to the stimulus 3) though the effects of the stimulus on interaction pathways though the two observed neurons 4) though any combination of these [Perkel et al., 1967a]

It is clear that the observed cross-correlation function between two neurons A and B will be in general different under the "stimulus-on" and "stimulus-off" conditions. In this case, the basic assumption is that the modifications of the cross-correlation produced by the stimulus are sufficiently independent to be additive [Perkel et al., 1967a].

One way to predict the contribution of shared inputs from the stimulus is to isolate the effects that are time-locked to the stimulus by thoroughly shuffling the responses of one of the neurons and calculating a shuffled cross-correlation between the original spike train of A and the shuffled spike train of B. The shuffling destroys all significant time relationships between the two trains except those related to the stimulus presentations [Perkel et al., 1967a].

4.3 Synchrony

The cross-correlation function can be used to detect synchronous spiking activity in neuron pairs. Synchrony can be defined as coincident spiking events that occur across multiple neurons [Tchumatchenko et al., 2011], and its role in neural signaling is still debated. On one hand, it is thought that synchrony is involved in the transmission of temporally precise signals and that ensembles of neurons become more correlated during a behavioral task or during sensory stimulation [Hatsopoulos et al., 1998; Panzeri et al., 2001; Riehle et al., 1997]. On the other hand, if neural signals are propagated by rate codes, then the precise timing of action potentials is irrelevant, and spiking synchrony may even compromise rate coding [Shadlen and Newsome, 1998]. However, much evidence suggests that synchronous spiking events are used in a variety of situations, including stimulus selection [Cardoso de Oliveira et al., 1997], information binding [Friedrich et al., 2004], and attention [Fries et al., 2001; Steinmetz et al., 2000].

Synchrony is thought to influence the transmission of activity from one group of neurons to another, in addition to possibly coding information related to stimulus parameters. However, factors that affect the reliability of spike trains will also affect the synchrony of spike trains. For example, neurons receiving identical input may respond with very different spiking responses, and therefore, even during highly uniform stimulus conditions, subsets of neurons with similar tuning properties may be synchronized [Ermentrout et al., 2008]. Similarly, network connectivity may bias the responses in favor of or against synchronous responses [Whittington et al., 1995].

How does spiking synchrony actually influence the transmission of information from one populations of neurons to another? Synchronous spiking allows groups of neurons with the same post-synaptic targets to more effectively depolarize these targets, leading to better information propagation to downstream targets [Salinas and Sejnowski, 2001]. Indeed, evidence suggests that synchronization of thalamocortical inputs maximizes the transfer of information from the thalamus to the cortex [Wang et al., 2010a], suggesting that precise temporal correlation could be used as an additional information channel from thalamus to visual cortex [Dan et al., 1998].

5 Information Theory Approaches

Information theory approaches have been used to determine how correlations in populations of neurons impact the overall amount of information that can be coded by a population of neurons. As one might anticipate, the effects of intrinsic noise correlations on the information encoded by a population of neurons are diverse and largely depend on the tuning of the constituent neurons. In these studies, information is usually estimated by either calculating Shannon's mutual information [Shannon, 1948] or the Fisher information.

The mutual information provides an absolute bound on the performance of a classifier that must discriminate between several different stimuli [Gordon et al., 2008], and it is defined as the information gained about the stimulus by knowing the neural response [Quiroga and Panzeri, 2009]. In the case where the the stimuli and the responses are completely independent, the mutual information equals zero, otherwise it takes positive values. Alternatively, knowledge about the stimulus can be measure using Fisher information, where the inverse of the Fisher information provides a lower bound to the mean decoding error obtained with an unbiased decoder [Quiroga and Panzeri, 2009]. Fisher information is often utilized to evaluate the impact of correlations on the accuracy of a population code [Abbott and Dayan, 1999; Smith and Kohn, 2008; Sompolinsky et al.,

2001], or the network efficiency of a population of neurons [Gutnisky and Dragoi, 2008; Quiroga and Panzeri, 2009].

Early studies, which examined the effect of noise correlations in a population of homogeneously tuned neurons, found that even weak (positive) covariation among pairs of neurons can greatly limit the signal-to-noise ratio of any stimulus represented by the activity of a pool neurons, and that this effect is not counteracted by increasing the neuronal pool size [Zohary et al., 1994].

Other studies using population network models found that the effect of intrinsic noise correlations depend on the polarity of the correlations: while positive correlations could decrease the estimation capacity of the network, negative correlations could substantially increase the information capacity [Sompolinsky et al., 2001].

In heterogeneously tuned neuron populations, the most relevant requirement for improved accuracy in information coding is that neurons must have different selectivities to the variables they jointly encode [Abbott and Dayan, 1999]. The information capacity of a heterogeneous network is not limited by shared intrinsic noise, but rather scales linearly with the number of cells in the population [Shamir and Sompolinsky, 2006].

Finally, [Ecker et al., 2011] studied the effect of limited range correlations in heterogeneous population models, and found that reducing spike count correlations does not necessarily improve encoding accuracy, and rather, for populations of neurons greater than several hundred, limited range correlations could substantially improve encoding accuracy. Specifically, for biologically plausible parameters, such as a population of thousands of neurons with heterogeneous tuning curves, increasing spike count correlations increased the overall Fisher information.

In closing, although there is not a clear answer to how noise correlations affect the amount of information available to a population of neurons, it seems that the ability of a population of neurons to code information depends on the details of the correlations, i.e., the polarity of the noise correlations and their relationship to signal correlations of the neurons.

Chapter 3

The Auditory Brains of Songbirds

1 Introduction

Songbirds have long captivated humankind with their ability to socially communicate through their beautiful and complex songs and calls. Birdsong is not only music to our ears, but also a highly complex acoustic signal that changes dynamically over time. And like any communication signal, birdsong transmits information between the sender of the signal and the receivers. The songbird's auditory system must therefore be able to interpret a continuous stream of acoustic information - including the vocalizations of other birds and other animals in an environment filled with ambient noise.

The ability to process these acoustic signals, to identify the signal in the noise, and to arrive at a behaviorally relevant communication message is one reason that animals depend on auditory processing for survival. Such auditory processing allows a songbird to discriminate between mate and non-mate, relative and non-relative, and neighbor and stranger - purely on the basis of hearing a song [Sherman et al., 1997].

But why study auditory processing in songbirds, specifically? Surely mice and cats are able to hear as well as birds do and must also rely on auditory processing to make smart behavioral decisions. This is true, and a large amount of work in the field of auditory processing has been done in rodents [Christianson et al., 2011; Hromádka et al., 2008; Linden et al., 2003; Rothschild et al., 2010], cats [Eggermont, 2006; Moshitch

This chapter is based on the following publication: Ondracek, J. M. and Hahnloser, R. H. R. (2013). "Advances in Understanding the Auditory Brain of Songbirds." In Fay, R. R. and Popper, A. N. (Eds.), **Springer Handbook in Auditory Research: Insights from Comparative Hearing Research.** (*in publication*).

et al., 2006], and ferrets [Atiani et al., 2009]. Songbirds, however, have one advantage. While most animals rely on the ability to process auditory signals in order to classify the source of the signal (i.e., whether it originates from a conspecific, a stranger, or a mate), songbirds - like humans - also depend on auditory processing to learn their vocalizations. The ability to memorize and learn from the auditory input it receives in order to shape its own vocalizations implies that the songbird brain, not unlike the human brain, is uniquely built to process and discriminate complex sounds.

2 What Makes Songbirds So Special?

Of all the birds that produce sounds, only 3 of the 23 major bird orders possess the ability to learn from the environment and produce new sounds [Jarvis, 2004]. These birds are songbirds (order *Passeriformes*), parrots (order *Psittaciformes*), and hummingbirds (order *Trochiliformes*). Unlike the 20 other major bird groups which rely solely on basal brain structures to produce their vocalizations, only the birds in these three groups have the necessary brain anatomy to produce learned vocalizations [Jarvis, 2004].

Being born with the necessary brain structures, however, is not enough for these birds to learn their songs. Like humans, songbirds, parrots, and hummingbirds must hear the sounds and songs of the adults of their species to develop normal adult vocalizations. If a songbird (or a human) is raised in isolation or deafened early in its life and never allowed to hear adult songs or speech, it develops abnormal vocalizations [Konishi, 1965].

In contrast, when birds such as chickens and pigeons are deafened or raised in auditory isolation, their adult vocalizations are normal. The songs and sounds of these birds are considered *innate* vocalizations because they do not require auditory feedback to be learned [Wilbrecht and Nottebohm, 2003]. A male chicken can live its whole life without ever hearing another rooster, and it still maintains the ability to sing its own“Cock-a-doodle-do.” On the other hand, songbirds, parrots, hummingbirds, and humans (as well as dolphins, whales, and bats) are considered vocal learners because they must develop their vocalizations through a sensory-motor learning program that crucially depends on the ability to hear the vocalizations of themselves and of others [Jarvis, 2004].

In this chapter, the strategy that songbirds use to learn their songs will be discussed, highlighting its dependence on auditory feedback for successful song learning. We will define the elements of birdsong, and then discuss the patterns of auditory processing that occur in the songbird's brain, beginning with the midbrain and thalamic structures that are common to all birds, moving up to the primary and secondary auditory areas in

the songbird cerebrum involved in the discrimination of behaviorally relevant complex sounds in birdsong, and finally to the unique pre-motor area HVC (used as a proper name) that is involved in song production.

3 Song Learning in Songbirds

For over 50 years, researchers have examined the areas of the songbird brain associated with song learning and song production. Of the more than 4,000 species of songbirds, many different types of songbirds have been studied in the laboratory, including chaffinches (*Fringilla coelebs*), canaries (*Serinus canaria*), white crowned sparrows (*Zonotrichia leucophrys*), zebra finches (*Taeniopygia guttata*), Bengalese finches (*Lonchura striata domestica*), swamp sparrows (*Melospiza georgiana*), and European starlings (*Sturnus vulgaris*) to name just a few. Most of what is known about birdsong neurophysiology, however, comes from studies of one particular songbird, the zebra finch. These boisterous birds breed easily in captivity, reach sexual maturity by 90 days post hatch (dph), learn a single, stereotyped song, and sing readily in the laboratory [Brenowitz and Beecher, 2005].

Although each of the aforementioned songbirds learns its song in a slightly different way, these birds are vocal learners in the sense that as juveniles, they memorize the target vocalizations by listening to adults during a passive sensory phase. During the subsequent sensory-motor phase, the young birds practice and perfect their songs until the sensitive period ends, after which the songs are considered crystallized [Konishi, 1965]. Although this general pattern of song learning applies to all songbirds, it is important to remember that there is a tremendous diversity in the individual learning strategies among species of songbirds, and factors such as the timing of song learning, the number of different songs that are learned, and the sexual patterns of song production vary widely between species [Brenowitz, 2002].

Indeed, song learning can be characterized along many dimensions. The timing of song learning can range from early sensitive period learners, called close-ended learners, such as the zebra finches, which learn their songs in under 90 days, to life-long, or open-ended learners, such as European starlings and canaries. The number of songs a bird learns to sing may range from one single stereotyped song, as performed by the zebra finch, to repertoires of more than 1,000 songs, like those of the brown thrashers (*Toxostoma rufum*). The manner of song imitation also varies. Some birds, such as the zebra finches, closely imitate the model song almost identically, whereas other birds, such as the sedge wrens (*Cistothorus platensis*), improvise completely new song elements to create novel songs. Similarly, some birds are able to copy almost anything that

they hear, such as the northern mockingbird (*Mimus polyglottus*), whereas other birds can only copy a song if it fits a set of tightly constrained, species-specific parameters. These differences highlight the difficulty in identifying a single typical birdsong learning program [Brenowitz and Beecher, 2005] but reveal the marvelous variety of song learning capabilities of songbirds.

3.1 The Auditory Template Theory of Song Learning

Despite the variety present in the song learning strategies of different species, the auditory template theory of song learning provides a good starting point for understanding how songbirds learn to sing. According to this model, a young bird is born with a rough innate template that defines the basic species-specific features of the song, such as tempo and frequency range. During the initial sensory phase, the young bird may be exposed to a variety of different bird songs, and only the songs that match the template are memorized for later use [Konishi, 1985]. Once the young bird begins to sing during the sensory-motor phase, the bird attempts to match its own singing to the template it has memorized in its early youth. The exact timing of the onset of singing depends on the species of bird, but it usually begins before the bird is one year old [Catchpole and Slater, 2008].

The bird does not immediately begin to sing a perfect replica of the memorized template song on its first day of singing; rather it is a gradual process which is critically dependent on the bird hearing its own attempts at vocalizing. The young bird begins to sing a version of the song known as subsong, as depicted in Figure 3.1A. This is a squeaky, noisy, low amplitude version of the adult song, which can be compared to the babbling of human babies [Aronov et al., 2008; Doupe and Kuhl, 1999; Marler, 1970b]. As the young bird practices its song, syllable structure develops, as shown in Figure 3.1B, and the bird begins to sing plastic song, a few examples of which are presented in Figure 3.1C and Figure 3.1D. During the plastic song phase, the song becomes louder, gains rhythm, and becomes more organized and structured; however, at this stage it is still not a perfect replica of the tutor's song. Finally, the bird produces crystallized song, shown in Figure 3.1E a song form that is highly stereotyped and faithfully matches the song of the juvenile's tutor, shown in Figure 3.1F [Marler and Peters, 1982].

3.2 Effects of Auditory Exposure During Song Learning

Copious evidence emphasizes the importance of auditory feedback and experience during the song learning process. Birds that are removed from their nests as nestlings

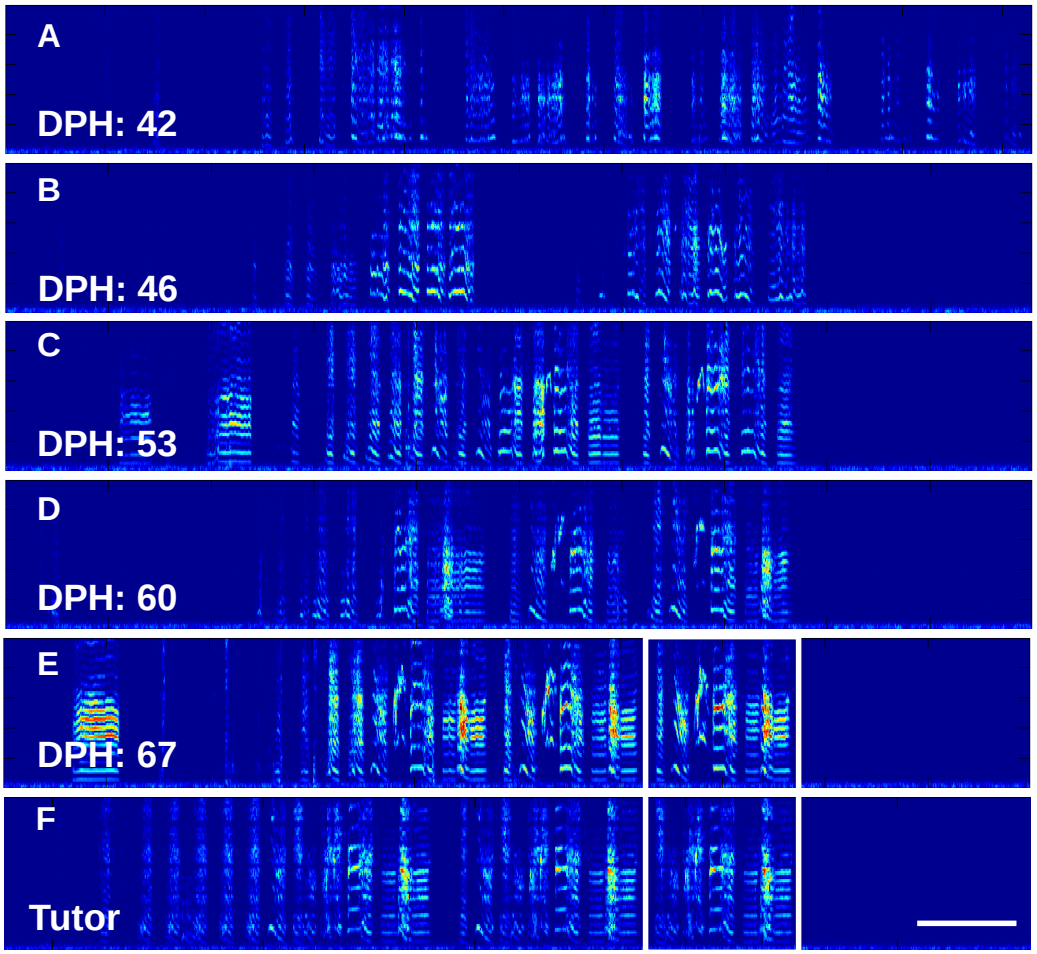

FIGURE 3.1: **The stages of song learning in a juvenile zebra finch. A-E)** Spectrograms of vocalizations recorded during the song learning of a juvenile zebra finch ages 42 to 67 dph. **A)** Vocalizations begin as noisy, unstructured versions of song called subsong. As the bird practices vocalizing, syllable structure develops (**B**), and the song develops into a more stereotyped vocalization known as plastic song (**C, D**). At the end of song learning (**E**), the vocalization becomes stereotyped and displays strong similarity to the tutor song (**F**). The white box indicates the motif in the song of the juvenile bird and the comparable motif in the tutor song. White horizontal line indicates 500 ms. All spectrograms are depicted in the range of 0 to 8 kHz. Data from Joshua Herbst.

and tutored with other adults of the same species will ultimately produce songs that resemble those of the foster parents [Marler, 1970a]. Furthermore, cross-fostering experiments, in which male zebra finch chicks are raised by Bengalese finches, reveal that zebra finches are able to learn Bengalese finch songs with as much accuracy as male birds learning from their natural father [Clayton, 1989; Immelmann, 1969].

Highlighting the importance of early auditory exposure to vocalizations, young birds that are raised in acoustic isolation produce abnormal isolate songs. In some of the first isolation experiments performed, Peter Marler hand-raised 5 day old male chaffinches

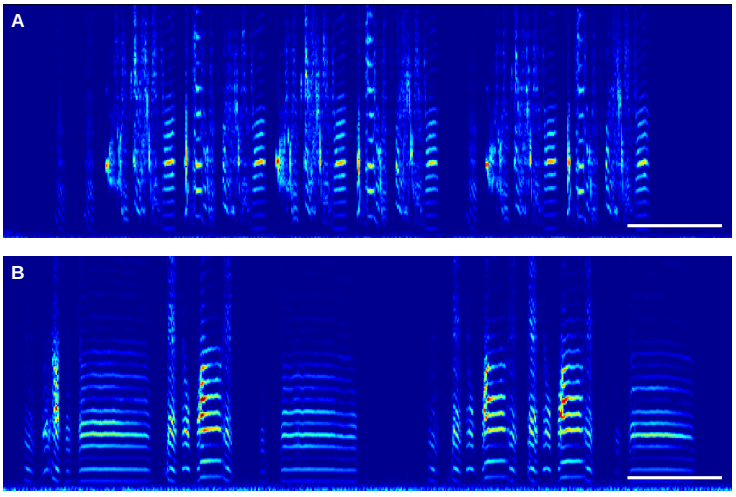

FIGURE 3.2: **Normal and isolate songs from adult zebra finches. A)** The result of unperturbed song learning results in a highly stereotyped and complex song containing motifs with a variety of syllables. **B)** When a juvenile bird is prevented from hearing other birds during the sensory phase of song learning, the bird develops a highly simplified isolate song containing syllables that resemble innate calls. Spectrograms are depicted in the range of 0 to 10 kHz. White horizontal lines represent 500 ms.

in acoustic isolation, such that the young birds did not hear any other bird sounds other than their own vocalizations [Marler, 1970a]. After 8 months of isolation, the birds began to sing, and after 12 months, the birds had developed stereotyped and crystallized song. Although the songs were roughly the same duration as those of naturally reared birds, all of the fine detail of a chaffinch song was lacking from the songs of these isolate birds. The song elements that did appear in the isolate song consisted largely of calls and whistles that developed normally in isolation [Marler, 1970a].

One can similarly compare the effects of isolation in zebra finches. Figure 3.2A shows the final song of a zebra finch that underwent normal song learning: the motifs are highly stereotyped and consist of numerous complex syllables. In contrast, the isolate song of a zebra finch, shown in Figure 3.2B, is a simple song composed of short motifs that contain simple syllables which resemble the innate calls of these birds.

3.3 Effects of Auditory Feedback During Song Learning

Although it is important for the bird to hear conspecific vocalizations early in its life, it is equally important for the bird to be able to hear itself as it starts to vocalize. Songbirds normally exposed to conspecific vocalizations early in life but deafened at different stages of the sensory-motor period develop songs that are correlated with the birds vocal experience prior to being deafened [Nottebohm, 1970]. That is, a bird that is deafened after exposure to the songs of adult birds but before the onset of the sensory-motor

phase develops highly abnormal songs (Fig. 3.3) that lack any evidence of song learning [Konishi, 1965]. Once a bird has begun to sing, however, the juvenile bird no longer needs to hear adult vocalizations in order to develop a normal song, indicating that a representation of the template song has been memorized by the juvenile [Brainard and Doupe, 2002]. The template theory postulates that birds use auditory feedback to compare their developing vocalizations with the template song - the bird hears itself and corrects errors in its vocal output until its song matches the intended pattern of the template song [Konishi, 1985].

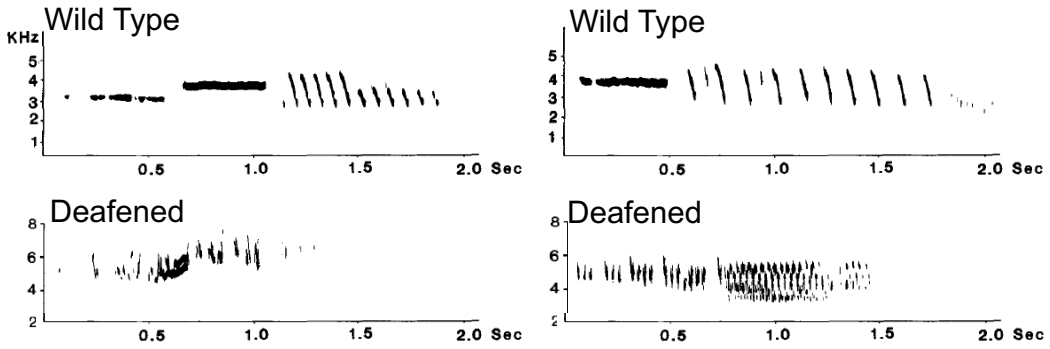

FIGURE 3.3: **The effect of deafening on white crown sparrow songs.** Wild type songs of white crown sparrows (top) from Berkley, California (top left), and Inverness, California (top right). Songs from deafened white crown sparrows (bottom). In comparison to the wild type songs, the songs of deafened birds lack the structure of learned songs, highlighting the necessity for the bird to hear itself as it begins to vocalize. Figure adapted from [Konishi, 1985].

3.4 The Elements of Songbird Vocalizations

Although the crowing of roosters and the cooing of doves may sound like song, the term birdsong is reserved for the songs of passerine birds. Because the various vocalizations of songbirds are often used in experiments to probe the auditory system, it is useful to define the nomenclature that is used to describe songbird vocalizations. Songbirds use two different vocalizations to communicate information: songs and calls.

3.4.1 The Songbird's Song

Songs and calls differ in terms of their structure. Songs tend to be the longest and most elaborate vocalization produced by a bird, lasting anywhere from 2 seconds, as in most species, to tens of seconds [Konishi, 1985]. An example of a zebra finch song is shown in Figure 3.4A. In most species of songbirds, only the male bird sings, and in the wild, songs are typically produced by males during the breeding season. Thankfully, singing

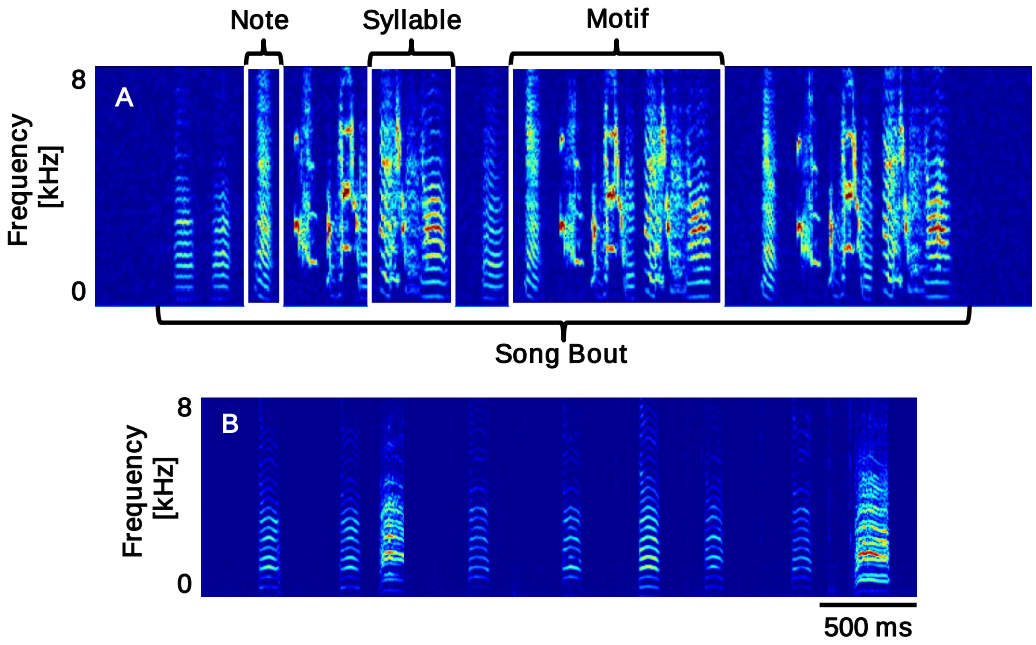

FIGURE 3.4: **The song and calls of a zebra finch. A)** The zebra finch song is a long and complex acoustic signal consisting of notes, syllables, and motifs (see text for explanation). This song bout contains 3 repetitions of the motif, which contains 4 different syllables. **B)** Zebra finch calls, by contrast, are simple vocalizations of shorter duration.

also occurs spontaneously in the laboratory, and birds that are isolated will start and stop singing without any external cues.

Birdsong has several features that make it attractive to study. Most importantly, the songs have a well-defined acoustic structure that is characteristic of the bird species. By breaking down the song into its constituent parts, the song components can be analyzed individually, making study and quantification easier.

The components of a zebra finch song are identified in Figure 3.4A. The most basic acoustic unit in song is called an *element* or *note*. One or more notes may group together to form a *syllable*, which is a stereotyped note sequence that is separated from other syllables by approximately 25 ms of silence. In some cases, when the syllable is very simple, notes and syllables are identical. The syllables are usually repeated in a certain order, and a single syllable can be identified by its fixed position in the temporal structure of the song [Konishi, 1985].

It should be emphasized that the syllables used in birdsong analysis are different from syllables in human speech. In human speech, a syllable usually refers to a single vowel sound and the immediately co-articulated consonant sound, and successive speech syllables need not be separated by a silent interval [Wilbrecht and Nottebohm, 2003]. In

contrast, the syllable in birdsong is a continuous sound preceded and followed by a silent interval and is relatively brief, lasting 50-300 ms.

A series of syllables that is repeated in a predictable sequence is referred to as a *motif*, identified in Figure 3.4A. Motifs may contain identical or dissimilar syllables, depending on the species, but completely random sequences of syllables and notes rarely occur. Some songbirds sing only one motif (e.g., zebra finches), whereas others sing several non-identical motifs (e.g., European starlings).

A series of motifs that occur regularly constitute a song type, and the various song types make up a bird's song repertoire. A repertoire might contain only one song, as for the white-crowned sparrow and the zebra finch, or it can contain hundreds, as for the winter wren (Troglodytes hiemalis; [Konishi, 1985]). A succession of multiple motifs is called a song bout. Depending on the species, a songbird can sing thousands of song bouts in a single day. The timing and sequencing of motifs and syllables is called the syntax of a song, and syntax rules are unique to the species of bird [Konishi, 1985].

3.4.2 The Songbird's Call

In contrast to songs, calls are usually short, monosyllabic utterances of a simple frequency pattern [Marler, 2004]. An example of a train of calls is depicted in Figure 3.4B. Since it is beyond the scope of this chapter to discuss in detail the rich and diverse meaning of bird calls, interested readers are directed to [Marler, 2004]. Unlike songs, both males and females use calls to communicate, and a call can either be innate or learned. For example, zebra finches use distance calls to locate and reunite with each other in the wild after a disturbance of the nest. Interestingly, the distance calls of female zebra finches are innate, whereas the distance calls of males have a learned component [Zann, 1985]. Calls usually occur in response to a specific stimulus, oftentimes the vocalizations of another bird.

4 Methods in Songbird Neurophysiology

4.1 Auditory Stimulus Design

A wide range of stimuli have been used to elicit responses from auditory neurons in the songbird's brain. While some of the first studies of auditory processing in songbirds used simple stimuli, such as click trains [Biederman-Thorson, 1970] or pure tones [Konishi, 1970; Leppelsack, 1974], more recently, scientists have used natural stimuli

such as songbird vocalizations [Margoliash, 1983; Theunissen et al., 2000]) in addition to synthetic stimuli designed to capture specific features of birdsong [Grace et al., 2003; Nagel and Doupe, 2006; Woolley and Casseday, 2005]. The following sections briefly introduce some of the stimuli commonly used in auditory neurophysiology experiments.

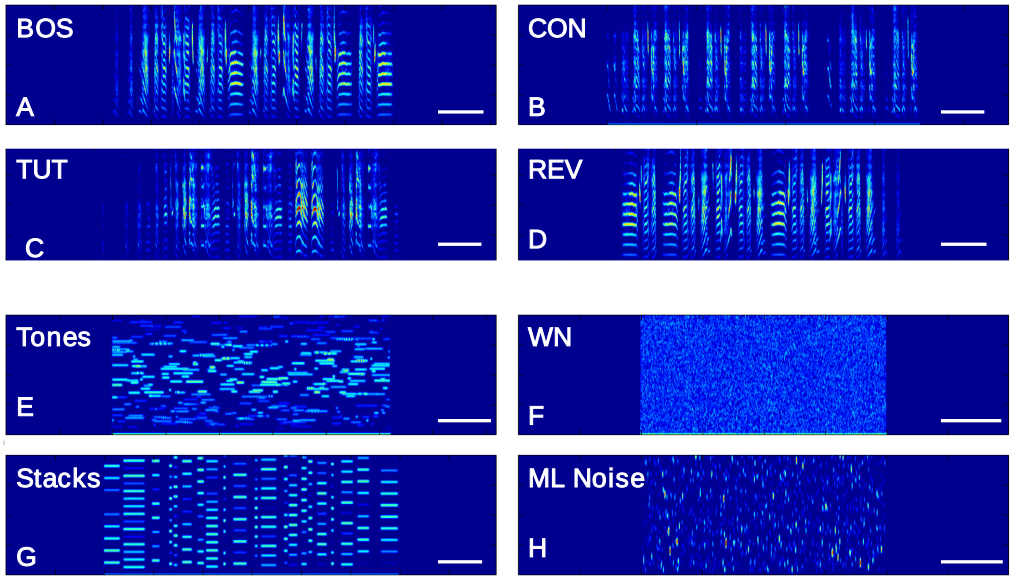

FIGURE 3.5: **Natural and artificial stimuli used in songbird neurophysiology. A-D)** Spectrograms of natural stimuli. **A)** The bird's own song (BOS) is the experimental bird's autogenous song. **B)** Songs from another bird of the same species are also used (conspecific stimuli, CON). **C)** The song of the experimental bird's tutor (TUT) is another commonly used stimulus. **D)** A temporally modified version, such as a reversed (REV) version of the BOS, is another useful stimulus. **E-H)** Spectrograms of synthetic stimuli. **E)** Pure tones can be composed into tone chords. **F)** White noise (WN) is a random synthetic stimulus that contains no correlations in time or frequency. WN is often band pass filtered to match the hearing range of the experimental subject; in this case, it is filtered from 0-8 kHz. **G)** Artificial harmonic stacks are stimuli used to match harmonic syllables common in finch song, such as the natural harmonic stack observed in A. **H)** Modulation-limited (ML) noise is a stimulus that matches the time-varying modulations of frequency and time present in birdsong. All spectrograms are depicted in the range of 0 to 8 kHz. White horizontal lines represent 500 ms.

4.1.1 Natural Stimuli

Perhaps the simplest stimulus to use in a playback experiment is the bird's own song (BOS). An example of a zebra finch BOS is depicted in Figure 3.5A. The BOS can be recorded from the bird on one day and used as a stimulus in a subsequent experiment. Depending on the age and the species of the bird, the BOS can be variable (as in juvenile birds singing subsong or plastic song) or very stereotyped (as in adult birds), varying only in the number of repetitions of the motifs per song bout.

Songs from other birds can also be recorded and used as playback stimuli. These stimuli can be songs from other adult birds of the same species (conspecific songs, CON), and example of which is depicted in Figure 3.5B, or they can be songs of adult birds from different species, known as heterospecific songs. Another useful stimulus is the song of the adult male bird that was used to tutor the experimental bird as a juvenile. The tutor song (TUT), or the song of the male bird who tutored the juvenile may often resemble the BOS, but may also contain slightly different song elements depending on how well the juvenile bird was able to copy the tutor song. An example of TUT is depicted in Figure 3.5C,

Temporal modifications of the BOS are also common stimuli used in experiments probing song-selectivity. A song can be temporally reversed (REV), such that both the sequence of syllables as well as the temporal order within the syllables is altered while the overall power spectrum of the whole song is preserved, as depicted in Figure 3.5D. Similarly, the order of the syllables can be reversed (reverse order, RO-BOS), which disrupts the global sequence of syllables but preserves the local syllable structure [Doupe, 1997].

4.1.2 Synthetic Stimuli

A variety of synthetic stimuli have been used in auditory neurophysiology experiments to ascertain to which song features neurons may be tuned. These stimuli often sound nothing like natural birdsong, but they contain certain statistics that match those of birdsong.

Pure tones are one of the simplest stimuli used in auditory neurophysiology experiments. To be effective, a pure tone stimulus should have a frequency that is within hearing range of the bird. By playing back various tone stimuli, it is possible to determine which frequencies might best drive the auditory neuron. One caveat of this stimulus is that it may not adequately drive higher-order auditory neurons. Pure tones can be organized in chord-like structures, as depicted in Figure 3.5E, to better drive auditory neurons.

White noise (WN) stimuli have been exceptionally useful in characterizing the responses of neurons in the visual system, and they can be used to a similar extent in auditory neurophysiology experiments [Grace et al., 2003]. White noise is a random acoustic signal that has a flat power spectral density and does not contain correlations in time or frequency: the frequency composition of WN at one point in time is unrelated to the sounds preceding it or following it. In neurophysiology experiments, WN stimuli

are usually band-pass filtered to match the frequencies usually present in birdsong; an example is depicted in Figure 3.5F.

Synthetic stimuli can also be created to match specific syllabic structures commonly observed in birdsong, and the harmonic stack, or ripple [Grace et al., 2003], is one such stimulus. Zebra finch songs, as well as the songs of other finches, contain many song syllables that are composed of harmonically related frequency components. The synthetic harmonic stacks try to capture this harmonic structure in order to probe the response of auditory neurons to slow-varying oscillations in the temporal structure of song. An example is presented in Figure 3.5G.

Modulation-limited (ML) noise is used to match both the temporal and spectral modulations that are present in birdsong. Temporal modulations, like those represented by harmonic stacks, are oscillations in power across a frequency spectrum at specific points in time. Spectral modulations, on the other hand, are oscillations in power, or amplitude, of a song over time [Woolley and Casseday, 2005]. An example of ML noise is presented in Figure 3.5H.

4.2 Methods for the Analysis of Neurophysiological Data

Neurophysiologists have recorded neural responses to vocalizations and other complex sounds throughout most of the songbird's auditory system, and a large goal of this research is to quantify the firing responses of neurons in response to auditory stimuli. The following provides a brief summary of the basic quantitative tools used to characterize neural responses in the songbird field. Interested readers should see [Theunissen et al., 2004b] for an in-depth review of these quantitative tools.

Auditory responses can be analyzed using several general methods that quantify the spiking activity in response to a stimulus. The best frequency response of a single neuron describes the loudness (sound pressure level, SPL) required for detection of a pure tone stimulus as a function of frequency. These quantifications are useful for characterizing low-level auditory neurons, but may not be appropriate for higher-functioning neurons which may prefer more complex sound features.

Visualizing the spiking responses of auditory neurons as a raster plot can also be extremely informative in identifying the gross stimulus features to which a neuron responds. Figure 3.6 displays the spiking responses of an auditory forebrain neuron recorded in an urethane-anesthetized zebra finch to WN (upper left spectrogram) and the same neuron's response to BOS (upper right spectrogram), for which spiking responses are aligned to the BOS motif. The middle panels of Figure 3.6 show rasterplots

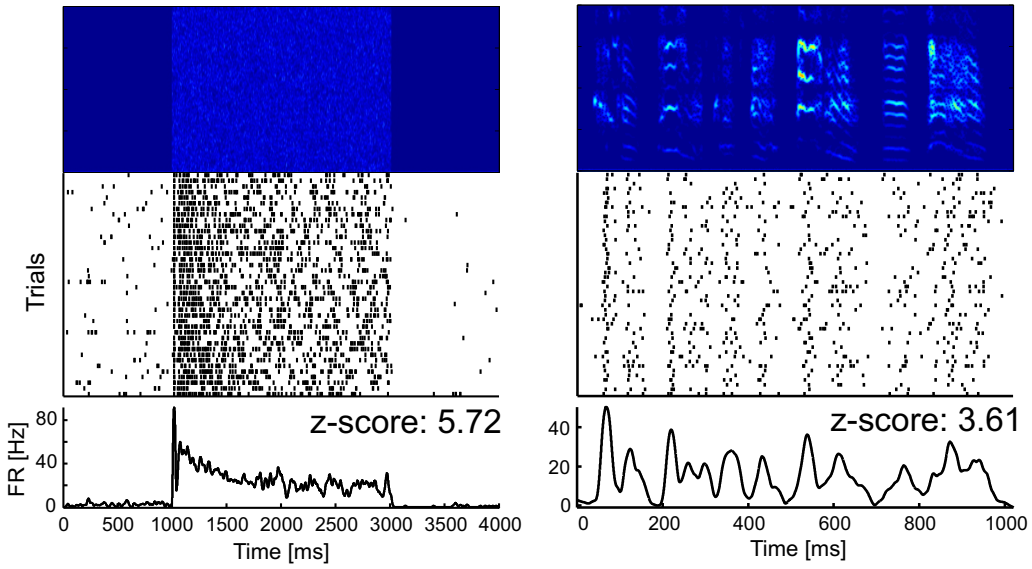

FIGURE 3.6: **Raster plot representation of auditory response dynamics.** Responses of an auditory forebrain neuron to WN (upper left spectrogram) and BOS (upper right spectrogram) stimuli. See text for details. The middle panels depict the spiking responses for each trial of auditory stimulation. Each horizontal row represents the spiking response to one stimulus presentation, and each black line represents a spike. The lower panels displays the neuron's integrated firing response averaged over all stimulus presentations as a function of time. Spectrograms are depicted in the range from 0 to 8 kHz.

of the spiking responses to several identical stimulus repetitions, for which each spike is represented as a black line. The lower panels of Figure 3.6 depict the neuron's integrated firing response averaged over all trials as a function of time.

The auditoriness of a neuron can quantified by the z-score, which is a unit-less, unbounded comparison of the neuron's normalized stimulus-evoked firing rate compared to its normalized baseline firing rate:

$$Z = \frac{(\mu_{Stim} - \mu_{Base})}{\sqrt{(\sigma^2_{Stim} + \sigma^2_{Base} - 2Cov(Stim, Base))}}$$

where μ_{Stim} is the mean stimulus-evoked firing rate, μ_{Base} is the mean baseline firing rate, σ^2 is the variance of the response, and $2Cov(Stim, Spont)$ is the covariance between the stimulus and the response [Theunissen et al., 2004b].

Positive z-scores indicate that the neuron's spiking activity increased during the auditory playback, whereas negative scores indicate that the spiking activity was suppressed by the playback. A z-score of 1 would indicate that the mean firing rate during the stimulus is 1 standard deviation higher than the baseline firing rate. In the case of the neuron in Figure 3.6, the z-score calculated for the neuron in response to WN is 5.72 and for BOS

3.61. The interpretation of these values is that this neuron responds to both WN and BOS playback with increased spiking activity compared to quiet spontaneous firing, and it has a larger auditory response to WN than to BOS. An important caveat in the interpretation of z-scores is to note that neurons with very different spiking response patterns may have nearly identical z-scores, as illustrated in Figure 3.7.

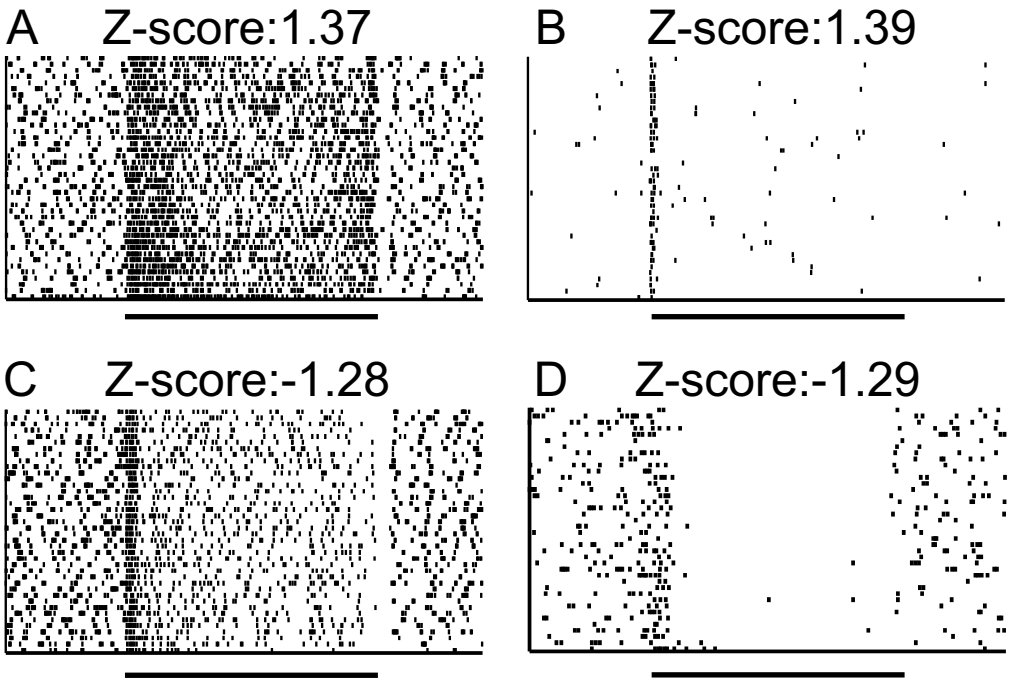

FIGURE 3.7: **Raster plots and corresponding z-scores for four auditory neurons. A-D)** Responses of 4 auditory neurons to 2 s of WN stimulus (black horizontal line). Rasterplots depict the spiking responses of four different neurons, black lines represent spikes. 1 s of spontaneous firing occurring before and after the WN stimulus is also depicted. **A, B)** Both neurons have similar z-scores, despite the obvious differences in their response to WN. The neuron in A has a high spontaneous firing rate and responds robustly to the WN stimulus for the duration of the stimulus. The neuron in B has a much lower spontaneous firing rate and responds only to the onset of the stimulus. **C, D)** Similarly, the neurons in C and D have similar z-scores despite the different responses to WN stimulus. The neuron in C has a high spontaneous firing rate and phasically increases its response to the onset of WN. The neuron in D also has a high firing rate during silence, and its spiking activity is completely suppressed by WN.

When several different auditory stimuli are used in an experiment, it is useful to quantify whether one stimulus is more effective in driving a response than another stimulus. The d-prime score for a two-alternative, forced choice test [Swets, 1961] can be used to quantify the selectivity preference of a neuron between two stimuli A and B. Positive d-prime values indicate that stimulus A elicited a greater response than stimulus B and oppositely if the d-prime value is negative. A d-prime value of zero indicates no difference in the responses evoked by the stimuli. For normal probabilities, a d-prime value

of 1 corresponds to a probability of correct discrimination of 85% [Theunissen et al., 2004b].

Much in the same way that visual neurons can be characterized by calculating their receptive fields, auditory neurons can also be characterized by their spectro-temporal receptive fields (STRFs). The STRF characterizes the spectral and temporal features of a sound to which a neuron best responds. It is beyond the scope of this section to provide an in-depth description of the mathematics behind the STRF estimation, and interested readers are referred to [Theunissen et al., 2001, 2004b] for an extensive explanation. It is, however, worthwhile to describe how STRFs describe the features of sound which are encoded by auditory neurons.

The estimated STRF for the neuron whose responses to WN and BOS are depicted in Figure 3.6 is shown in Figure 3.8A. When an STRF is plotted with time increasing along the x axis, one can achieve an overall sense of frequencies and latencies that have an effect on the neural responses of the auditory neuron. The easiest way to interpret an STRF plotted like this is to imagine that a sound stimulus is presented at time 0, and that this sound has some specific frequency content. For this STRF example, red areas indicated frequencies that would drive neural firing and blue areas indicate frequencies that would suppress neural firing. By interpreting the STRF in Figure 3.8A , one can see that this neuron responds to a fairly broad frequency range. Sounds that that occurred 10-20 ms in the past and that have energy in the frequency range of 1 to 5 kHz tend to drive the neuron to fire, whereas sounds that occurred approximately 20 to 30 ms in the past and that contain energy from 3-4 kHz or 5 to 7 kHz tend to suppress the firing of this neuron.

In addition to providing a visual representation of the spectral and temporal features to which a neuron responds, the STRF can also be used to make predictions of the neural responses to novel stimuli. The success of a STRF in correctly predicting the neural responses to novel stimuli is used to validate the STRF. In Figure 3.8B, a prediction of the neural response by the STRF in Figure 3.8A to the CON stimulus depicted in the upper spectrogram is plotted in red. The measured neural response of the neuron to the CON stimulus is plotted in black. The correlation coefficient (CC) can be calculated between the predicted and real responses in order to estimate how similar the two curves are, where a value of 1 implies perfectly correlated responses and a value of -1 implies perfectly anticorrelated responses. In this case, the CC is 0.37, which indicates that this linear STRF fairly accurately estimates the real neural response of the neuron.

In summary, the STRF is an exceptionally useful tool because it provides a visual representation of the frequency tuning bandwidth of a neuron, the spike latencies for both

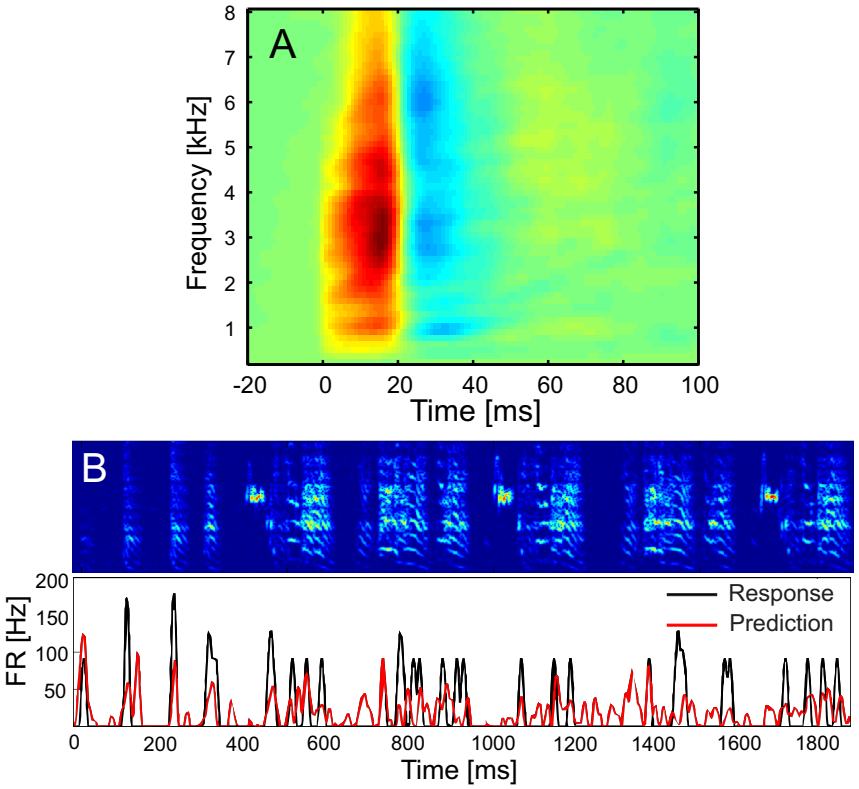

FIGURE 3.8: **STRF and prediction validation of auditory response to CON stimuli. A)** STRF of an auditory neuron, the responses of which are displayed in Figure 3.6. See text for details. Red areas indicate regions of excitation, blue areas indicate areas of suppression. **B)** Predicted response of STRF in A to CON stimuli (upper spectrogram) withheld from the STRF estimation. Black line, real response of the neuron, red line, predicted response generated by convolving the STRF with the stimulus spectrogram. The correlation coefficient can be calculated between the real and predicted responses in order to validate the STRF; for this STRF, the CC value is 0.37.

excitation and suppression as function of frequency, and the temporal and spectral modulations that best drive the neuron [Theunissen et al., 2004b]. The STRF provides an accurate linear representation of the basic underlying computations performed by the cells, and many classical characterizations of auditory responses can be extracted from the STRF. It is important to understand, however, that the STRF is only a linear estimation of the firing response of a neuron. Since many auditory neurons have non-linear responses to auditory stimuli, the picture one obtains with the STRF is not complete, but it nonetheless remains a useful tool in the analysis of neural responses to complex auditory stimuli. A freely available Matlab (Mathworks) toolbox for the estimation of STRFs is available for general use at `http://strfpak.berkeley.edu/`.

5 Auditory Processing in the Songbird

5.1 An Overview of the Anatomy of the Songbird Auditory Pathway

Songbirds are adept at recognizing other conspecific birds solely on the basis of their vocalizations and often in unfavorable acoustic environments. Recent studies investigating the physiology of natural sound processing in the avian auditory system use an impressive array of experimental techniques, including extracellular recordings in anesthetized birds [Grace et al., 2003; Sen et al., 2001], intracellular recordings in anesthetized birds [Lewicki, 1996; Mooney, 2000] extracellular recordings in awake and behaving birds [Keller and Hahnloser, 2009], and functional magnetic resonance imaging [Boumans et al., 2008]. The cumulative results of these studies suggest that the auditory system of songbirds shows a specialization for processing natural sounds and vocalizations.

The avian auditory system is old, and it shares features with all avian groups and other vertebrates, including mammals. The most obvious similarity pertains to the number of neural processing stages and the feed-forward connections from the cochlear nuclei to the auditory forebrain [Theunissen and Shaevitz, 2006]. Afferents from the inner ear project to the cochlear nucleus in the brainstem. Similar to the connectivity in mammals, there is both a direct and an indirect route connecting the cochlear nucleus and the auditory midbrain, as depicted in Figure 3.9.

In the midbrain, these pathways converge in the dorsal lateral nucleus of the mesencephalon (MLd), which is analogous to the inferior colliculus in mammals [Woolley and Casseday, 2004]. The auditory midbrain projects to Ovoidalis (Ov), an auditory nucleus in the thalamus, just as the inferior colliculus projects to the medial geniculate body in mammals. Ov, in turn, sends projections to the primary auditory area in the telencephalon (pallium), called Field L, which is considered analogous to the mammalian auditory cortex [Vates et al., 1996]. Field L can be divided into several subregions on the basis of cytoarchitecture and connectivity [Fortune and Margoliash, 1992]. Input from Ov projects to the input regions of Field L, which in turn project reciprocally to the output regions of Field L. These output regions make reciprocal connections with two secondary auditory areas, the nidopallium caudal medial (NCM) and the caudal mesopallium (CM; [Vates et al., 1996]). Information from the secondary areas projects indirectly to HVC, a unique pre-motor area that is involved in song production and which contains some of the most stimulus-selective neurons ever observed [Margoliash, 1986].

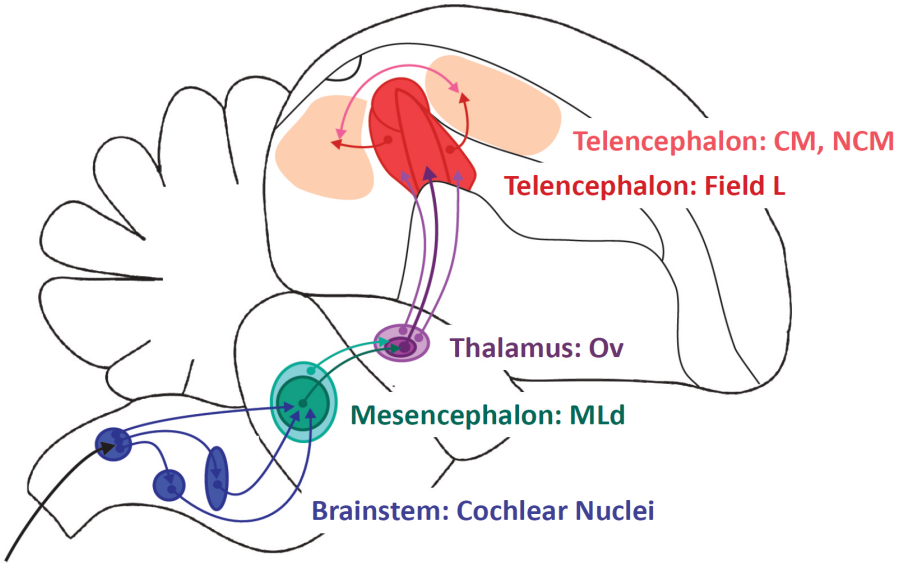

FIGURE 3.9: **Auditory pathway of the songbird.** Schematic illustrates the feed-forward connections of the auditory pathway in the songbird. See text for details. Afferents from the inner ear project to the cochlear nuclei in the brainstem (blue). Information from the cochlear nuclei converges in the auditory mesencephalon, in area MLd (green). Auditory information from MLd travels to the auditory thalamus, Ov (purple), and then to the auditory forebrain area in the telencephalon (red). The auditory forebrain is composed of the Field L complex, as well as the large secondary auditory areas, CM and NCM.

6 The Auditory Midbrain MLd

The auditory midbrain region has only recently gained attention from songbird neurophysiologists. While a large amount of research has been devoted to studying the mechanisms of sound localization in the auditory midbrain of a non-songbird, the barn owl (*Tyto alba*; for a review, see [Knudsen, 1999]), fewer studies have focused on the auditory processing capabilities of these neurons in songbirds.

6.1 Anatomy and Connectivity of MLd

The auditory midbrain region MLd is a brain structure that is conserved across birds and is homologous (derived from a common ancestor) to the inferior colliculus in mammals and the torus semicircularis in amphibians [Woolley and Casseday, 2004]. Anatomically, MLd is located within the intercollicular complex (ICo) and can be differentiated into an inner and an outer region distinct from the surrounding ICo on the basis of calcium-binding protein staining [Logerot et al., 2011].

MLd is the first central processing site of auditory information, and almost all of the brain stem pathways converge there. The brain stem nuclei which project to MLd include the superior olivary nucleus [Wild et al., 2010], the nucleus angularis, and the nucleus laminaris [Krützfeldt et al., 2010]. This convergence of such various brainstem signals makes MLd a likely first site in the pathway of auditory processing, where simple signals that carry information about "what" a signal is and "where" it is coming from are integrated by the MLd neurons. This integrated information from MLd forms the output to the next area of auditory processing, the auditory thalamus Ov.

6.2 Neural Responses of MLd Neurons

In some of the first studies of the neural responses of midbrain neurons in songbirds, extracellular recordings were obtained from MLd neurons in urethane-anesthetized zebra finches in response to pure tones [Woolley and Casseday, 2004], a variety of artificial stimuli [Woolley and Casseday, 2005], and song stimuli [Woolley et al., 2009]. Although urethane anesthesia depresses neural excitability in the songbird auditory midbrain, causing lower spontaneous activity, it does not affect the spectral tuning of these neurons [Schumacher et al., 2011]. The following section discusses some of the important findings regarding the neural responses and processing abilities of MLd neurons.

6.3 MLd Neurons: Onset Detectors that Encode the Rhythm of Song

How do MLd neurons treat the low-level information they receive from the brainstem nuclei? MLd is the site of converging parallel auditory processing streams, and it could potentially integrate simple signals to create more specialized tuning properties.

In terms of tonotopic organization, neurons in MLd are organized in the same way as those found in mammalian inferior colliculus: the deeper, more ventral neurons of MLd have higher best frequency responses, whereas the neurons located in the more superficial regions of MLd tend to have lower best frequency responses [Woolley and Casseday, 2004]. This dorsoventral tonotopy closely matches that found in the mammalian inferior colliculus.

A wide range of frequency tuning across cells is present in MLd neurons, and this frequency tuning in MLd could encode aspects of bird songs and calls. Most MLd neurons have simple, V-shaped frequency tuning curves [Woolley and Casseday, 2004], indicating that MLd neurons respond to specific spectra of neighboring frequencies [Woolley and Casseday, 2004]). Interestingly, a subset of MLd neurons have complex tuning

curves, indicating that even at this lower processing level, a portion of neurons in MLd have complex, non-linear responses to pure tones.

Supporting the idea that MLd responses may be tuned to specific aspects of songs and calls, some MLd neurons respond to specific acoustic features of white noise [Woolley and Casseday, 2005]. The majority of MLd neurons respond robustly to a wide range of synthetic stimuli, indicating that these neurons are insensitive to stimulus class. A subset of neurons does display response properties to white noise, indicating that some MLd neurons could be capable of discriminating between different acoustic elements. These findings suggest that in some MLd neurons, selectivity exists for particular spectro-temporal features, if not for broad classes of stimuli [Woolley and Casseday, 2005].

The temporal responses of MLd neurons to ongoing stimuli reflect the importance of identifying the starts of sounds, and a majority of MLd neurons encode the onsets of stimuli with a temporally-locked bout of spikes. The remaining neurons respond tonically throughout the duration of the stimulus, with some neurons displaying sustained firing throughout the duration of the stimulus and others responding robustly to the onset of the stimulus and then decreasing their firing for the rest of the stimulus duration [Woolley and Casseday, 2004]. Overall these results show that the onset of a sound stimulus is a feature that is encoded across a variety of neuron types in MLd; the pure onset neurons as well as the other tonically firing neurons are equipped to indicate the onsets of sound stimuli by their firing patterns.

STRF estimation using MLd responses to song stimuli confirmed that most neurons in MLd had strong onset characteristics and that these song STRFs were similar across the population of MLd neurons [Woolley et al., 2009]. The majority of the MLd neurons sampled (74 of 110) responded to a wide range of frequencies within a narrow time window. These neurons, which displayed broad spectral tuning, were classified as "broadband" neurons (BB; Fig. 3.10A [Woolley et al., 2009]). In contrast, the second largest group of neurons (25 of 100) responded to only a narrow range of frequencies within a narrow time window and was classified as "narrowband-temporal" neurons (NB-T, Fig. 3.10B). Both the broadband and narrowband-temporal STRFs were characterized by bands of excitation followed in time by bands of inhibition, indicating that the neurons are maximally driven by sound onsets [Woolley et al., 2009]. The broadband neurons would be adept at detecting the onsets of sounds containing a wide range of frequencies, whereas the narrowband-temporal neurons could detect the onsets of sounds with specific, perhaps behaviorally relevant, frequency content.

This preponderance of onset responses suggests that MLd neurons may be particularly good at marking the temporal relationships or rhythm that exist within birdsong. As a population of onset detectors, MLd neurons may facilitate accurate encoding of stimuli

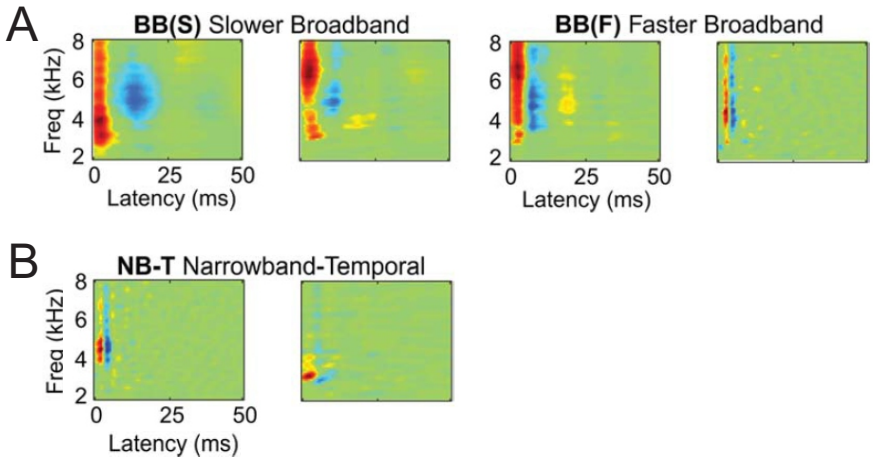

FIGURE 3.10: **Commonly observed STRFs in MLd. A).** Neurons with broadband (BB) tuning represent the largest class of STRFs in MLd. Broadband neurons respond to a wide range of frequency and integrate information over long (Slower) or short (Faster) periods of time. **B)**. Neurons with narrowband-temporal (NB-T) tuning represent the second largest class of STRFs in MLd. These neurons responded to narrow ranges of frequencies within a narrow time window. Figure adapted from [Woolley et al., 2009].

and produces a powerful and temporally precise population response to song [Woolley et al., 2006]. While encoding the rhythm of birdsong is most likely a major function of MLd neurons, their responses may also allow them to encode other features of birdsong. For example, neurons in MLd are for the most part insensitive to the sound intensity level of the stimulus [Woolley and Casseday, 2004], and MLd neurons could function to stabilize the precise encoding of temporally complex signals like birdsong by preserving the meaning of the signal regardless of sound intensity level.

Overall, neurons in MLd respond robustly to pure tones, complex tones, noise, and songs, show a wide range of frequency tuning, an overall lack of sensitivity to intensity changes, and display temporal response patterns that reflect the encoding of the temporal patterns of the stimulus. The precise onset responses suggest a role in temporal processing, such that the unique characteristics of birdsong are preserved in a neural representation of the temporal pattern of the song.

7 The Auditory Nucleus of the Thalamus

The nucleus Ov was classically defined as the avian auditory relay nucleus of the thalamus [Karten, 1968] and is analogous (similar in function but not in evolutionary origin) to the mammalian medial geniculate body. Ov is a tiny, oval shaped nucleus located deep within the avian forebrain. While it has been traditionally thought of as a simple,

feed-forward thalamic relay nucleus, recent studies [Amin et al., 2010; Lei and Mooney, 2010] have challenged this notion, suggesting that Ov also has a role in the way that auditory feedback shapes learned vocalizations by encoding information about the quality of vocal performance.

7.1 Anatomy and Connectivity of Ov

Ov stands out from the surrounding tissue in Nissl stained slices due to its darkly colored and densely packed neurons. Ov can be divided into a core and a surrounding shell, and the shell itself can be divided into several subregions [Zeng et al., 2004]. Because both the core and the shell respond to auditory stimuli [Durand et al., 1992], both regions are likely to be involved in auditory processing, but unfortunately, neurophysiological studies rarely distinguish between Ov shell or Ov core when recording.

Fibers from MLd enter the Ov core rostroventrally [Bigalke-Kunz et al., 1987]. The Ov shell receives input from neurons located on the medial border of MLd and the area surrounding the robust nucleus of the arcopallium (RA cup; [Zeng et al., 2004].

The Ov core projects primarily to the input area L2a of Field L [Vates et al., 1996]. The Ov shell is more widely connected than the Ov core: it connects to regions of the pallium including Field L (L2b, L1 and L3), NCM, as well as several hypothalamic areas such as the nucleus vetromedialis hypothalami and the nucleus anterior medialis hypothalami [Vates et al., 1996; Zeng et al., 2004]. Considering its projections to the hypothalamic areas, Ov shell might serve to integrate auditory input with other sensory systems, affecting neurosecretory and reproductive activities through its hypothalamic targets, and may help to process behaviorally relevant vocal cues [Zeng et al., 2004]. Supporting this idea, met-enkephalin immunoreactive fibers or cells were found in Ov shell but not in Ov core [Zeng et al., 2004]. Met-enkephalin is a neuromodulator that inhibits the release of several classical neurotransmitters, such as acetylcholine. In addition, both the core and the shell regions of Ov contain GABA-positive cells [Pinaud and Mello, 2007].

7.2 Neural Responses of Ov Neurons

Perhaps because Ov is a tiny nucleus, few attempts have been made to characterize the auditory responses of Ov neurons. In possibly the first study of this kind, Bigalke-Kunz et al. [1987] used a range of artificial stimuli to analyze the auditory responses of neurons in head-fixed, unanesthetized European starlings. In a study more than 20 years later, Amin et al. [2010] obtained extracellular recordings from urethane-anesthetized

zebra finches and reported the first STRFs for Ov neurons. In a study using chronically-implanted awake and behaving zebra finches, Lei and Mooney [2010] used white noise to perturb auditory feedback during singing and recorded multi-unit activity in Ov. The following section summarizes this work and discusses the role of Ov neurons in representing complex sounds.

7.3 Nucleus Ovoidalis: More Than a Simple Relay Station

What function does the thalamic auditory nucleus serve in the processing of complex auditory stimuli? Is it simply a nucleus that forwards auditory information from MLd to the next auditory brain area, Field L? Or does it act as more than a simple relay nucleus, restructuring and reorganizing auditory information before sending it onwards? Recent evidence indicates that Ov does indeed act as more than a simple thalamic auditory relay nucleus, suggesting that Ov may play a more substantial role in the processing of auditory information involving the recognition and gating of behaviorally relevant sounds [Amin et al., 2010]. Neurons in Ov are tonotopically organized, but in a gradient opposite of that observed in afferent MLd. In Ov, neurons with higher best frequency responses are located more dorsally and neurons with lower best frequencies are located more ventrally [Bigalke-Kunz et al., 1987].

Like the afferent neurons in MLd, neurons in Ov are insensitive to stimulus class. Bigalke-Kunz and colleagues reported that all of the Ov neurons that they tested responded similarly to all classes of artificial stimuli, including bandpassed noise, pure tones, and frequency and amplitude modulated tones [Bigalke-Kunz et al., 1987]. In contrast to the responses of MLd neurons, which primarily encode the onset of the stimuli, the majority of neurons in Ov respond tonically throughout the duration of the stimulus. Phasic onset response, like those observed from MLd neurons were also reported in Ov, although they were less prevalent than the tonically firing neurons [Bigalke-Kunz et al., 1987]. Interestingly, a small subset of neurons in Ov were suppressed by sound and displayed rebound excitation at the offset of the stimulus [Bigalke-Kunz et al., 1987], responses that have not been reported in afferent MLd neurons.

Using STRF estimation as a tool, Amin and colleagues investigated whether the receptive field features of Ov neurons are more similar to the afferent MLd neurons or the efferent Field L neurons [Amin et al., 2010]. Most of the neurons in Ov (23 of 45) are tuned to either a specific temporal feature of song or to a narrow range of frequencies. STRFs for these "narrowband-temporal" (NB-T; Fig. 3.11A) and "narrowband-spectral" (NB-S; Fig. 3.11B) neurons look similar, except that the NB-S neurons have inhibitory sidebands in the frequency domain as well as the time domain that is, these

neurons are maximally driven by a narrow range of frequencies and inhibited by frequencies neighboring that narrow frequency range. While many neurons in MLd are also reported to have narrow temporal tuning useful for detecting onsets [Woolley et al., 2009], neurons tuned to narrow spectral features, which could be used to recognize sounds with a particular harmonic structure, were not observed in MLd and only first arise in its downstream target Ov.

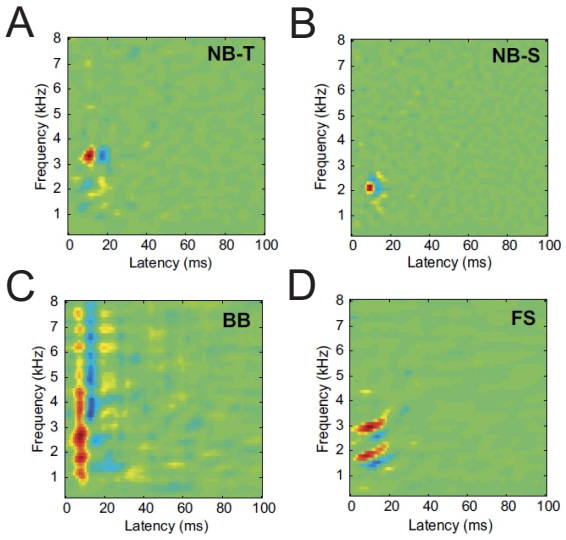

FIGURE 3.11: **Commonly observed STRFs in Ov. A, B)** Neurons with narrowband (NB) tuning represent the largest class of STRFs in Ov. These classes were found to be either temporally sensitive (**A**) or spectrally sensitive (**B**; see text for details). **C)** Broadband responses were also observed in Ov. These neurons are tuned to a wide range of frequencies. **D)** A small group of neurons in Ov were also tuned to frequency sweeps (FS). Figure adapted from [Amin et al., 2010]

Neurons with broadband responses (BB; 10 of 45; Fig. 3.11C) were also present in Ov and resembled broadband neurons found in both MLd and Field L. These neurons have broad frequency tuning and integrate information over short or long time windows. Neurons that integrate information over long time windows could encode features of the stimulus that are important for timbre, such as the amplitude envelope of a stimulus, whereas the faster broadband neurons able to integrate information over short time windows could encode the onset of a stimulus, which is important for detecting the overall rhythm of the song [Amin et al., 2010]. Finally, some neurons in Ov (7 of 45) are selective for frequency sweeps (FS; Fig. 3.11D), sounds that are a common feature of birdsong. These neurons are more common in Ov than in both MLd and Field L. [Amin et al., 2010].

STRFs from Ov neurons suggest that there is a preservation of the type of receptive fields inherited from the MLd, such as the narrowband-temporal and broadband tuning, but that Ov neurons also generate new types of receptive fields, specifically the

narrowband-spectral and frequency sweep tuning [Amin et al., 2010]. Furthermore, although the mean response latency of Ov neurons is 10.2 ms, several neurons were found in Ov with response latencies of around 20 ms or more, suggesting that feedback from Field L may shape these responses. Indeed, the STRFs for Ov neurons with short mean latencies were more similar to MLd STRFs (due to fast feed-forward connections) and the STRFs for neurons with long mean latencies were more similar to Field L STRFs [Amin et al., 2010]. These results suggest that there might be modulated feedback to Ov from Field L.

Experiments using perturbed auditory feedback during singing also support an active role for Ov in the encoding of auditory information about the signing performance. Lei and Mooney [2010] played a white noise perturbation that blocked out a portion of the motif while the bird was singing. This perturbing stimulus significantly increased the multiunit Ov activity in the region of the motif targeted by noise. Furthermore, song-triggered electrical stimulation in Ov gradually decrystallizes targeted regions of the bird's song over time, indicating that singing-related auditory feedback travels though Ov [Lei and Mooney, 2010]. Therefore, Ov plays a substantial role in the processing of auditory information and additionally has a role in the singing behavior of juvenile songbirds through its interactions with song motor networks.

8 The Auditory Cortex of Songbirds: Field L

The term "Field L" was first mentioned by Maximilian Rose in his 1914 study of the avian forebrain [Rose, 1914]. He described Field L as an area of darkly-staining, densely-packed neurons that was present in every bird he examined, including songbirds. A physiological role for Field L was first proposed by Harvey Karten in 1968 [Karten, 1968]. By selectively destroying Ov in pigeons and tracing the resulting anterograde degeneration to Field L, Karten proposed the existence of an auditory pathway leading from the inferior colliculus (area MLd in songbirds) over the thalamic nucleus Ov to the auditory forebrain area Field L. Nearly a hundred years later, Rose's Field L is now considered to be one component of a larger auditory processing complex that is considered analogous (similar in function but not in evolutionary origin) to the mammalian auditory cortex.

8.1 Anatomy and Connectivity of Field L

Field L is a large complex of neurons in the middle of the telencephalon (nidopallium). Based on the cytoarchitectural organization, connectivity, and morphology of

cells, Field L can be divided into several subregions: L1, L2a, L2b, and L3 [Fortune and Margoliash, 1992].

Field L receives input from both the core and shell of the thalamic nucleus Ov. L2a receives projections from the Ov core, and L2b receives input from a thalamic nucleus adjacent to Ov, the nucleus ovoidalis medialis (see Figure 3.12; [Vates et al., 1996]. Areas L1 and L3 also receive weak input from the Ov shell. Overall these results suggest that L2a and L2b constitute the input region of Field L, whereas L1 and L3 can be thought of as the output areas of Field L. After auditory information reaches Field L, it is distributed via a branching network of projections to brain areas including secondary auditory areas NCM and CM, the avian striatum, and the shelf area of HVC [Kelley and Nottebohm, 1979].

The subregions of Field L are densely interconnected and form a web of connections that process incoming auditory information. Projections from L2a form a dense fiber network with neighboring regions L2b, L1, and L3. L2a strongly projects to NCM as well as to striatal targets such as the globus pallidus, and the lateral striatum [Mitchell and Hall, 1984; Reiner et al., 2004; Vates et al., 1996]. Interestingly, there seems to be no connection from L2a to the immediately adjacent song control area, nucleus interface of the nidopallium (NIf; [Vates et al., 1996]).

Projections from L2b also form reciprocal connections to neighboring areas within Field L, including L1, L2a, and L3. The most prominent projection from L2b is a reciprocal connection to the lateral part of the caudal mesopallium (CLM; see Figure 3.12). Like L2a, L2b also targets the avian lateral striatum, but unlike L2a, it does not reach into avian globus pallidus [Vates et al., 1996].

The output regions of Field L, L1 and L3, also project to several areas in the avian brain. L1 projects to several nuclei important for song control. The most prominent projection from L1 is to an area just beneath HVC known as the HVC shelf region (see Figure 3.12;[Vates et al., 1996]. A small but direct pathway may also exist between L1 and HVC proper [Shaevitz and Theunissen, 2007]. Additionally, L1 also projects to the region just anterior to the robust nucleus of the arcopallium (RA) known as the RA cup. Like L2a, L1 also projects to secondary auditory areas CLM and NCM, the avian striatum, and shares a reciprocal connection to L3 [Vates et al., 1996]. L3 is reciprocally connected to L1, L2a, L2b, and projects to both CLM, and HVC shelf (see Figure 3.12; [Vates et al., 1996]. It projects to NCM much more robustly than does L1. Retrograde tracer injections into HVC suggest that a weak connection between L3 and HVC proper may also exist [Fortune and Margoliash, 1995].

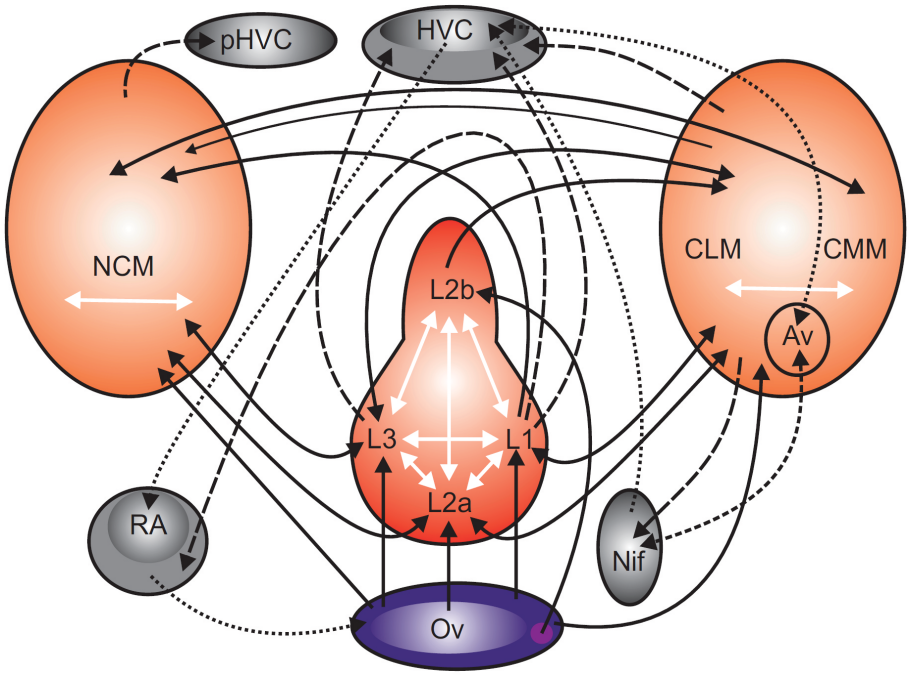

FIGURE 3.12: **Connections between Ov, the auditory forebrain, and song control nuclei.** See text for details. White arrows indicate reciprocal connections within an auditory area. Solid black arrows indicate projections from one auditory area to another auditory area. Dashed arrows indicate connections from auditory areas to song control nuclei. Dotted lines indicate projections from song control nuclei. Ov core (light purple), shell (dark purple), and nucleus ovoidalis medialis (light purple circle) send projections to the Field L complex, NCM, and CM. Field L is highly interconnected (see text for details), and sends output to NCM and CM as well as to HVC shelf. NCM projects to para HVC (pHVC) and CM projects to song control nuclei HVC and HVC shelf, and NIf. Field L projections to striatal targets are not indicated, nor are Ov projections to hypothalamic nuclei.

Overall, anatomy studies indicate that L2a and L2b act as the principal point of access for auditory information entering the nidopallium from the Ov core. L2a and L2b then relay this information to L1, L3, CLM, and NCM in a dense web of interconnecting reciprocal connections.

8.2 Neural Responses of Field L Neurons

Perhaps because its size and darkly stained neurons make it easy to locate, neurophysiologists have studied the auditory responses of neurons in the avian Field L complex for more than 40 years in many different birds, both songbirds and non-songbirds alike. While early studies focused on characterizing the auditory responses to synthetic stimuli such as noise and tones [Biederman-Thorson, 1970; Leppelsack, 1974], it quickly became clear that neurons in Field L are also well-driven by natural auditory stimuli such as birdsong [Leppelsack and Vogt, 1976; Scheich et al., 1979]. An abundance of

research examined the auditory responses of neurons in Field L to synthetic and natural stimuli in the hopes of revealing which acoustic features drive these neurons. The following sections will highlight some of the important response properties of neurons in the Field L complex. Overall, results suggest that neurons in Field L act as "feature extractors" that are spatially segregated and tuned for specific features common in birdsong.

8.3 Field L Neurons: Spectro-Temporal Feature Detectors

Extensive electrophysiological mapping of Field L revealed that it, like most auditory nuclei, is tonotopically organized. The caudal areas of Field L display more robust responses to lower frequencies, whereas the rostral areas of Field L are characterized by responses to higher frequencies [Müller and Leppelsack, 1985; Zaretsky and Konishi, 1976].

A long series of electrophysiological studies in various laboratories gradually revealed that neurons in Field L respond robustly and preferentially to complex auditory stimuli such as conspecific birdsong vocalizations rather than pure tones and other simple stimuli. Early experiments using pure tones failed to drive Field L neurons in a non-songbird, the ring dove (*Streptopelia risoria*), although neurons did respond to more "complicated" sounds, such as "trains of clicks, hisses, squeaks, and rattling of keys" [Biederman-Thorson, 1970]. In perhaps one of the first studies to characterize the auditory responses of Field L neurons in songbirds, Leppelsack [1974] used European starlings to examine auditory responses to complex noise and tone playback. The majority of neurons (50 of 103) responded with a strong onset response followed by tonic firing for the duration of the stimulus and suppression of firing marking the offset of the stimulus [Leppelsack, 1974].

Unlike the responses reported for MLd and Ov neurons, many responses of Field L neurons (27 of 103) were characterized by periods of phasic suppression to sound [Leppelsack, 1974]. A large group of these neurons (21 of 27) were tonically suppressed by the noise stimuli and only increased their firing at the offset of the stimulus [Leppelsack, 1974]. The third most common response in Field L (18 of 103) was tonic activation throughout the duration of the stimulus, followed by offset suppression.

The phasic suppression that encodes the onsets and offset of stimuli in Field L neurons is in stark contrast to the responses of MLd neurons, which primarily encode the onsets of noise and tone stimuli with increases in spiking activity. While a few neurons in Ov show phasic suppression in response to auditory stimuli, most neurons respond tonically to a wide range of stimuli. Field L neurons, on the other hand, seem to be equipped to

encode both the onset and offsets of complex noise stimuli with increases and decreases in firing activity, suggesting that inhibitory networks might have a significant role in modulating the responses of neurons to auditory stimuli in Field L. Indeed, high densities of GABA-positive cells are present in L1, L2a, and L3 [Pinaud and Mello, 2007] and may provide the source of this inhibitory modulation.

In the mid eighties, researchers began to wonder whether Field L neurons might also have the same BOS-selective tuning that had recently been discovered in HVC neurons. It had been discovered that while some Field L neurons responded to a wide variety of stimuli, others seemed to only respond to specific features of natural vocalizations [Leppelsack and Vogt, 1976; Scheich et al., 1979], suggesting that some neurons in Field L are tuned to detect species-specific features of birdsong [Amin et al., 2007].

Experiments comparing the responses of Field L neurons to behaviorally relevant stimuli including BOS, CON, and TUT revealed a defining characteristic of neuronal responses in Field L. Unlike higher-order HVC neurons, which show an exquisite tuning preference to BOS playback, Field L neurons showed no preference for BOS or TUT compared to CON stimuli [Margoliash, 1983, 1986]. In addition, neurons in Field L showed much less sensitivity to temporal manipulation of the BOS stimuli than did neurons in HVC, and the responded equally robustly to BOS, REV, and reversed order BOS [Lewicki and Arthur, 1996]. Indeed, neurons throughout Field L prefer CON over a variety of synthetic stimuli designed to match the acoustic features of song, as measured by d-prime scores [Amin et al., 2004], and these CON-selective neurons were found in greater number outside of the thalamorecipient subregion L2 [Grace et al., 2003]. Overall, these results suggest that the neurons in the auditory forebrain are able to process a large ensemble of sounds, but are selective for particular spectro-temporal patterns commonly found in conspecific song [Theunissen and Shaevitz, 2006].

In addition to analyzing the auditory responses of Field L neurons with quantitative measure such as z-scores and d-prime scores, many experimental groups have used STRF estimation as a tool to assess the features of birdsong to which Field L neurons best respond. It seems that as auditory information passes from MLd to Ov and then finally to Field L, both the preservation of tuning and the generation of novel, more complex tuning occur. Using conspecific stimuli to estimate STRFs, Woolley and colleagues [Woolley et al., 2009] found that almost all types of STRFs found in afferent Ov were also found in Field L, including BB (41 of 137; Fig. 3.13A) and the NB-T and NB-S responses (35 of 137; Fig. 3.13B; [Woolley et al., 2009]).

In addition to inheriting or regenerating tuning responses from Ov, two novel groups of STRFs were found from the responses of Field L neurons. One group (7 of 137) had broadband offset responses (Off, Fig. 3.13C), where the excitatory band followed the

inhibitory band in time. Another hybrid group of neurons (Hy; 10 of 137; Fig. 3.13D) showed responses that were a combination of the NB-T and BB responses and included the frequency sweep neurons found in Ov [Woolley et al., 2009]. These findings highlight the fact that in Field L, there is either preservation or regeneration of the feature detectors present in Ov in addition to the generation of novel tuning responses.

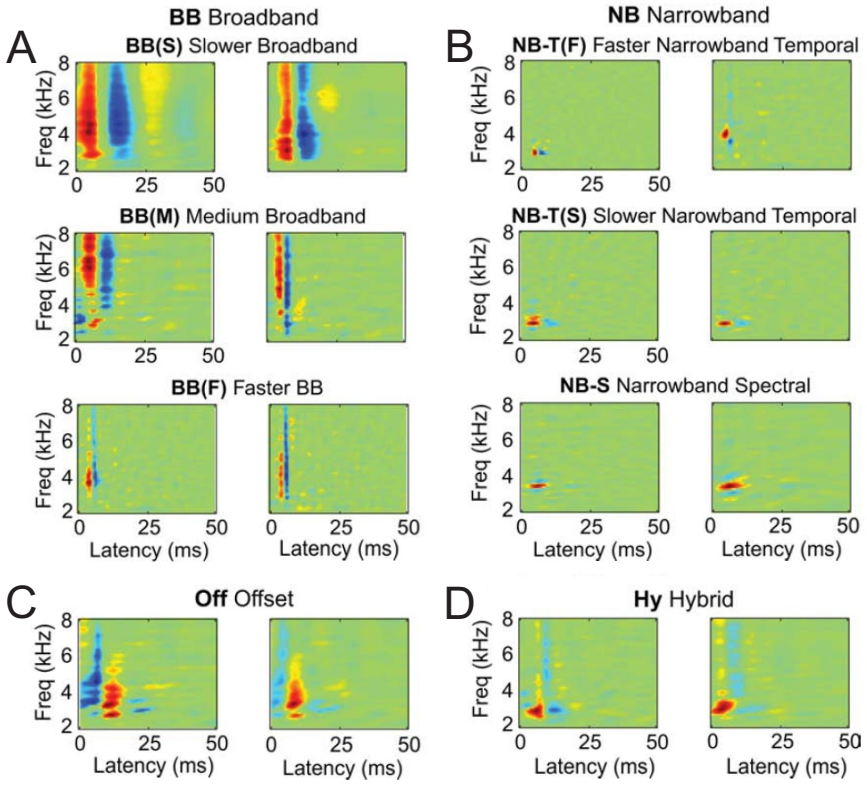

FIGURE 3.13: **Commonly observed STRFs in Field L. A)** Neurons with broadband (BB) tuning represent the largest class of STRFs in Field L. These neurons were tuned to a wide range of frequencies and were capable of integrating information over a range of latencies, including large (slow), medium, and short (fast) time windows. **B)** Neurons with narrowband (NB) tuning represent the second largest class of STRFs found in Field L. These neurons were temporally tuned (NB-T) to narrow ranges of frequencies. Like the BB neurons, these NB-T neurons could integrate information over a range of latencies, including large (NB-T slow) and short (NB-T fast) time windows. An additional class of NB neurons, the NB-S neurons, had inhibitory sidebands in the frequency ranges neighboring the best frequency tuning. **C)** A small group of neurons in Field L had broadband offset (Off) responses. **D)** Neurons in Field L were also tuned to hybrid (Hy) combinations of narrowband-temporal and broadband responses. Figure adapted from [Woolley et al., 2009].

In a parallel STRF estimation study, Nagel and Doupe [2006, 2008] used a synthetic stimulus (amplitude-modulated noise) instead of conspecific songs to estimate STRFs in Field L neurons. Amplitude-modulated noise captures the temporal frequency and amplitude distribution statistics of natural birdsong. Although the authors classified the resulting STRFs more generally than Woolley et al. [2009], Nagel and Doupe [2008]

found three repeating patterns of STRFs: those responsive to 1) temporal features, 2) spectral features, or 3) temporal and spectral features. The group of neurons responsive to the temporal features of song overlaps well with the NB-T and BB neurons STRF groups reported by [Woolley et al., 2009]. The second group of neurons, those that are sensitive to the spectral features of song, corresponds well with the NB-S neurons.

Analysis of the anatomical distribution of the different STRF types suggested that a form of spatially segregated processing may be present in Field L. STRFs with fast temporal response properties are localized to the input area L2, whereas STRFs with slower response properties are more prevalent in output regions L1 and L3 [Kim and Doupe, 2011; Nagel et al., 2011]. Furthermore, a mediolateral broadening of spectral tuning is also present across the subregions of Field L, suggesting that parallel channels with different degrees of spectral integration may exist in Field L [Kim and Doupe, 2011]. Altogether, these results suggest that fast and slow neuronal responses may arise from neurons with distinct morphological and/or electrophysiological properties [Nagel and Doupe, 2008], and the processing of auditory information related to rhythm, pitch, and timbre could occur in a spatially organized and segregated manner in Field L [Kim and Doupe, 2011].

Nagel and Doupe [Nagel and Doupe, 2006, 2008] characterized the sensitivity of Field L neurons to sound intensity using amplitude-modulated noise played back at 2 different sound levels to estimate STRFs. The quiet sound level of 30 dB was similar to a whisper at a quiet library, whereas the louder sound level of 63 dB was similar to the loudness of normal conversation. The temporal responses patterns of the STRFs of most Field L neurons changed systematically with the stimulus intensity [Nagel and Doupe, 2006]. At low sound intensities, STRFs were characterized by mostly positive, excitatory fields, indicating that Field L neurons acted as "integrators" of sound over time by responding whenever the stimulus amplitude was high. At higher sound intensities, the negative, inhibitory regions of the STRFs became larger, indicating that the neurons were responding more selectively to amplitude changes and acting more as stimulus "differentiators" rather than integrators [Nagel and Doupe, 2008].

These finding suggest that intensity-sensitive Field L neurons receive input from at least two neurons - one that is excitatory and one that is inhibitory - and that this input interacts non-linearly to give rise to the intensity-dependent receptive fields observed in Field L [Nagel and Doupe, 2008]. Such non-linear response properties could improve the ability of Field L neurons to effectively encode different stimulus contexts such as birdsong sung by distant birds versus birdsong sung by a nearby tutor [Nagel and Doupe, 2006].

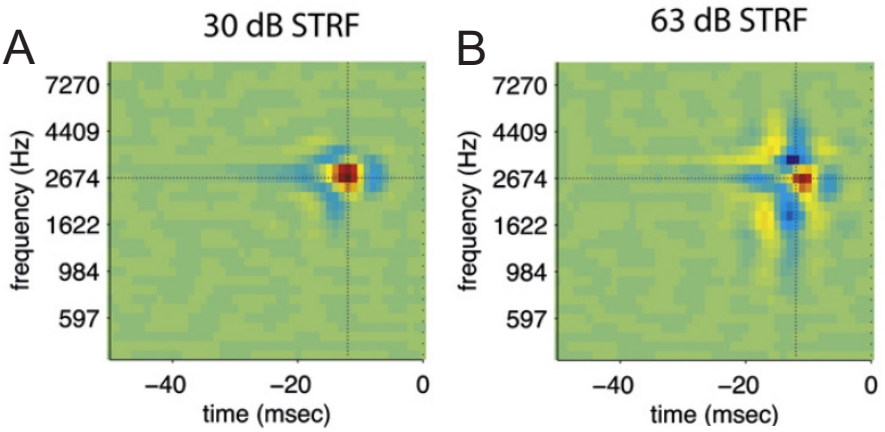

FIGURE 3.14: **STRFs from an intensity-sensitive neuron in Field L. A, B)** Two STRFs calculated for the same neuron in response to stimulus played back at two different sound intensity levels. **A)** At low sound intensities (30 dB) STRFs had mostly positive, excitatory fields (red bands). **B)** At higher sound intensities (63 dB), STRFs were characterized by large and negative inhibitory sidebands (blue bands). Figure adapted from [Nagel and Doupe, 2008].

In summary, it seems that the auditory information from Ov is received by neurons in Field L and is passed via densely interconnected parallel networks to specific subregions of Field L, where spatially segregated processing preserves and generates tuning for features necessary for the recognition of species-specific sounds. These networks most likely involve inhibitory interneurons, allowing Field L neurons to encode sound features with both suppression of firing activity in addition to increases in firing. Overall, neurons in Field L are equipped to encode for information pertaining to rhythm, timbre, intensity, and pitch of birdsong.

9 Secondary Auditory Area NCM

NCM is a large brain area in the nidopallium, and neurons there show selectivity for complex sounds [Theunissen and Shaevitz, 2006]. Molecular gene expression studies of an immediate-early gene encoding a transcriptional regulator (ZENK) involved in memory consolidation implicate NCM as a potential storage site for the neural representation of the tutor song [Hahnloser and Kotowicz, 2010]. Many neuron in NCM show a rapid up-regulation of ZENK expression in response to conspecific songs, but not to heterospecific songs or tone bursts [Mello et al., 1992], and ZENK expression to tutor song in NCM is significantly and positively correlated with the bird's ability to imitate the tutor song [Bolhuis et al., 2000]. Overall, these results suggest that molecularly, neurons in NCM are sensitive to the sensory experiences of songbirds and may have a

role in detecting novel, behaviorally relevant song features [Hahnloser and Kotowicz, 2010].

9.1 Anatomy and Connectivity of NCM

Anatomically, NCM is a large area in the nidopallium that surrounds Field L, and it is bounded dorsally and medially by the ventricular zone and overlying hippocampus [Maney and Pinaud, 2010]. Studies suggest that the dorsal and ventral regions of NCM may differ in functionality [Thompson and Gentner, 2010], although these areas of NCM do not appear different in terms of their cytoarchitectural organization or cell morphology [Maney and Pinaud, 2010]. NCM receives input from the Ov shell, L2a and L3 (see Figure 3.9; [Vates et al., 1996]). In addition, NCM receives a strong reciprocal connection from the medial region of CM (CMM) in addition to a weaker connection from the lateral region of CM (CLM; [Vates et al., 1996]).

Dorsomedial regions of NCM project reciprocally to an area adjacent to HCV known as para-HVC, and these portions of para-HVC that overlap with NCM could serve as an anatomical link between NCM and HVC [Foster and Bottjer, 1998]. NCM is richly interconnected with itself and sends a strong reciprocal projection to CMM [Vates et al., 1996]).

High densities of GABA-positive cells are prevalent in NCM [Pinaud and Mello, 2007]. Recent studies indicate that GABA-A receptors may have a crucial role in auditory processing in NCM, since the expression of a GABA-A receptor subunit increases in NCM after exposure to 30 minutes of conspecific song [Jeong et al., 2011]. Furthermore, significant populations of aromatase-positive cells are exclusively located in the caudal portions of NCM, contrasting with MLd, Ov, Field L and CMM, which are all devoid of estrogen-sensitive or aromatase-positive labeling. These results suggest that estradiol-mediated regulation of auditory input is present in NCM [Maney and Pinaud, 2010], and furthermore, that these estradiol-generated responses in NCM involve GABAergic neurotransmission [Tremere et al., 2009].

9.2 Neural Responses of NCM Neurons

Although the study of basic neurophysiological responses of NCM neurons to auditory stimuli has not been conducted with the same rigor as it has been conducted in other auditory brain areas, some studies have investigated the fundamental auditory response properties of neurons in NCM. Chew and colleagues used awake, restrained zebra finches to record multiunit activity in response to conspecific and heterospecific

songs [Chew et al., 1995], bird calls, human speech, and tone sequences [Chew et al., 1996]. Terleph et al. [2006, 2007] used pure tones and synthetic whistles to drive auditory neurons in awake and restrained male and female canaries while they recorded multiunit activity in NCM. Results from these studies suggest that neurons in NCM have responses that are more complex than neurons in afferent Field L and may be involved in the neural encoding of vocal recognition.

9.3 NCM Neurons: A Role in Vocal Recognition

NCM has a clear tonotopic organization, with higher frequencies represented more ventrally, matching the tonotopy observed in the adjacent Field L [Terleph et al., 2006]. NCM neurons respond less robustly to pure tones than do neurons in L2, and a typical NCM response to tone stimuli is characterized by an initial burst of phasic onset activity that is lower than that observed in L2 neurons, followed by lower tonic firing for the duration of the stimulus [Terleph et al., 2006]. Frequency tuning curves of NCM neurons are wider than those observed in L2 and often contained multiple excitatory peaks, suggesting that information from several sites in L2 may converge onto a single site in NCM, although input from other auditory areas like CM cannot be ruled out [Terleph et al., 2006].

Neurons in caudal NCM respond strongly to conspecific vocalizations and other natural complex stimuli such as human speech [Chew et al., 1996]. Some neurons display phasic bursting to specific elements of conspecific song [Chew et al., 1995]. Unfamiliar heterospecific songs and artificial stimuli produce significantly weaker responses than conspecific song. Responses to BOS are similar in strength to CON, and REV produces responses that are almost identical to BOS [Chew et al., 1996].

There are species-specific differences in the tuning properties of NCM neurons. In zebra finches, NCM neurons respond with broader, more sustained responses to tones than do NCM neurons in canaries. This is potentially an effect of the selectivity of these neurons to the species-specific vocalizations, since zebra finches typically have songs with a broader harmonic structure than those of canaries [Terleph et al., 2007].

Neurons in caudal NCM show marked habituation to repeated tones and complex stimuli. Initially, NCM neurons respond robustly to novel heterospecific and conspecific songs, but the amplitude of the responses rapidly habituates to 40 percent of the initial response [Chew et al., 1995]. Furthermore, habituation to the stimulus types occurs independently for each type. For example, habituation to BOS and REV stimuli occurs at different rates, suggesting that REV stimuli, although they match the BOS stimuli

in terms of its spectral features, are treated by NCM as a different stimulus than BOS [Chew et al., 1996].

In addition, the duration of habituation also differs for each stimulus. Habituation to CON persisted for 38 hours after onset, whereas habituation to heterospecific stimuli was reset as early as 4 hours after exposure [Chew et al., 1995]. The habituation rate was not influenced by gender and appears to document a process of "forgetting" by NCM neurons, suggesting that conspecific songs were quite well "remembered" for 20 h, whereas the other types of stimuli were "forgotten" as early as 4 hours [Chew et al., 1995]. These results suggest that NCM is very sensitive to specific acoustic features of birdsong vocalizations and most likely has a role in the recognition of conspecific songs.

Using a clever experimental paradigm, Thompson and Gentner [2010] showed that neurons in NCM display a learning-induced decrease in firing response to familiar conspecific stimuli. European starlings were trained to recognize 2-3 conspecific songs in a Go/No go operant procedure. After the birds had learned to correctly discriminate the songs, extracellular recordings were made from neurons in NCM during playback of the familiar learned songs and unfamiliar novel songs. Neurons in the ventral region of NCM responded more strongly to the unfamiliar, novel songs than to the familiar learned songs. Furthermore, the neural responses to learned songs were significantly weaker than the responses to neutral passive songs that were played during the training, indicating that the decreased response of NCM neurons to familiar stimuli is not merely the result of song exposure.

These NCM habituation responses may be an effect of learning, as neurons in ventral NCM fire less in response to familiar learned songs when compared to novel, unfamiliar songs. This sparse and selective firing to familiar stimuli may be the result of efficient coding strategies that minimize firing responses to such stimuli [Blättler and Hahnloser, 2011]. It is likely that neurons in NCM have a behaviorally relevant function beyond simple auditory processing; specifically, the increased response to unfamiliar songs in NCM may provide a mechanism by which novel information is integrated into the auditory system [Thompson and Gentner, 2010].

10 Secondary Auditory Area CM

CM is a secondary auditory area that, like NCM, is responsive to features of birdsong vocalizations [Theunissen et al., 2004a]. Neurons in CM and Field L both respond similarly to birdsong [Grace et al., 2003], although neurons in CM are more sensitive

than Field L neurons to the natural order in time and frequency of the acoustic features of birdsong [Hsu et al., 2004]. Like NCM, CM has been implicated in sensory experience-dependent associative learning, whereby CM neurons respond robustly to songs that the bird has learned to recognize [Knudsen and Gentner, 2010]. These results suggest that CM neurons may encode behaviorally relevant features of birdsong.

10.1 Anatomy and Connectivity

CM encompasses a large portion of the anterodorsal part of the brain. CM receives most of its input from the underlying brain areas, and the lateral and medial parts of CM each receive distinctly different input. CLM receives reciprocal input from the all regions of the underlying Field L complex (see Figure 3.9), including L2a, L2b, L1, and L3. Of these subregions, CLM projects most strongly back to L2b, more so than to L1, L2a, or L3. In contrast the medial portion of CM, CMM, receives its input from NCM rather than Field L. Additionally, CM is richly interconnected with itself [Vates et al., 1996].

In addition to the input it receives from primary and secondary auditory areas, CM receives innervations from other brain regions. A weak projection exists from Ov shell to CM [Vates et al., 1996]. Furthermore, a restricted area within the caudal ventral mesopallium, known as the avalanche nucleus (Av; [Akutagawa and Konishi, 2010]) has a reciprocal connection with HVC. Av was first identified by Nottebohm and colleagues in canaries [Nottebohm et al., 1982] and has since been implicated in the song control system [Jarvis, 2007; Jarvis and Nottebohm, 1997]. Since neurons in Av have been reported to display BOS-selective responses [Akutagawa and Konishi, 2010], this nucleus may constitute a pathway for auditory information to reach HVC.

CLM projects to HVC shelf, overlapping with projections from L3 and L1. Dorsal and ventral regions of CM may also project directly to HVC [Bauer et al., 2008]. CLM has a strong projection to NIf, which itself projects directly to HVC, and as well as to the RA cup region [Vates et al., 1996]. As such, strong anatomical evidence implicates CM as a potential source of auditory input to HVC.

10.2 Neural Responses in CM

Neurons in CM are considered higher-order auditory neurons, and as such, it is difficult to assess the basic tuning functions of these neurons. Although CM responses have not been probed to the same extent as other auditory areas in regards to spectral-temporal tuning, several studies have made use of the birdsong as a stimulus to investigate coding in CM of behaviorally-relevant auditory information. Grace et al. [2003] and Amin et al.

[2004] measured multiunit activity in urethane-anesthetized zebra finches in response to synthetic and natural stimuli. Bauer et al. [2008] used chronically implanted, awake behaving zebra finches to record multiunit activity in CM during song playback. Several experiments used European starlings to assess behaviorally relevant activity in CM after an operant discrimination training paradigm, including [Gentner and Margoliash, 2003; Jeanne et al., 2011; Meliza et al., 2010].

10.3 CM Neurons: A Role in Auditory Object Recognition

CM neurons prefer conspecific stimuli over synthetic stimuli such as tone pips, white noise, and artificial stacks [Grace et al., 2003]. Although the response strength to natural sounds is similar for neurons in Field L and CM, recording sites in CM show weaker auditory responses (measured by mean z-scores) than Field L neurons to both synthetic and natural stimuli, [Grace et al., 2003]. Neurons in CM also show a small preference for BOS versus REV, RO-BOS, and CON, as measured by d-prime score [Amin et al., 2004]. Selectivity for BOS was also observed by Bauer and colleagues [Bauer et al., 2008], who reported that 70% of multi-unit recording sites were significantly excited by BOS. These results suggest that a significant fraction of CM neurons are strongly BOS-selective, especially the population of neurons that are excited by song stimuli [Bauer et al., 2008]. Considering the fact that areas of CM project to HVC [Bauer et al., 2008], this BOS selectivity may be a source of auditory input that shapes the BOS selectivity observed in HVC.

Gentner and Margoliash trained European starlings to accurately discriminate between different conspecific songs [Gentner and Margoliash, 2003]. After the training period, the bird was anesthetized and single unit recordings were made in CMM during playback of the familiar songs heard during training and of novel unfamiliar songs. As a population, neurons in CMM responded much more strongly to familiar songs heard during training compared to unfamiliar, novel songs [Gentner and Margoliash, 2003]. Roughly 64% of CMM neurons sampled were selective for one of the conspecific stimuli, and of these neurons, 93% preferred one of the familiar training songs [Gentner, 2004]. These song-selective neurons tend to have lower firing rates than the non-selective neurons and display phasic responses to specific acoustic elements within a motif and suppression of activity for all other motifs. Additionally,the song-selective neurons tend to respond to overall fewer motifs than the non-song-selective neurons [Gentner, 2004]. When compared to responses in neighboring CLM, CMM neurons respond more selectively to motifs than do neurons in CLM, song-evoked neural activity was more variable in CMM neurons than in CLM neurons, and neurons in CMM were able to encode more information about motif identity than did neurons in CLM [Jeanne et al., 2011].

The "song selectivity" response of CMM neurons is the result of an increased phasic response to specific acoustic features that only appear in a small number of motifs, i.e., it is the result of selective tuning for spectro-temporal features centered at the level of the motif [Gentner, 2004]. These data suggest that a subpopulation of CMM neurons are shaped by the bird's previous experience with conspecific songs, and may represent a neural correlate of learned object recognition [Gentner and Margoliash, 2003]. Since behavioral recognition of individual songs is driven in part by the acoustics of the motif, the experience-dependent neuronal selectivity displayed by the CMM neurons may contribute substantially to individual song recognition behavior in starlings [Gentner, 2004].

Responses of CMM neurons are qualitatively different from the responses of the reciprocally connected ventral NCM. Like in NCM, neurons in CMM respond to a few specific motifs within the training set, but unlike neurons in ventral NCM, which respond strongly to novel unfamiliar stimuli [Thompson and Gentner, 2010], unfamiliar songs evoke very weak responses in CMM neurons [Gentner and Margoliash, 2003]. Similarly, familiar learned songs evoke strong responses in song-selective CMM neurons and weak responses in ventral NCM neurons. While the responses of CMM neurons can be interpreted as a result of a feed-forward sensory hierarchy [Thompson and Gentner, 2010] that selects for increasingly complex features of birdsong, it is more difficult to understand the mechanism behind the response properties of ventral NCM neurons.

Why would neurons in NCM be driven preferentially by a large set of unfamiliar features in novel songs? Thompson and Gentner [Thompson and Gentner, 2010] speculate that the selectivity of ventral NCM neurons results from selective suppression of specific motifs in leaned songs, and that CMM could be a source of this selective suppression. In any case, it seems that neurons in both NCM and CM are tuned to experience-based features of birdsong, and could contribute to encoding the behavioral relevance of incoming auditory signals useful for identification of neighbors and kin.

11 Pre-Motor Area HVC

HVC was one of the first brain areas thoroughly researched in songbirds due to its role in song production. Early studies revealed that bilateral lesions of HVC completely eliminated the bird's ability to sing, although the accompanying behavioral aspects of singing were all still present [Nottebohm et al., 1976]. Birds with unilateral lesions were able to sing some fraction of syllables that were previously sung, but with overall marked deterioration of phase structure and instability of syllables [Nottebohm et al., 1976]. Even small microlesions in HVC were sufficient to cause destabilization of song

motifs, resulting in gross deficits in spectral and temporal organization of the motifs [Thompson and Johnson, 2007].

HVC is a sexually dimorphic brain area, appearing much more reduced in size and containing different populations of neuron types in female songbirds compared to males [Fortune and Margoliash, 1995]. In zebra finches, the total volume of HVC is five times larger in males than in females [Nottebohm and Arnold, 1976]; however, in species of songbird where both the male and the female bird sing, HVC is only about 1.1 - 1.5 times larger in males than in females [Brenowitz and Arnold, 1986].

In female birds that do not sing, HVC is thought to have a role in conspecific song perception, and bilateral lesions to HVC in female canaries produced abnormal indiscriminate sexual solicitation behavior to both conspecific and heterospecific male songs [Brenowitz, 1991]. These anatomical sex differences are thought to be functionally related to sex differences in vocal behavior [Wade and Arnold, 2004].

11.1 Anatomy and Connectivity

In the older songbird literature, HVC was (erroneously) thought to occupy the caudal part of the hyperstriatum ventrale [Nottebohm et al., 1976], and thus was referred to as the hyperstriatum ventrale, pars caudale (or HVc). Since the abbreviation HVC had become entrenched in songbird literature, recent nomenclature revisions retained the acronym HVC, which is now used as a proper name to describe this brain area [Reiner et al., 2004].

The fiber-rich ventral border of HVC is referred to as the HVC shelf [Fortune and Margoliash, 1995]. Auditory responses recorded in this area seem to resemble those of HVC itself [Katz and Gurney, 1981]. The HVC shelf receives a projection from Field L [Kelley and Nottebohm, 1979], but it is unclear how HVC shelf interfaces with HVC proper. HVC shelf may be an additional stage of processing between Field L and HVC, but it is unclear whether projections from L1 and L3 synapse onto shelf neurons that innervate HVC or if L1 and L3 neurons synapse directly onto HVC dendrites that extend into the shelf [Lewicki and Arthur, 1996].

HVC receives input both from motor nuclei involved in song production as well as input from auditory areas. It receives a projection from the medial magnocellular nucleus of anterior nidopallium (MMAN), Nif, and the thalamic uvaeform nucleus (Uva), all nuclei implicated in song production [Fortune and Margoliash, 1995]. L1 and L3 project to the HVC shelf area. It has also recently been shown that HVC receives a projection from CM [Bauer et al., 2008]. HVC projects to two song control nuclei, RA and area X.

HVC has 4 populations of neurons with distinct neurophysiological properties. It has three types of projection cells, the RA-projecting neurons, the X-projecting neurons [Nottebohm et al., 1976], and the Av-projecting neurons. Additionally a population of interneurons also exists in HVC and projects locally within HVC.

11.2 Neural Responses in HVC

Neurophysiological recordings have been made in HVC for several decades. The studies first established that auditory neurons do exist in HVC, and furthermore, that they exhibit exceptional selectivity to the BOS. While many studies have also focused on the role of HVC neurons during singing behavior [Kozhevnikov and Fee, 2007; Long et al., 2010; McCasland and Konishi, 1981], and some have revealed beautifully sparse spiking responses locked to specific elements of the motif during singing [Hahnloser et al., 2002], the following section will summarize the neurophysiological work that investigates the auditory processing capabilities of these neurons.

11.2.1 HVC Neurons Are BOS-Selective

Early studies in HVC revealed that HVC neurons respond phasically to the onsets and offsets of noise stimuli, but that tone bursts were ineffective in driving auditory responses [Katz and Gurney, 1981]. Furthermore, several neurons identified as either X-projecting or RA-projecting neurons failed to respond to either the noise or the tone bursts [Katz and Gurney, 1981]. Margoliash later reported that although most neurons in HVC responded weakly or not at all to auditory stimuli - including birdsong - he identified a population of cells in HVC that did respond vigorously to song [Margoliash, 1983, 1986].

The song-specific neurons were maximally stimulated by the BOS, and furthermore, even birds that were raised in isolation preferred the isolate BOS to other stimuli. BOS also elicited a stronger response than the TUT song the bird was repeatedly exposed to during song learning [Margoliash, 1986]. These results suggest that the neurons are selective for the self-produced BOS, and that neurons in HVC are sensitive to the slight differences in self-produced songs compared to the songs of other birds, including the tutor [Margoliash, 1986].

11.2.2 Intracellular Recordings Reveal Distinct HVC Neuron Types

HVC has three populations of projection neurons. One neuron type (HVC_X) innervates Area X, a basal ganglia homologue that has a role in vocal learning [Nottebohm et al., 1976]. Another type of projection neuron (HVC_{RA}) innervates the robust nucleus of the archistriatum (RA; [Nottebohm et al., 1976]. Both neuron types generate song-selective firing via distinct subthreshold processes. A subset of neurons project to AV, and additionally, a population of inhibitory interneurons (HVC_{Int}) modulates the subthreshold responses of the projection neurons and shapes the auditory output of these cells [Mooney, 2000].

Supporting evidence obtained in slices [Dutar et al., 1998], *in vivo* intracellular recordings revealed that each HVC neuron type had distinct intrinsic properties and distinct subthreshold auditory responses to song playback that allowed the neurons to be easily identified [Mooney, 2000]. Both the HVC_{RA} and HVC_X neurons responded to the BOS stimulus with phasic spikes occurring at specific points during the motif. In contrast, HVC_{Int} neurons fired tonically throughout the duration of the BOS [Mooney, 2000].

HVC_{RA} neurons displayed robust subthreshold depolarizations locked to the BOS, contrasting with the relatively restricted suprathreshold spiking patters of the neurons. Indeed, it seems that if one considers the magnitude of subthreshold depolarizing responses, HVC_{RA} neurons receive much more extensive auditory input than is revealed by the suprathreshold spiking patterns of these neurons [Mooney et al., 2002]. Interestingly, when the local HVC circuit was reversibly inactivated with GABA, sustained responses to BOS and REV were still detectable in HVC_{RA} subthreshold responses, indicating that the afferents to these HVC_{RA} neurons are themselves activated by both BOS and REV.

HVC_X neurons, in contrast, displayed both depolarizing as well as hyperpolarizing components in the subthreshold response to BOS [Mooney, 2000], suggesting that HVC_X neurons receive both excitatory and inhibitory inputs strongly activated by BOS but do not received inputs that are strongly activated by REV. Interestingly, during BOS playback, HVC_{Int} firing closely matched periods of HVC_X neuron hyperpolarization, suggesting that interneurons may be the source of the song-selective inhibition that shapes the subthreshold HVC_X cell response [Mooney, 2000]. Overall the data suggest that there are distinct populations of neurons that, based on their different subthreshold responses to BOS, could transmit different auditory representations of song to their respective targets [Mooney, 2000].

11.2.3 HVC_X Neurons Are Tuned to Note Duration

While it had been shown that neurons in HVC respond preferentially to acoustic features of BOS, it was not clear what these features were or whether the neurons were tuned to the spectral features of song or to the temporal context of the syllables. Margoliash had noted that neurons in HVC were driven preferentially by pairs of syllables, and often failed to respond to either syllable presented in isolation [Margoliash, 1983]. This type of nonlinear response is known as *temporal-combination-sensitivity*, and cells with this type of tuning may respond to combinations of syllables more than the linear sum of their responses to the individual syllables or to a particular combination of syllables, e.g., AB, but not by the reversed combination BA (Fig. 3.15; [Lewicki and Konishi, 1995].

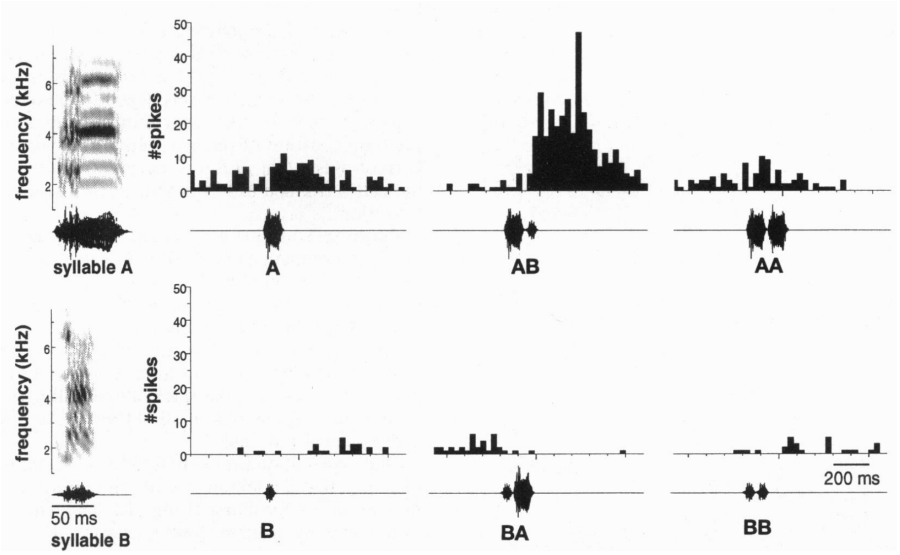

FIGURE 3.15: **Temporal-combination sensitivity of an HVC neuron.** Extracellular responses to syllables from the BOS (syllable A and syllable B; spectrogram and oscillogram on left part of panel) are represented as peristimulus time histograms. This neuron is combination-sensitive, because the neuron's response to the syllable pair AB is greater than the sum of responses to syllables A and B alone. This neuron is also sensitive to the temporal order of the syllables, since it shows no response to the syllable pair BA. Figure adapted from [Lewicki and Konishi, 1995].

Intracellular recordings in anesthetized zebra finches revealed that temporally selective responses arose from an interaction of excitatory and inhibitory post synaptic potentials [Lewicki and Konishi, 1995], and furthermore, that these responses depended on the temporal context of the motif beyond a single syllable, indicating that HVC neurons are able to integrate the auditory context over periods longer than the duration of syllable pairs [Lewicki and Arthur, 1996].

Supporting the findings of these earlier studies, the auditory responses of HVC_X neurons in a wild songbird, the swamp sparrow (*Melospiza georgiana*), were found to be

strongly dependent on the auditory context of the song, such as the sequence of notes or syllables present in a song [Prather et al., 2009]. Swamp sparrows, unlike zebra finches, sing 3-5 different song types that consist of repetitions of a single trilled, multi note syllable. HVC_X cells were unresponsive to conspecific swamp sparrow songs, mainly because the conspecific songs played back to swamp sparrow often lacked elements similar to the BOS of the test subject. When CON sounds were used that were very similar to the BOS of the experimental bird, in some cases the CON elements drove the HVC_X response better than the BOS. Therefore, the selective auditory responsiveness of HVC_X cells can also extend to similar vocalizations produced in other birds [Prather et al., 2008].

The selective auditory responses of HVC_X neurons reflect the importance of note duration as a song feature in swamp sparrows. When artificial songs were generated in which original song notes were replaced with artificially shortened or lengthened notes that match behavioral perceptibility, the auditory responses of HVC_X neurons were highly sensitive to changes in note duration. In contrast, the HVC_{Int} neurons responded similarly to the artificial notes and to the original notes, while HVC_{RA} neurons were unresponsive.

HVC_X neurons were driven by syllables containing synthetic notes that closely replicated the features of the natural note that they replaced just as effectively as the natural syllable. Importantly, syllables containing synthetic notes with durations similar to those of natural notes, but with different frequency modulations or bandwidths were also highly effective at eliciting responses from HVC_X neurons, suggesting that HVC_X neurons are exquisitely sensitive to note duration on the millisecond time scale. Furthermore, HVC_X neurons seem to categorically encode note duration rather than spectral features of the syllables [Prather et al., 2009]. These data indicate that neither frequency modulation nor frequency bandwidth has a primary role in the HVC_X response, but rather that note duration is the song feature important in the categorical perception of the note types [Prather et al., 2009].

11.2.4 Vocal-Auditory Correspondence is Present in HVC_X cells

In an early HVC study, McCasland and Konishi probed the auditory responses of neurons to birdsong both while the bird was singing itself and during passive awake listening [McCasland and Konishi, 1981]. Interestingly, they found that the individual neuronal responses during passive listening to BOS did not match the neuronal responses during singing of the BOS, suggesting that there is an interaction in HVC between motor and auditory neural activity.

Furthermore, when auditory stimuli were played while the bird sang, the motor response dominated, and little auditory response could be elicited from the neurons for up to 20 seconds after singing, even in response to songs which had previously elicited strong responses at other times. These results suggest that auditory activity in HVC is suppressed during and slightly after singing [McCasland and Konishi, 1981], and imply that HVC would we able to relay both auditory and vocal motor information directly to its targets [Mooney et al., 2002].

Prather et al. [2008] used the swamp sparrow to further investigate the interplay between vocal and auditory information in HVC. Using antidromic stimulation to identify X-projecting HVC neurons, they found that a subset of HVC_X neurons (21 out of 60) responded robustly to song playback, and of these 21 neurons, 16 HVC_X neurons responded selectively to only one song type of the swamp sparrow's repertoire. This single song from the song sparrow's repertoire was defined as the "primary song" for a particular neuron if it was the only song that could robustly drive the firing. The primary song varied across cells from the same bird, so that based on the firing response, different neurons were assigned different primary song types. These response were sparse and phasic, displaying only a few spikes at a precise phase relative to the onset of the syllable [Prather et al., 2008].

HVC_X neurons were active during both singing and listening. The most robust firing responses usually occurred while the bird sang the primary song type, which was defined by the auditory response of that neuron during song playback. Singing-related activity generally matched the auditory-evoked activity with regard to timing of phasic spiking responses but was more robust during singing. Singing-evoked activity contained bursts of spikes, whereas auditory-evoked activity consisted of single action potentials. The highly similar, temporally-precise activity patterns between the singing state and the listening state suggest that there is sensorimotor correspondence in the HVC_X cells [Prather et al., 2008].

11.2.5 HVC_X Neurons Act as Auditory Vocal Mirror Neurons

Evidence indicates that HVC_X activity during singing is a motor-related discharge. Background multiunit activity in HVC increases before the onset of singing, and the auditory response of HVC_X cells to the primary song type were suppressed during this pre-singing phase, suggesting that neurons switch from an auditory state to an auditory-insensitive state several hundred milliseconds before the onset of singing [Prather et al., 2008]. Furthermore, auditory evoked activity remained suppressed for up to 250 ms after the offset of singing [Prather et al., 2008], which is much shorter than what has been

reported in zebra finches [McCasland and Konishi, 1981]. Interestingly, a second song type also usually existed, which drove motor responses in the HVC_X cell but would not elicit auditory responses.

In cases of auditory interference, where singing overlapped with playback or vice versa, the singing-related activity was unaffected by the distortion. These results indicate that singing-related HVC_X activity is due to a motor-related corollary discharge rather than an auditory feedback-related signal, and suggests that HVC_X cells are gated to exist in either a purely auditory or purely motor states [Prather et al., 2008].

12 What Do the Non-Auditory Neurons Do?

In the previous sections we have discussed the encoding of acoustic information based on results from experiments that were made in anesthetized, sleeping, and awake, head-fixed songbirds. While these experiments are important for understanding which acoustic features of song may be behaviorally relevant, these passive listening experiments do not assess the functionality of these auditory neurons during a behaviorally relevant situation, such as during singing. A few experiments have used WN stimuli to actively perturb singing, thereby creating a distortion in the auditory feedback that the bird perceives. These experiments have revealed that a subset of auditory neurons also have a role in interpreting the live auditory feedback evoked by singing.

When noise was used to perturb active singing of the bird, Ov neurons responded with increased firing during the parts of song targeted with noise perturbation, indicating that Ov is sensitive to singing-related feedback and can encode information about the quality of vocal performance [Lei and Mooney, 2010]. Similarly, in addition to responding to spectro-temporal features of birdsong, neurons in Field L and CLM are responsive to perturbations of the song during singing [Keller and Hahnloser, 2009]. Interestingly, in some cases, and unlike the neurons in Ov, the neural responses to perturbed singing were very different from the responses to BOS playback perturbed in the same way [Keller and Hahnloser, 2009].

Figure 3.16 shows two examples of such perturbation-selective neurons. In Figure 3.16A, the neuron responded robustly to perturbation during singing and during BOS playback. In contrast, the neuron in 3.16B fired robustly in response to perturbed singing, but not to perturbed BOS playback. Based on its low firing response during singing, one could claim that this Field L neuron is not very auditory; however, its response to perturbed auditory feedback reveals that such "non-auditory" neurons have a clear function and could be involved in the top-down monitoring of auditory feedback.

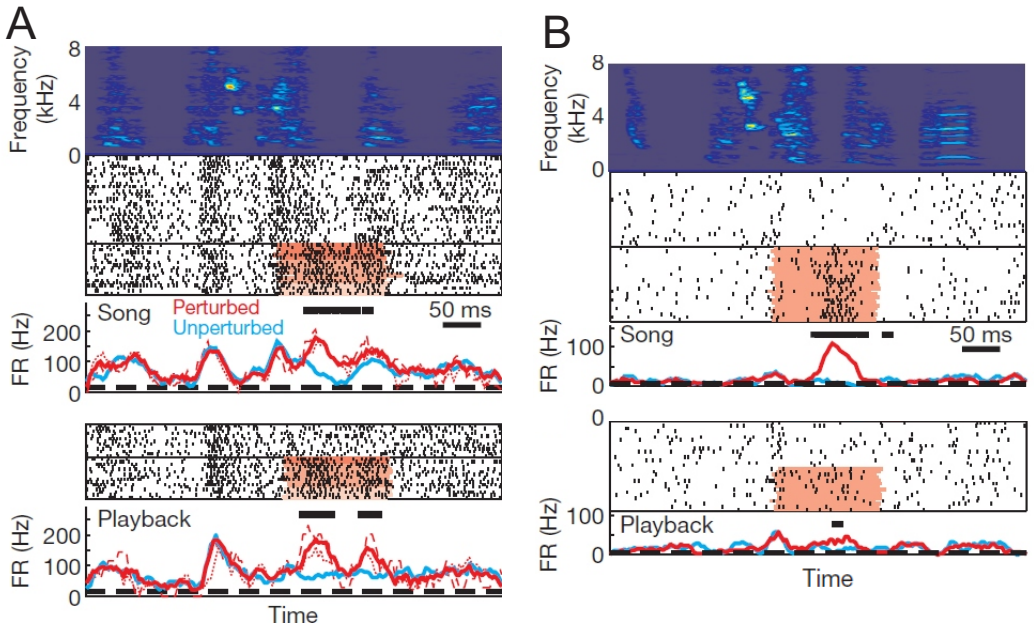

FIGURE 3.16: **Perturbation-selective neurons in the auditory forebrain. A)** Spectrogram indicates the BOS motif. When singing (top raster plot) was perturbed with a long-call stimulus (red shading), this neuron responded with an increase in firing rate. Blue line indicates the integrated firing rate during unperturbed singing, red line indicates the integrated firing rate during perturbation. Thick black bars indicate the time of significant perturbation. Black dotted lines at the bottom of the plot indicate the spontaneous firing rate. During auditory playback of BOS (lower raster plot), the neuron also responded to playback perturbation (red shading) with an increase in firing rate. **B)** Like the neuron in A, this neuron also responded to perturbation during singing (top raster plot, red shading) with an increase in firing rate (red line compared to blue line). However, unlike the neuron in A, this neuron was not selective to perturbation during BOS playback (lower raster plot). Figure adapted from [Keller and Hahnloser, 2009].

These results highlight the fact that many neurons in auditory brain areas may have non-traditional auditory responses. In MLd and Ov, both of which can be considered low-level auditory processing areas, non-trivial numbers of neurons have responses that are often considered too complex to include in the population analysis [Amin et al., 2010; Woolley et al., 2009]. When one moves to higher auditory areas like Field L, this non-trivial number of neurons explodes. Out of 647 responsive recording sites in Field L and CM, only 352 sites (or 55 percent) were found to be significantly responsive to BOS or WN [Grace et al., 2003]. In secondary auditory areas, the number of responsive neurons is often even lower. What are these complex and non-responsive neurons doing? Results suggest that some neurons in auditory brain areas from MLd to CM may have functions beyond simple feature detection. In some cases, these neurons may also serve as monitors of actual auditory feedback compared to mirrored feedback derived from an internal model of the template song [Keller and Hahnloser, 2009; Lei and Mooney, 2010]. Overall, these findings suggest that although neurons throughout the

songbird auditory forebrain are involved in detecting acoustic features of behaviorally-relevant stimuli, they may also be have functions not easily assessed with traditional auditory playback experiments.

13 Summary of the Songbird Auditory System

By examining responses from the songbird's auditory brain areas, from the midbrain MLd to the secondary auditory areas NCM and CM, one observes how information is transformed from the simple detection of sound in MLd to the highly note-specific responses found in CM and NCM. Whereas neurons in MLd and OV tend to robustly respond to many different kinds of auditory stimuli, including synthetic sounds, these promiscuous responses become rare as one moves up to higher order auditory areas such as L1, L3 and the secondary auditory areas.

Beginning in Field L, parallel and perhaps redundant streams of auditory information create receptive fields capable of encoding species-specific, behaviorally-relevant auditory information. This tuning seems to be sensitive to auditory experience and remains active even in adult birds, as demonstrated by the operant conditioning paradigms used in European starlings, which revealed experience-based preference to specific notes and motifs in CM, and the opposite response in NCM.

Overall these results suggest that at lower levels in the auditory pathway, the songbird brain is wired to organize auditory information into spectro-temporal features that represent the rhythm, pitch, and timbre of birdsong. At higher levels of processing, neurons are not tuned as much to specific spectral-temporal features as they are to specific notes, syllables, or motifs which have a behavioral meaning and are learned though experience. Furthermore, non-linear neurons, such as the perturbation selective neurons found in Ov, Field L, and CM also have a role in the active monitoring of auditory feedback, suggesting that the neurons in the auditory nuclei have complex roles beyond the simple parsing of auditory information.

Part II

Methods

Chapter 4

Methods

1 Extracellular Recording from the Songbird Auditory Forebrain

All experiments were carried out in accordance with protocols approved by the Veterinary Office of the Canton of Zurich, Switzerland.

1.1 Animals

Data were collected from 16 urethane-anesthetized adult male zebra finches (dph $>$ 120). All birds were born in our breeding colony and raised by their mother and father until dph 60. During this time, male birds were tutored naturally by their fathers. After the male birds were at least 60 days old, they were separated from their parents and joined a colony of adult male zebra finches.

One week prior to the experiment, male birds were isolated in a sound-attenuating chamber and their songs were recorded. During isolation, animals were housed individually in standard cages with a light/dark cycle of 14 and 10 hours, respectively, and received food and water *ad libitum*. Amplified songs were digitized at 44 kHz and stored on a computer.

1.2 Surgery

Birds were anesthetized with 25-40 μl of urethane injected 3 times into the breast muscles for a total volume of 75-120 μl of urethane injected per bird. 2 injections occurred

prior to surgery and were separated by 20 minutes. The final injection occurred at the end of the surgery.

During surgery, birds were additionally sedated with isoflurane (1-3% dissolved in 100% oxygen). Birds were placed in a stereotaxic device, the feathers were removed from the head, 0.5% lidocaine gel was applied to the scalp, and the scalp was dissected along the midline. The angle of the skull was adjusted to 65° with respect to the horizontal plane. Specifically, a small metal post was placed at the intersection of the beak and the skull, and a protractor was used to measure the resulting angle. The head was tilted forward or backwards in order to achieve an angle of 65°.

A small area of skull (approximately 2 x 4 mm) was removed from the upper bone layer over the auditory forebrain, and a metal plate was fastened to the skull with dental acrylic. At the completion of the surgery, the animal was administered the final urethane injection, placed in a small fabric restraint, and moved to the recording apparatus. Four small windows (approximately 500 x 500 μm) were created in the inner bone layer according to predetermined coordinates and small holes were made in the dura to allow entry of the micropipettes.

1.3 Extracellular recording

Glass electrodes (borosilicate glass, BF120-69-10; Sutter Instruments Co.) were pulled to 10-20 MΩ when filled with 0.2 M potassium chloride. Four independent electrodes (Fig. 4.1) were advanced into the auditory forebrain to the desired ventral coordinates (targeting area L2 of Field L). Extracellular signals were amplified using custom-built amplifiers, high-pass filtered at 300 Hz, and low-pass filtered at 10 kHz. Voltage signals were recorded with custom-written LabView software at a sampling rate of 30 kHz and digitized to 16-bits precision.

The auditory stimulus ensemble was played at approximately 80 dB SPL (measured at the location of the bird) through a Control 1 Pro loudspeaker system (JBL professional) placed 90 cm in front of the bird. A 130 D21 ICP Microphone (PCB Piezotronics) microphone was placed near the bird and used to record the auditory stimulus. The microphone signal was amplified and low-pass filtered at 10 kHz and high-pass filtered at 300 Hz.

To search for auditory neurons, the auditory stimulus ensemble was played continuously as one of the microelectrodes was advanced into the auditory forebrain until a neuron was encountered. This procedure was repeated until a neuronal signal was present on each of the microelectrodes. Neuronal responses were recorded if they were modulated

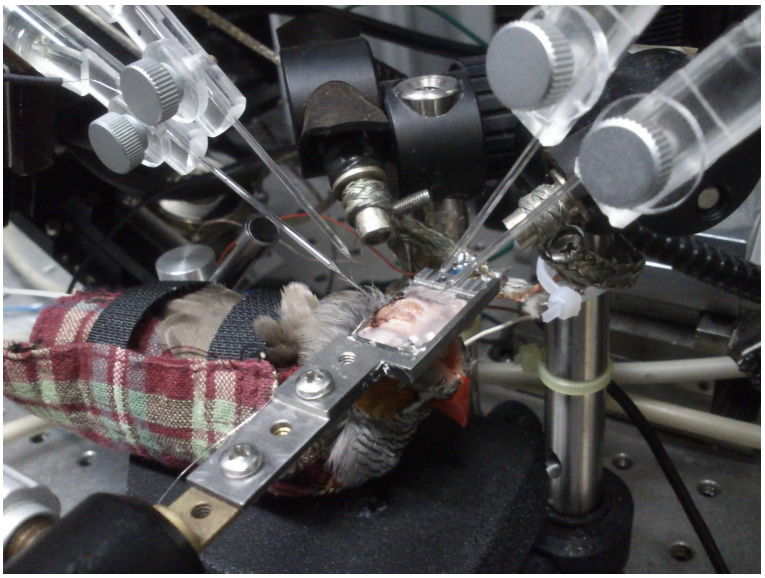

FIGURE 4.1: **Experimental animal during recording.** The bird was wrapped in a fabric restraint. A metal head-plate was fixed to the skull with dental acrylic prior to the experiment, and 4 independent glass micropipettes were used to record extracellular signals from the brain.

by auditory stimuli or if they had a large S/N ratio (S/N $>$ 5). Two Axoclamp 2B amplifiers were used in bridge mode to simultaneously record the extracellular signals.

We recorded from 229 neurons (272 pairs) in the auditory forebrain of 16 birds.

1.4 Histology

At the end of the recording session, small electrolytic lesions (10 μA, 5-second duration) were made with a 1.2 MΩ metal electrode in some birds to aid in the histological verification of the electrode tract. Putative neuron locations were reconstructed using the electrode coordinates and the depth measurements recorded during the experiment. Putative neuron localization was achieved using the coordinates from Fortune and Margoliash [1992]; Nixdorf-Bergweiler and Bischof [2007], as well as the spiking responses to BOS and REV stimuli.

2 Auditory Stimulus Design

A unique auditory stimulus ensemble was created for each experimental bird using custom-written Matlab software (Florian Blättler) and included both natural and synthetic sounds. The entire stimulus ensemble lasted 20-25 minutes.

The natural stimuli included the BOS, REV, CON, quiet BOS (qBOS; the BOS played at a sound pressure level reduced by approximately 10 dB), and BOS perturbed by the call of another bird (perturbed BOS; pBOS).

CON and BOS are natural stimuli which have both been shown to effectively drive neurons in the auditory forebrain [Amin et al., 2004]. REV is a temporally reversed version of BOS, such that both the sequence of syllables as well as the temporal order within the syllables is reversed, changing the sound envelope of the stimulus, while preserving the overall power spectrum of the song.

CON songs were chosen from our song archive, which contained songs from naturally tutored male birds recorded during isolation as described for the experimental birds. In order to control for differences in loudness that may occur across birdsong, the BOS, REV and CON songs included in the ensemble were normalized to the identical loudness by calculating the root mean square (rms) of the songs and normalizing them to a fixed loudness of 0.002. All songs were high-pass filtered at 400 Hz.

In addition to the natural stimuli used in the stimulus ensemble, synthetic stimuli were also included. Synthetic stimuli included WN, tones, and stacks. WN was included at the start (WN-S) and end (WN-E) of the stimulus ensemble so that the stationarity of the recording could be assessed by comparing firing rates at the start of the file with those at the end of the file.

2.1 Synthetic Stimulus Design

In order to compare responses between natural and artificial stimuli, we used the acoustic statistics of the natural stimuli in the generation of the artificial stimuli. Synthetic stimuli were created following methods described by [Grace et al., 2003].

Tones

The power spectrum of the BOS was randomly sampled to obtain frequencies of pure tones, 20 of which were added together and normalized to achieve the same overall power spectrum as the BOS. These tones were used to assess whether the response of an auditory neuron could be reduced to its frequency tuning.

Stacks

Artificial harmonic stacks, used to imitate the harmonic stacks heard in bird song, were used to assess the auditory neuron's response to slow-varying spectral modulations.

This stimulus was created by adding harmonic frequencies upon a fundamental frequency of 700 $\pm$ 100 Hz, which matches the fundamental frequencies found in zebra finch song. The amplitude of each harmonic was modulated with a cosine function (mean period equal to 4 kHz $\pm$ 3 kHz), such that some frequency components were enhanced while others were suppressed. The overall power spectrum of the stacks were flat from 700 Hz to 8 kHz.

White Noise

WN was chosen because it is a random acoustic signal that is uncorrelated in time and frequency and was used to drive neurons which may not be responsive to sounds found in bird song or the other synthetic stimuli. WN samples were band-passed from 16 Hz to 8 kHz. The WN power spectrum was flat between these boundaries, and the overall power summed across all frequencies matched that of the BOS.

Each artificial stimulus lasted approximately 2 seconds, and individual stimuli were separated by 2 seconds of silence. The entire auditory stimulus ensemble lasted approximately 20-25 minutes. Spectrograms of each of the auditory stimuli are presented in Figure 4.2.

2.2 Stimulus Ensemble Sequence

The stimulus ensemble consisted of blocks of stimuli in the following order: WN-S, Tones, Stacks, REV, CON, BOS, qBOS, pBOS, and WN-E (Fig. 4.3). The Tones, Stacks, qBOS, and pBOS stimulus blocks were not included in all experiments. Each stimulus block contained repetitions of the stimulus (see Fig. 4.4), and each repetition was separated by 2 seconds of silence.

The WN-S and the WN-E blocks consisted of 10 unique versions of WN repeated 3 times each or 20 unique versions repeated 2 times each for a total of 30 or 40 WN repetitions, respectively (see Fig. 4.4, WN-S v1 and v2). The WN-S and WN-E blocks at the beginning and end of the stimulus ensemble were identical.

The Tones and Stacks block consisted of 5 unique versions of Tones or Stacks repeated 4 time each for a total of 20 Tones or Stacks repetitions (Fig. 4.4).

For the natural stimuli constructed from the BOS, 6 unique versions of BOS were repeated 3, 5, or 10 times each, for a total of 18, 30, or 60 repetitions, respectively (Fig. 4.4, BOS v1, v2, v3). The REV block used the same song versions selected for the BOS. For the REV block, the 6 or 10 unique REV versions were repeated 3 or 10 time each for a total of 18, or 60 repetitions of REV, respectively (Fig. 4.4, REV v1 and v2).

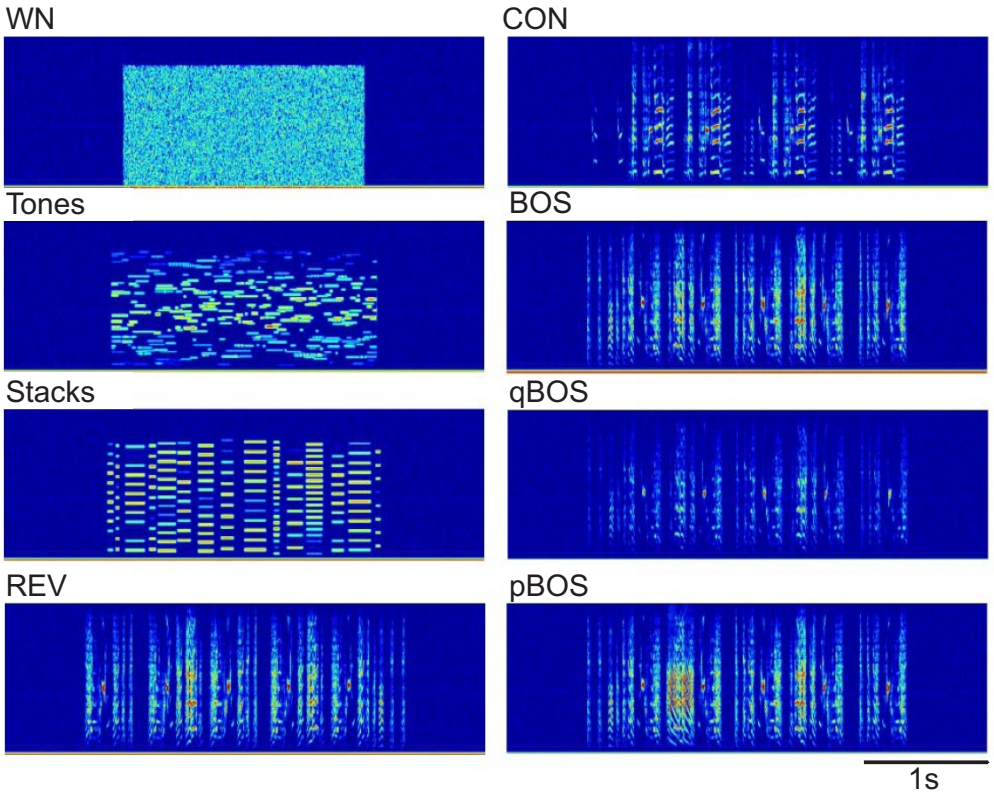

FIGURE 4.2: **Auditory stimuli included in the stimulus ensemble.** Spectrograms depict examples of auditory stimuli included in the auditory stimulus ensemble (see text for details). All spectrograms are depicted in the range of 0 to 10 kHz.

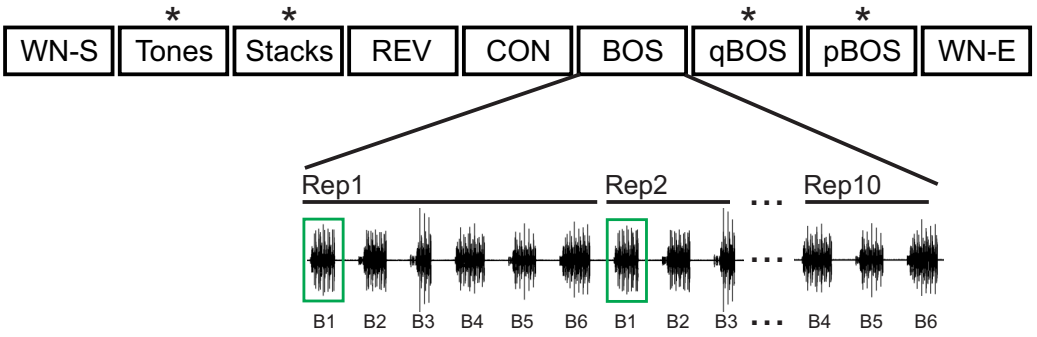

FIGURE 4.3: **Repetition sequence for stimulus ensemble.** The auditory stimulus ensemble consisted of blocks of stimuli including WN that appeared at the start (WN-S) and end (WN-E) of the playback file, Tones, Stacks, REV, CON, and BOS, qBOS, and pBOS (see text for details). Each stimulus block contained unique stimulus versions that were repeated non-consecutively and were separated by 2 s of inter-stimulus silence. The lower part of the panel depicts a part of the BOS stimulus block. A sequence of 6 different versions of BOS (B1-B6; depicted as oscillograms) was played, after which the entire sequence repeated until 10 repetitions of each version had been played. The green box indicates a repetition of identical BOS versions. The asterisks indicate stimulus types that were not included in all experiments.

Stimulus	Unique Versions	Reps/Version	Total Repetitions
WN-S (v1)	10	3	30
WN-S (v2)	20	2	40
Tones*	5	4	20
Stacks*	5	4	20
REV (v1)	6	3	18
REV (v2)	6	10	60
CON (v1)	40	1	40
CON (v2)	75	1	75
CON (v3)	100	1	100
BOS (v1)	6	3	18
BOS (v2)	6	5	30
BOS (v3)	6	10	60
qBOS*	6	3	18
pBOS*	1	3	3
WN-E (v1)	10	3	30
WN-E (v2)	20	2	40

FIGURE 4.4: **Numbers of repetitions for auditory stimulus ensemble.** For each stimulus block (Stimulus), different numbers of unique versions of the stimulus (Unique Versions) were used in the stimulus ensemble. These unique version were repeated a certain number of times (Reps/Version) to achieve a total number of repetitions per stimulus block (Total Repetitions). See text for details. Asterisks indicate stimuli that were not included in all experiments.

Within the stimulus block, the sequence of repetitions was such that each unique stimulus version did not repeat consecutively. That is, the first unique version of BOS (B1, see Fig. 4.3) was followed by the second unique version of BOS (B2), and so on, until all unique version had been played, after which the song sequence repeated (see Fig. 4.3 for details).

The CON stimulus block contained 2, 3, or 4 songs from 20 or 25 birds, for a total of 40, 75, or 100 unique CON songs, respectively (Fig. 4.4; CON v1, v2, v3). CON songs were not repeated.

3 Spike Time Extraction

Spike times were extracted from the voltage signal using custom written Matlab software (Richard Hahnloser). Briefly, spike waveforms were detected by setting a voltage threshold. Waveform clusters were created from the detected spike waveforms using K-means clustering. Generally, waveform clusters were Gaussian-distributed and the detected spikes were characterized by large signal-to-noise ratio (median S/N = 17.30 $\pm$ 9.74 (std)) Multi-unit signals were infrequent and were discarded unless they could be reliably discriminated into single units. In order to verify single-unit signals, we

calculated both the inter-spike-interval and the autocorrelation function for all of the detected spikes, and eliminated double detections accordingly. If movement artifacts were observed during the recording, all of the spikes that occurred during the stimulus repetition containing the movement artifact were deleted during the offline spike extraction.

4 Stimulus Alignment

Our stimulus ensemble consisted of stimulus repetitions interleaved with 2 s of silence. One of our major interests was to compare neural responses that occurred spontaneously during the 2 s of silence to responses that were evoked by the auditory stimulus. Because we used a long, continuous auditory ensemble, it was necessary to align the neural responses to stimulus onsets and offsets for our subsequent analysis. In the following sections, we will discuss neural activity that occurred spontaneously during the silent interval (Spont) and neural responses that occurred during the WN, REV, CON, or BOS stimuli (Stim). Because were we also curious about neural responses that occurred during the REV, CON, and BOS motifs, we will denote responses that were collected during the motif as $\text{Stim}_{\text{Motif}}$ (e.g., $\text{BOS}_{\text{Motif}}$) and the responses that were collected for the longer song bout as $\text{Stim}_{\text{Bout}}$ (e.g., BOS_{Bout}).

4.1 Stimulus Onset and Offset Definition

For each of the auditory stimulus ensemble used in these experiments, the onset and offset times of each stimulus were assigned off-line using custom-written Matlab software. The energy and the spectrogram of the auditory stimulus were visualized, and the onset and offset times were defined based on the crossing of the energy threshold. Because the inter-stimulus intervals of silence preceded and followed each stimulus repetition, the segmentation of unique stimuli was easily achieved.

To ensure that identical stimulus repetitions had identical lengths, the onset and offset times were set once for each unique rendition of the stimulus from which the stimulus length of that unique stimulus was defined. For all subsequent repetitions, only the stimulus onsets were assigned, and the stimulus offsets were defined by adding the calculated stimulus length for the respective stimulus to the manually set onset time.

4.2 Spontaneous Interval Onset and Offset Definition

Once the onsets and offsets were defined for the stimuli at the bout level, the spontaneous onsets and offsets were defined as the time point that occurred 500 ms after the offset of the previous stimulus. We chose to use spontaneous responses that occurred 500 ms after the offset of the previous stimulus in order to avoid offset responses, which can last up to 150 ms [Leppelsack, 1974]. We used two different lengths of spontaneous intervals, so the spontaneous offset was defined as the time point occurring 1 s or 1.5 s after the corresponding spontaneous onset.

4.3 Motif Onset and Offset Definition

The onsets and offsets of the motifs for the REV, BOS, and CON bouts were defined using the same technique used in the song bout onset and offset definition. A motif was defined here as an identically repeating pattern of syllables. In the case of the BOS and REV motifs, the onsets and offsets of all of the motifs were manually defined using the energy threshold crossing for each of the 6 unique BOS or REV bout versions that were included stimulus ensemble, and the mean length of the motif was calculated from these motif repetitions. For subsequent motifs, only the onsets were manually defined using the energy threshold crossing. We defined the motif offset for all motifs by adding the calculated mean motif length to the manually defined motif onset, so as to ensure that all the motif lengths were identical.

The CON motifs were defined in the same way as for the REV and BOS motifs; however, since there were overall fewer repetitions of CON motifs, we manually set the onset and offset of all motif repetitions in order to determine the mean unique CON motif length. In the same way as done for the REV and BOS motifs, the stimulus offsets were then redefined by adding the mean unique CON motif length to the CON motif onset, in order to ensure that identical repetitions of CON motifs had identical lengths.

5 Single Neuron Analysis

5.1 Z-Score

The z-score represents the normalized difference between a neuron's stimulus-evoked mean firing rate its spontaneous mean firing rate. For our analysis, we compared the stimulus-evoked mean firing rate to the full 2 s of spontaneously evoked inter-stimulus

silence preceding the stimulus. The z-score was calculated for each neuron in response to each stimulus as follows:

$$Z = \frac{(\mu_{Stim} - \mu_{Spont})}{\sqrt{(\sigma^2_{Stim} + \sigma^2_{Spont} - 2Cov(Stim, Spont))}}$$

where μ_{Stim} is the mean firing rate for the stimulus-evoked firing rate, μ_{Spont} is the mean firing rate for the 2 seconds of spontaneously-evoked firing, σ^2 is the variance of the response, and $2Cov(Stim, Spont)$ is the covariance between the stimulus and the response [Theunissen et al., 2004b].

5.2 D-Prime Score

The d-prime score for a two-choice, forced-alternative test [Swets, 1961] can be used to quantify the selectivity preference of a neuron between two stimuli A and B. Positive d-prime values indicate that stimulus A elicited a greater response than stimulus B and oppositely if the d-prime value is negative. A d-prime value of zero indicates no difference in the responses evoked by the stimuli. Neurons are considered positively selective if their d-prime values are greater than 0.5 [Amin et al., 2004] and suppressed if their d-prime values are less than -0.5. The d-prime score was calculated for each neuron in response to each stimulus as follows:

$$d_{A-B} = \frac{2(\mu_A - \mu_B)}{\sqrt{\sigma^2_A + \sigma^2_B}}$$

where μ_A is the mean firing rate for the stimulus A, μ_B is the mean firing rate for the stimulus B, and σ^2 is the variance of the responses to the stimuli A and B.

5.3 Coefficient of Variation, CV

We calculated the coefficient of variation to characterize distributions of values (such as z-scores or firing rates) across stimuli. The coefficient of variation (CV), defined as the ratio of the standard deviation to the mean, was calculated as follows:

$$CV = \frac{\sigma}{\mu}$$

where σ is the standard deviation and μ is the mean.

5.4 Spectro-Temporal Receptive Field Estimation

The STRFs of selected neurons were estimated using the STRFPak version 5.3, available from http://strfpak.berkeley.edu/. We estimated STRFs from the CON repetitions included in the auditory stimulus ensemble. Custom-written matlab software was used to create stimulus wav files and corresponding spike time text files from the raw data files and the sorted spike times.

The CON stimuli were first preprocessed with the STRFPak using a short time Fourier transform. The frequencies of the stimulus were divided into 59 bands of 125 Hz, and ranged from 250 Hz to 8000 Hz. Each filter width was 1.273 ms in duration.

We calculated the STRFs over 200 ms using the the mean-subtracted model (r+STRF) and the space-time non-separable STRF algorithm. We manually withheld 10% of the stimuli for STRF validation. Peri-stimulus time histograms (PSTH) were smoothed with an 11-ms smoothing window, and the correlation coefficient was calculated between the smoothed PSTH and predicted response of the estimated STRF in order to validate the estimated STRF.

5.4.1 STRF Similarity Index, SI

We used the STRF similarity index (SI; [DeAngelis et al., 1999; Escabi and Schreiner, 2002; Graña et al., 2009]) to quantify the similarity between different STRFs. We treated the STRF pixel values as vectors by reading the values from each STRF matrix down its columns. These vectorized STRFs were then used to calculate the similarity index

$$SI = \frac{\langle STRF_A, STRF_B \rangle}{|STRF_A| \cdot |STRF_B|}$$

where where $\langle , \rangle$ is the inner product of the vectors and $|\cdot|$ is the vector norm operator, which normalizes the index to the range from -1.0 to 1.0. Values of SI close to 1 indicate strong similarity between the two STRFs. If the two STRFs are identical but one is the inverse of the other, then SI = -1.0. If, for example, the two STRFs differ in phase by 90°, then the SI = 0. The SI is numerically identical to the Pearson correlation coefficient.

6 Pairwise Correlation Analysis

6.1 Spike Count Correlations, R_{SC}

Spike count correlations (R_{SC}) have been used extensively to estimate the level of covariation in firing rate that is shared across neurons pairs. Considering that R_{SC} have been measured over stimulus durations ranging from several seconds to several hundreds of milliseconds [Cohen and Kohn, 2011], we calculated spike count correlations for 2 different time scales. For the Spont and Stim-evoked (song bout) responses, we calculated R_{SC} from 1.5 s of data. For the Stim-evoked (song motif) responses, we calculated R_{SC} from 200 ms of data.

Generally, 1.5 s of a CON or BOS bout includes introductory notes and approximately 1-2 repetitions of the motif (Fig. 4.5). Because REV is a completely reversed version of the BOS, it is different BOS and CON in the sense that the end of the song is at the beginning, i.e., introductory notes are not usually present in the first 1.5 s of the song. 200 ms of a motif includes 1-2 syllables of the motif; it does not include introductory notes.

R_{SC} were calculated for on spike counts for identical repetitions using the Pearson's correlation coefficient. Before computing R_{SC}, we converted the data for each repetition into z-scores to normalize spike counts for each condition. The statistical significance of R_{SC} was determined using the Matlab function "corrcoef" and a correlation was considered to be significant if p was < 0.05. The R_{SC} for spike counts x and y was calculated as follows:

$$R_{SC} = \frac{Cov(x(T), y(T))}{\sqrt{(Var(x(T), x(T)) \cdot Var(y(T), y(T)))}}$$

where $x(T)$ and $y(T)$ are spike counts of neurons x and y in synchronous time bins of length T [Tchumatchenko et al., 2011]. If spike counts are identical, than the R_{SC} is equal to 1, and if the spike trains are independent of each other, than the R_{SC} is 0.

In order to statistically evaluate distributions of R_{SC} across stimuli, we converted R_{SC} values to z-scores using the Fisher transformation [Smith and Kohn, 2008; Zohary et al., 1994] as follows:

$$Z = \frac{1}{2} ln \frac{1 + R_{SC}}{1 - R_{SC}}$$

By converting the spike counts to z-scores, confounding variables such as stimulus response strength were eliminated [Zohary et al., 1994].

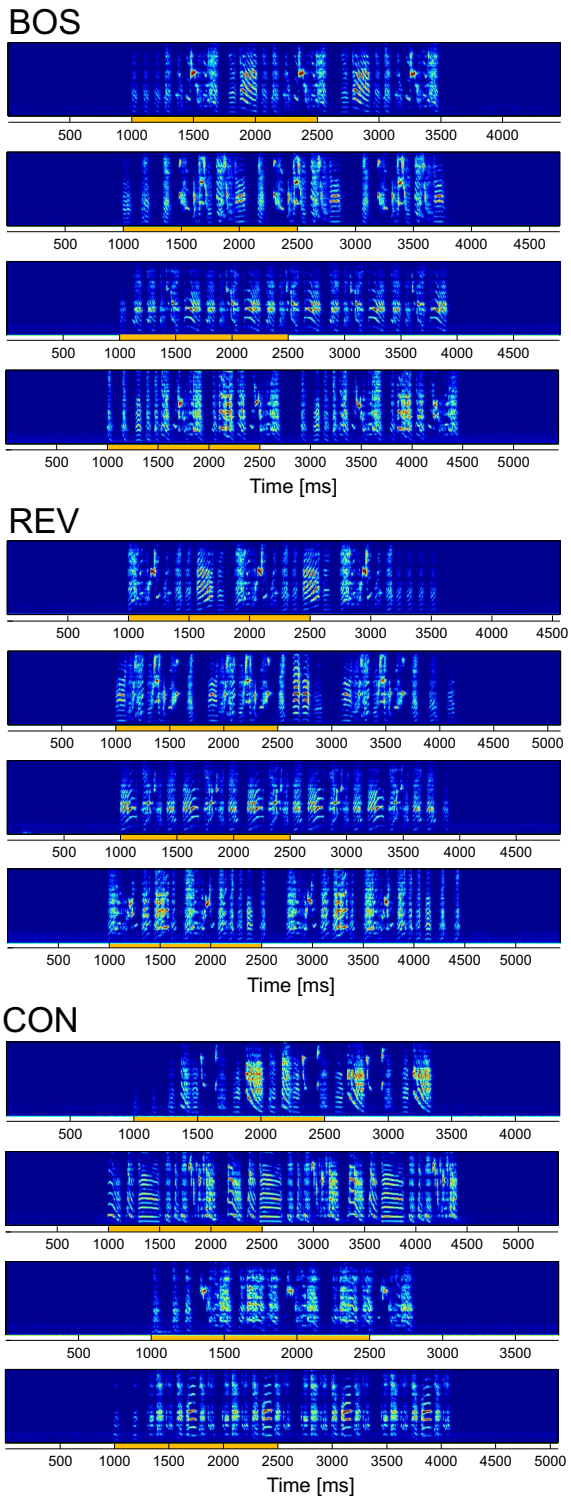

FIGURE 4.5: **Response duration used in spike count correlation analysis.** 1.5 s duration is highlighted in yellow for the different example spectrograms of BOS, REV, and CON. For most BOS and CON songs, 1.5 s includes introductory notes and 1-2 repetitions of the motif. Introductory notes were not included in the analysis for the REV stimulus. All spectrograms are depicted in the range of 0 to 8 kHz.

6.2 Spike Train Cross-Covariance, CCV

Cross covariances (CCV) were calculated to assess spike time correlations that exist among pairs of auditory forebrain neurons. The cross-covariance calculation is a normalized (mean-subtracted) version of the cross-correlation (or cross-correlogram of Perkel et al. [1967a]).

CCVs were calculated for both spontaneous and stimulus-evoked responses. For the spontaneous responses, we calculated CCVs from 1 s of data. For the $\text{Stim}_{\text{Bout}}$ and $\text{Stim}_{\text{Motif}}$ responses, we calculated CCV from the entire duration of the stimulus.

All spiking responses were first discretized into 5 ms bins. CCV were calculated on the binned spike vectors using the Matlab function "xcov" with the option set to "coef". Briefly, this function calculates the cross-correlation after subtracting the means from each trial and normalizes by the geometric mean of the autocorrelation.

We were interested in deconstructing the CCV correlations into component correlations induced by the stimulus and component correlations due to shared intrinsic noise. We achieved this in the following way.

6.2.1 CCV

We calculated pairwise correlations for each synchronous presentation of identical spontaneous and stimulus-evoked repetitions, and then averaged over all of the repetitions for each unique stimulus version. We will refer to these correlations as the "CCV" for clarity. The CCV correlations represent the sum of the independent correlations due to both correlations induced by the stimulus and intrinsic noise correlations.

The CCV was calculated for spike trains A and B as follows:

$$CCV_{A-B}^{j,i}(\tau) = \langle \frac{1}{T} \int_0^T \Big(r_A^{j,i}(t) - \bar{r}_A(t) \Big) \Big(r_B^{j,i}(t+\tau) - \bar{r}_B(t+\tau) \Big) dt \rangle$$

Spikes from train $r_A(t)$ were compared to spikes from train $r_B(t)$ as a function of time lag τ. In our analysis, we calculated CCVs for τ = 5 or 10 ms and for T = 1 s. $\bar{r}_A(t)$ and $\bar{r}_B(t)$ represent the time-varying mean firing rates, and $\langle\,\rangle$ indicates that the CCV calculation was averaged over all trials. The index i refers to the i^{th} repetition of stimulus version j.

6.2.2 Trial-Shuffled CCV

In order to tease apart the extrinsic correlations driven by the auditory stimulation from the intrinsic, noise-induced correlations, we shuffled among identical stimuli. We calculated the shuffle predictor by correlating the responses from neuron A during the i^{th} repetition (of N total repetitions) with the responses from neuron B during the i+1 repetition. For i=N, i+1 was set to equal 1. The trial-shuffled CCV represents the extrinsic, time-invariant correlations that are induced by the stimulus, and therefore we will refer to these correlations as the "trial-shuffled CCV" for clarity.

The trial-shuffled CCV was calculated as follows:

$$\check{CCV}^{j,i}_{A-B}(\tau) = \langle \frac{1}{T} \int_0^T \Big(r_A^{j,i}(t) - \bar{r}_A(t) \Big) \Big(r_B^{j,i+1}(t+\tau) - \bar{r}_B(t+\tau) \Big) dt \rangle$$

The notation for the trial-shuffled CCV is the same as for the CCV calculation, and only differs in the i+1 shuffling of responses.

The matlab "coef" option normalized the CCV calculations in the following way:

$$K(\tau) = \frac{CCV_{A-B}(\tau)}{\sqrt{CCV_{A-A}(0)}\sqrt{CCV_{B-B}(0)}}$$

where $CCV_{A-A}(0)$ and $CCV_{B-B}(0)$ represent the auto-covariance of spike trains A and B at $\tau = 0$.

6.2.3 Noise-Covariance

Assuming that the extrinsic stimulus-evoked correlations are independent from the intrinsic noise correlations, we subtracted the mean trial-shuffled CCV from the mean CCV in order to estimate the contribution of intrinsic noise to the CCV. We will refer to this as the "noise-covariance".

The noise covariance was calculated as the difference between the CCV and the trial-shuffled CCV, or

$$CCV^{j,i}_{A-B}(\tau) - \check{CCV}^{j,i}_{A-B}(\tau)$$

We assessed the significant peaks in the noise-covariance by calculating a confidence interval that was 3 times the standard deviation of the baseline values found at long time lags ranging from -1000 ms to -900 ms and +900 ms to +1000 ms [Bair et al., 2001; Huang and Lisberger, 2009; Kohn and Smith, 2005].

6.3 Figure Notation

Unless otherwise mentioned in the figure caption, all CCV, trial-shuffled CCV, and noise-covariance functions were smoothed using the Matlab "lowess" function using a span of 11 data points and corresponding to a window of 55 ms. The raw, unsmoothed data was used in all analysis.

7 Population Analysis

In order to get a sense of how spiking responses are correlated in a population of neurons, we averaged spike time cross-covariance functions and spike count correlations over pairs of neurons recorded in different birds.

7.1 Pooled Spike Count Correlations

For the spike count correlation R_{SC} values, we are aware that the number of repetitions included in the analysis is the prime determinant of the level of R_{SC} that is associated with statistical significance. Therefore, a neuron pair may show a small but functionally significant noise correlation that does not reach statistical significance simply because we did not record enough trials [Huang and Lisberger, 2009]. For this reason, we included all the pairs of neurons we recorded in our assessment of the structure of correlations across the population, and mention significant R_{SC} values ($p < 0.05$) accordingly.

7.2 Pooled Spike Train Cross Covariances

In order to correct for asymmetrical CCV functions that might result from one neuron leading another in firing, we symmetrized the mean CCV functions. That is, for each of the mean CCV functions included in the analysis, we added the original CCV calculation to a transposed version of the CCV calculation and divided the sum by 2.

8 Efficient Coding Simulations

We were curious to see whether our results could be confirmed by an efficient coding model recently proposed by Blättler and Hahnloser [2011]. In order to determine whether the correlations generated by our population of auditory neurons could reflect a general pattern of decorrelation that accompanies stimulus-selective neural responses,

we used a custom-designed GUI to run the efficient coding algorithms implemented in Blättler and Hahnloser [2011].

8.1 Song Selection and Preparation

11 birds for which we had recorded more that 1000 songs each were used in these simulations. Songs were recorded while birds were isolated in a sound-attenuating chamber. During isolation, animals were housed individually in standard cages with a light/dark cycle of 14 and 10 hours, respectively, and received food and water *ad libitum*. Amplified songs were digitized at 44 kHz and stored on a computer. Raw song files were pre-processed such that at most 200 ms of silence were present before the onset of the first note in the song and after the last note in the song. For each bird, we selected songs that were approximately of the same length and contained at least 3 motifs. In order to control for differences in loudness across songs, the rms of each song was normalized to a fixed loudness of 0.002.

8.2 Training the Networks

For each of the 11 birds, we created an archive of songs to be used in the training of the auditory network. For each of the 11 archives, we used 500 BOS songs and 100 CON songs (a BOS-to-CON ratio of 5:1). 20 songs selected from 5 different conspecific birds contributed to the 100 CON songs. Songs were counter-balanced across archives, such that songs used as BOS for one archive were used a CON songs in another archive. Specifically a selection of 20 BOS songs (out of 500) were used as CON songs in 5 other archives. Specific details about the birds used in each of the training archives can be found in Figure 4.6. Once archives containing a total of 600 songs were prepared, we used the custom-written *Flatclust* software to develop a network of 1000 neurons.

8.3 Testing the Network

Once we had trained the networks, we tested each network with archives containing 10 novel BOS versions not included in the training archive, 10 quiet versions of the BOS songs (qBOS), 10 WN versions, and 10 novel CON songs, 2 songs each from 5 conspecific birds not used during the training (see Fig. 4.6 for details). WN and qBOS songs were created using the same methods as described in the above *Auditory Stimulus Design* section. Responses to each of the 10 versions of the stimuli were collected from each of the 1000 neurons. REV responses were generated from transposed versions of the BOS songs. For each of the 11 networks, we analyzed responses from 1000 neurons

Training	b1r10-Network [1]	
	Bird	# Files
BOS Files	b1r10	500
CON Files	b16r10	20
	g3r10	20
	g5r10	20
	g8r1	20
	g8r10	20
Testing	b1r10-Network [1]	
	Bird	# Files
BOS Files	b1r10	10
CON Files	o3r8	2
	o11r18	2
	p2r10	2
	r7y5	2
	r10s4	2
REV Files	o1r10	10
qBOS Files	o1r10	10
WN Files	-	10

Training	b16r10-Network [2]	
	Bird	# Files
BOS Files	b16r10	500
CON Files	g3r10	20
	g5r10	20
	g8r1	20
	g8r10	20
	o3r8	20
Testing	b16r10-Network [2]	
	Bird	# Files
BOS Files	b16r10	10
CON Files	o11r18	2
	p2r10	2
	r7y5	2
	r10s4	2
	b1r10	2
REV Files	b16r10	10
qBOS Files	b16r10	10
WN Files	-	10

Training	g3r10-Network [3]	
	Bird	# Files
BOS Files	g3r10	500
CON Files	g5r10	20
	g8r1	20
	g8r10	20
	o3r8	20
	o11r18	20
Testing	g3r10-Network [3]	
	Bird	# Files
BOS Files	g3r10	10
CON Files	p2r10	2
	r7y5	2
	r10s4	2
	b1r10	2
	b16r10	2
REV Files	g3r10	10
qBOS Files	g3r10	10
WN Files	-	10

Training	g5r10-Network [4]	
	Bird	# Files
BOS Files	g5r10	500
CON Files	g8r1	20
	g8r10	20
	o3r8	20
	o11r18	20
	p2r10	20
Testing	g5r10-Network [4]	
	Bird	# Files
BOS Files	g5r10	10
CON Files	r7y5	2
	r10s4	2
	b1r10	2
	b16r10	2
	g3r10	2
REV Files	g5r10	10
qBOS Files	g5r10	10
WN Files	-	10

Training	g8r1-Network [5]	
	Bird	# Files
BOS Files	g8r1	500
CON Files	g8r10	20
	o3r8	20
	o11r18	20
	p2r10	20
	r7y5	20
Testing	g8r1-Network [5]	
	Bird	# Files
BOS Files	g8r1	10
CON Files	r10s4	2
	b1r10	2
	b16r10	2
	g3r10	2
	g5r10	2
REV Files	g8r1	10
qBOS Files	g8r1	10
WN Files	-	10

Training	g8r10-Network [6]	
	Bird	# Files
BOS Files	g8r10	500
CON Files	o3r8	20
	o11r18	20
	p2r10	20
	r7y5	20
	r10s4	20
Testing	g8r10-Network [6]	
	Bird	# Files
BOS Files	g8r10	10
CON Files	b1r10	2
	b16r10	2
	g3r10	2
	g5r10	2
	g8r1	2
REV Files	g8r10	10
qBOS Files	g8r10	10
WN Files	-	10

Training	o3r8-Network [7]	
	Bird	# Files
BOS Files	o2r8	500
CON Files	o11r18	20
	p2r10	20
	r7y5	20
	r10s4	20
	b1r10	20
Testing	o3r8-Network [7]	
	Bird	# Files
BOS Files	o3r8	10
CON Files	b16r10	2
	g3r10	2
	g5r10	2
	g8r1	2
	g8r10	2
REV Files	o3r8	10
qBOS Files	o3r8	10
WN Files	-	10

Training	o11r18-Network [8]	
	Bird	# Files
BOS Files	o11r18	500
CON Files	p2r10	20
	r7y5	20
	r10s4	20
	b1r10	20
	b16r10	20
Testing	o11r18-Network [8]	
	Bird	# Files
BOS Files	o11r18	10
CON Files	g3r10	2
	g5r10	2
	g8r1	2
	g8r10	2
	o3r8	2
REV Files	o11r18	10
qBOS Files	o11r18	10
WN Files	-	10

Training	p2r10-Network [9]	
	Bird	# Files
BOS Files	p2r10	500
CON Files	r7y5	20
	r10s4	20
	b1r10	20
	b16r10	20
	g3r10	20
Testing	p2r10-Network [9]	
	Bird	# Files
BOS Files	p2r10	10
CON Files	g5r10	2
	g8r1	2
	g8r10	2
	o3r8	2
	o11r18	2
REV Files	p2r10	10
qBOS Files	p2r10	10
WN Files	-	10

Training	r7y5-Network [10]	
	Bird	# Files
BOS Files	r7y5	500
CON Files	r10s4	20
	b1r10	20
	b16r10	20
	g3r10	20
	g5r10	20
Testing	r7y5-Network [10]	
	Bird	# Files
BOS Files	r7y5	10
CON Files	g8r1	2
	g8r10	2
	o3r8	2
	o11r18	2
	p2r10	2
REV Files	r7y5	10
qBOS Files	r7y5	10
WN Files	-	10

Training	r10s4-Network [11]	
	Bird	# Files
BOS Files	r10s4	500
CON Files	b1r10	20
	b16r10	20
	g3r10	20
	g5r10	20
	g8r1	20
Testing	r10s4-Network [11]	
	Bird	# Files
BOS Files	r10s4	10
CON Files	g8r10	2
	o3r8	2
	o11r18	2
	p2r10	2
	r7y5	2
REV Files	r10s4	10
qBOS Files	r10s4	10
WN Files	-	10

FIGURE 4.6: **Files used to train and test the simulated auditory networks.** The table displays the names of birds and numbers of files used during the training (light pink) and testing (light blue) for each of the 11 networks. All songs used during the testing were novel songs to the network. See text for details.

in response to 10 different versions of 5 different stimulus classes: BOS, REV, CON, qBOS, and WN.

8.4 Data Analysis

In response to each stimulus rendition, each neuron in the network responded with a unique synaptic current. The raw synaptic current for each neuron was mean-subtracted and thresholded in order to create discrete spiking responses and firing rates. We used thresholds of 0, 0.5, 1, 2, 3, 4, 5, and 6 times the standard deviation of each neuron's mean-subtracted response.

8.4.1 Individual Neuron Statistics

For each member of the 1000-neuron network, we calculated mean firing rates to each of the stimulus classes by averaging responses over each of the 10 stimulus versions. Similarly, d-prime scores were calculated for each neuron using the mean firing rate calculated for the 10 stimulus versions and the corresponding standard deviation. We

examined in detail d-prime scores involving the BOS song (i.e., BOS versus CON, BOS versus REV, BOS versus qBOS, and BOS versus WN).

8.4.2 Pairwise CCV calculations

We calculated pairwise CCVs between every neuron in the network and 10 randomly selected, non-identical members of the network.

In order to calculate CCVs from the simulated neuron responses, we first discretized the spike times obtained by thresholding into 10 ms bins. We calculated the raw pairwise CCVs for each of the 1000 neurons and its 10 connectivity pairs using the same method as described above in the *Spike Train Cross-Covariance, CCV* section. These pairwise calculations left us with 10,000 raw CCVs for each stimulus version (n = 10) per stimulus class (n = 5), and for each firing rate threshold (n = 8), or 4,000,000 raw CCV calculations in total.

We then calculated an average CCV function per pairwise comparison by averaging the CCVs over the stimulus versions within a stimulus class. These calculations left us with 10,000 mean CCV functions per stimulus class per firing rate threshold.

8.4.3 Population Analysis: Pooled Spike Train Cross Covariances

In order to compare our simulated results with our auditory forebrain results, we symmetrized our simulated mean CCV calculations as mentioned above in section *Pooled Spike Train Cross Covariances*. We then calculated mean and median network CCVs by averaging the 10,000 mean CCVs for each stimulus class and firing rate threshold. Mean and median network CCVs were pooled over all 11 birds and averaged to achieve an overall population CCV for each stimulus class and firing rate threshold.

Part III

Results

CH 5

Chapter 5

Results: Neural Correlations in Single Pairs

1 Neural Correlation in an Exemplary Pair

The following section provides an example of the correlations that we calculated for neuron pairs included in our analysis. The raw spiking responses for two auditory Field L neurons recorded simultaneously in response to BOS are shown in Figure 5.1.

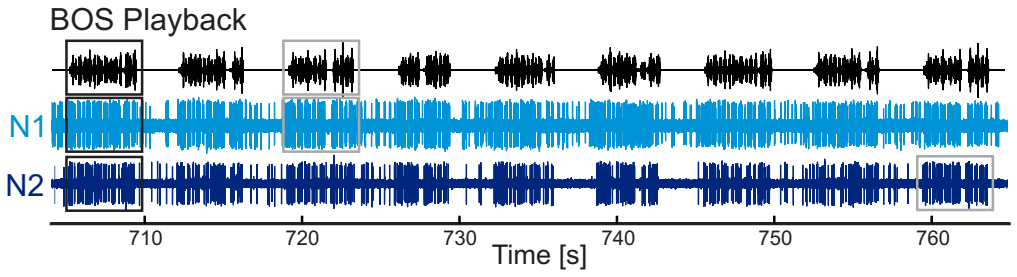

FIGURE 5.1: **Neural activity for two auditory neurons in response to BOS.** [g18r202] Black boxes indicate the responses used in the BOS (Bout) CCV calculation for BOS version 1. Gray box indicates an example of the shuffling used in the trial-shuffled CCV calculation. Specifically, the responses of Neuron 1 to BOS version 3 (N1; gray box) were compared to the responses of Neuron 2 to the next identical repetition of BOS version 3 (N2; gray box).

The spontaneous response dynamics that occurred during the silent periods between the BOS stimuli for the neuron pair are depicted as raster plots in Figure 5.2A (Neuron 1 (N1), left side; Neuron 2 (N2), right side). Figure 5.2B displays the z-normalized spike count responses for N1 and N2 during 1.5 s of spontaneous firing.

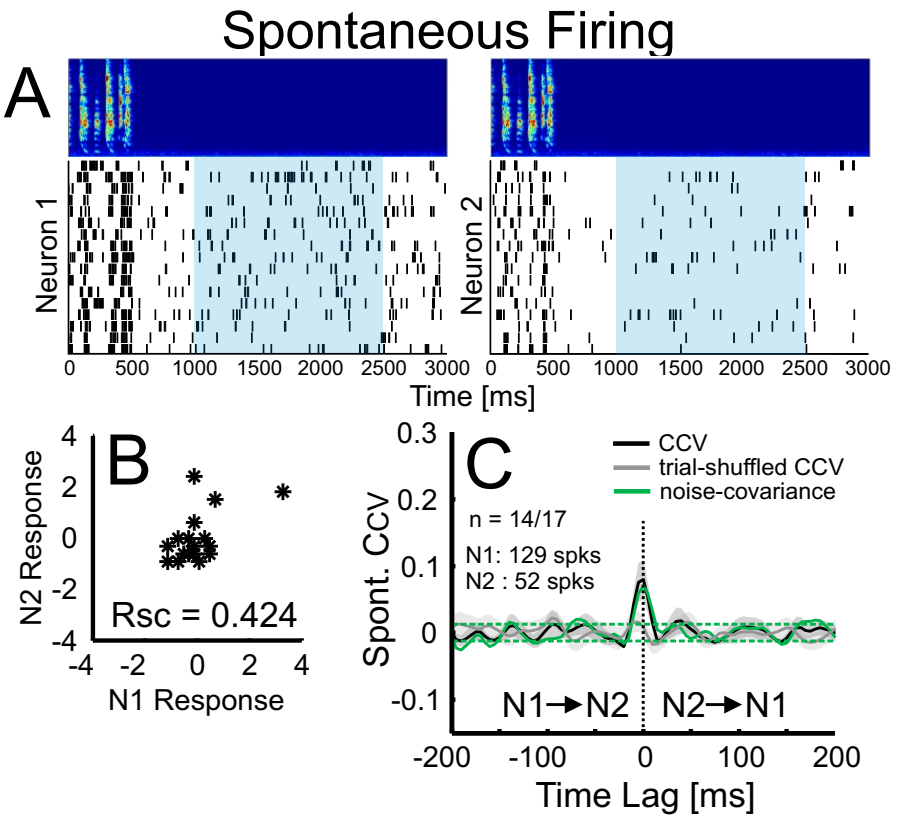

FIGURE 5.2: **Response dynamics and correlations during spontaneous firing.** [g18r2 02] **A)** Spike raster plots depict the spiking for neuron 1 (N1, left side) and neuron 2 (N2, right side) during silent periods in between the BOS stimuli (spectrogram, top). Blue shading indicates the 1.5 s period from which the spike count correlation was calculated. **B)** Z-normalized spike count responses of N1 and N2 during 1.5 s of spontaneous firing. The spike count correlation during spontaneous firing for this neuron pair was $R_{SC} = 0.42$ (not significant, $p = 0.089$). **C)** The cross-covariance function (CCV; black line) for this neuron pair during spontaneous firing exhibits a peak near zero time lag. Its trial-shuffled CCV (gray line) does not exhibit the same peak, but the noise-covariance (green line) does. Dark and light gray shading around the black and gray lines indicate the SEM. Green dotted line indicate the 99% confidence interval for the noise-covariance. N1 → N2 indicates that spikes from N1 lead spikes from N2 in time. N2 → N1 indicates that spikes from N2 lead spikes from N1. The number of trials (n = 14) for which spikes were present are indicated as a fraction of the total number of trials (17). The number of spikes (spks) for both neurons is indicated beneath. Spectrograms display frequencies from 0 to 8 kHz.

Figure 5.2C shows the CCV functions during spontaneous firing (Spont). The CCV (black line) has a large peak that is centered at zero time lag, indicating that spontaneously, this neuron pair fires near-synchronously at short time lags. The CCV peak is much higher than the trial-shuffled CCV peak (gray line), suggesting that correlations do not arise from similar auditory-induced reverberations in these two cells, but from correlations in the shared intrinsic noise. Indeed, the noise-covariance (green line) also has a large peak that is centered at zero time lag, and overall, the shapes of the CCV and the noise-covariance are highly similar. Although this neuron pair did not have significantly correlated spike counts during silence, a large component of the synchronous

firing during silence is presumably due to shared pre-synaptic input that drives both neurons.

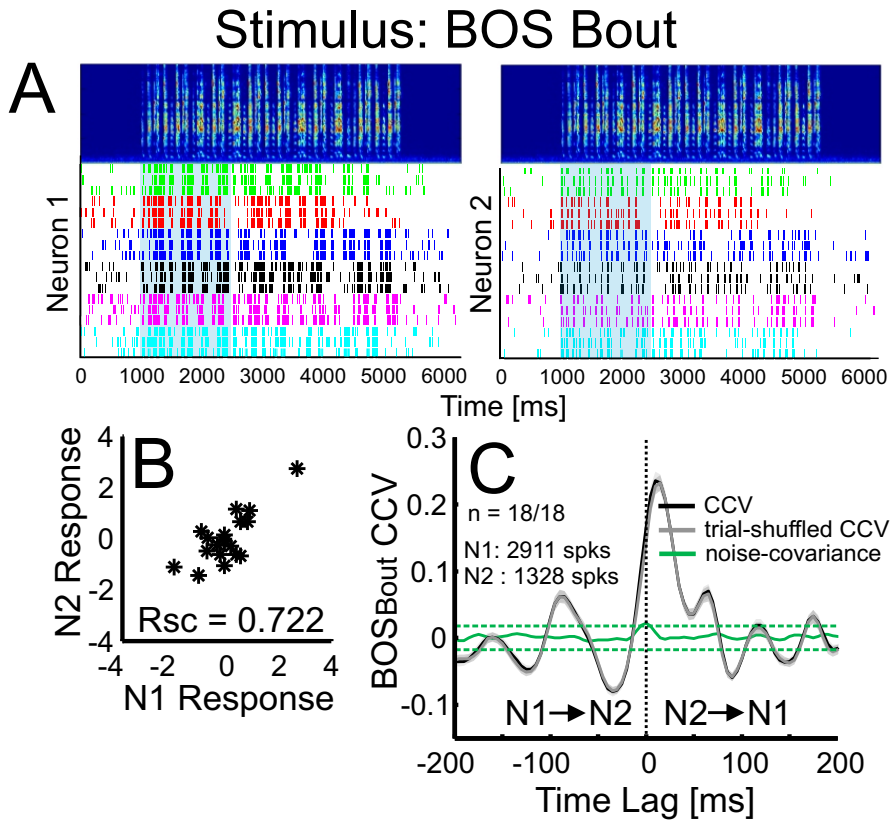

FIGURE 5.3: **Response dynamics and correlations during BOS (bout)-evoked firing.** [g18r2 02] Same figure conventions as described for Fig. 5.2. **A)** Spike raster plots depict the spiking responses for N1 and N2 during bouts of BOS stimulation (spectrogram, top). In the raster plot, the spikes are represented as lines, and the responses to each of the six different BOS versions is indicated by a different color. Both neurons displayed stereotyped spiking responses locked to the BOS stimulus. **B)** This neuron pair had a significant, positive spike count correlation (R_{SC} = 0.72, p <0.001) in response to BOS stimulation. **C)** The CCV for this neuron pair has a large asymmetric peak slightly offset from zero time lag. Its trial-shuffled CCV is almost identical to the CCV, indicating that features of the BOS stimulus drives correlated spiking patterns in this pair. In contrast, the noise-covariance is almost flat.

The BOS-evoked responses of the neuron pair are depicted in Figure 5.3A. Figure 5.3B shows the z-normalized responses of the neuron pair during 1.5 s of the BOS bouts. The spike count correlation during BOS was R_{SC} = 0.72, p <0.001, indicating that the spike counts in response to BOS are strongly and significantly correlated for this neuron pair.

Figure 5.3C displays CCV functions in response to bouts of BOS. The BOS_{Bout} CCV (black line) and its trial-shuffled CCV (gray line) are almost identical. Both have a large broad peak that extends from -30 ms to +50 ms time lags, indicating that the neurons tend to fire together within this time window. The peak of the BOS_{Bout} CCV and its trial-shuffled CCV are slightly offset from zero time lag and centered around +20 ms.

Because the trial-shuffled BOS_{Bout} CCV is so similar to the BOS_{Bout} CCV, it suggests that the BOS stimulus is driving the synchronous firing of this pair. Specifically, features of the BOS stimulus cause the spiking responses of N2 to precede the spiking responses of N1 by about 20 ms.

In addition to the large peak centered at +20 ms time lag, the CCV and trial-shuffled CCV also share many secondary peaks at long time lags, the significance of which will be discussed in the following paragraphs. The noise-covariance (green line) represents the part of the BOS_{Bout} CCV that is due correlations in shared input. Although a small peak is present, it does not cross the 99% confidence interval, suggesting that the shared pre-synaptic input that drives synchrony during spontaneous firing is reduced during auditory stimulation (see Fig. 5.2C for comparison).

We were curious whether we could extract more information about the correlations by examining the responses that occurred during the motif. The BOS-evoked responses aligned to the motif are depicted in Figure 5.4A. Figure 5.4B shows the z-normalized responses of the neuron pair during 200 ms of the BOS motif. The spike count correlation for the BOS motif was $R_{SC} = 0.24$, which was not significant ($p = 0.085$).

Figure 5.4C displays the CCV function in response to BOS motif, which closely resemble the CCV function in response to BOS bouts (see Fig. 5.3C). The BOS_{Motif} CCV (black line) and its trial-shuffled CCV (gray line) have a similar shape, suggesting that a large component of the CCV is driven by stimulus-induced correlations. As was true for the cross-covariance calculated on BOS bouts, peaks in the CCV and trial-shuffled CCV for BOS motifs are centered around +20 ms, suggesting that N2 leads N1 in firing. Unlike for the BOS_{Bout} noise-covariance, a significant peak is present in the BOS_{Motif} noise-covariance (green line), perhaps revealed by averaging over many more motif trials. Interestingly, the noise-covariance peak is slightly offset to positive time lags, suggesting that the activity of N2 might weakly - but directly - influence the firing responses of N1.

In addition to the large central peak present in both Fig. 5.3C and 5.4C, many smaller secondary peaks and troughs are also present at long time lags. Whereas peaks that are centered around zero time lag indicate correlations in pre-synaptic input, peaks and troughs at long time lags can indicate indirect influences between neurons. In the case of these two neurons, the peaks and troughs at long time lags are most likely driven by the spectral-temporal features of the BOS stimulus.

We were curious if the large asymmetric peaks in the trial-shuffled CCV could be explained by examining the STRFs of the two neurons. STRFs were estimated for each neuron from 40 CON songs (see Methods for details) and are presented in Figure 5.5.

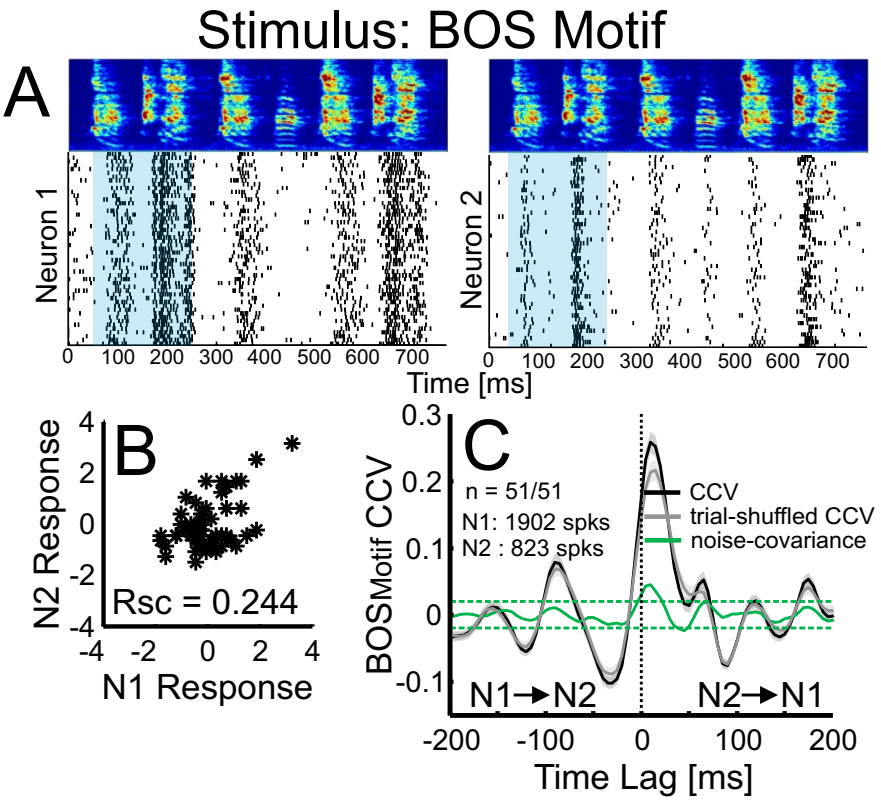

FIGURE 5.4: **Response dynamics and correlations during BOS (motif)-evoked firing.** [g18r2 02] Same figure conventions as described for Fig. 5.2. **A)** Spike raster plots depict the spiking for N1 and N2 during the BOS motif (spectrogram, top). In the raster plot, each spike is indicated by a black line. Blue shading indicates the 200 ms period from which spike count correlation was calculated. **B)** The spike count correlation during 200 ms of the BOS motif was positive but not significant (R_{SC} = 0.24, p = 0.085). **C)** The CCV, trial-shuffled CCV, and noise-covariance function calculated on the spiking responses during the BOS motif were nearly identical to those calculated for this neuron pair during the BOS bout. The CCV has a large asymmetric peak slightly offset from zero time lag. Its trial-shuffled CCV is almost identical to the CCV, indicating that the BOS stimulus drives the correlated spike timing of this pair. The noise-covariance is nearly flat, but has a small peak that passes the 99% confidence interval at slightly positive time lags, suggesting that spiking activity from N2 may influence the spiking of N1. This peak is present for the BOS motif case since these functions were averaged over more motif repetitions compared to bout repetitions.

We obtained high predictive CC values for both neuron pairs, indicating that both neurons respond linearly to auditory stimulation with birdsong. Features of both STRFs indicate that these neurons are receptive to a broad range of frequencies and would be good at detecting sound onsets. These STRFs closely resemble the "Broadband" neurons that have been previously reported for Field L neurons [Woolley et al., 2009].

The STRF similarity index (SI_{STRF}) for the two STRFs was SI_{STRF} = -0.25, indicating that the two STRFs are slightly anti-correlated in time. Indeed, when one examines the timing of the excitatory bands (Fig. 5.5, red bands), it is clear that the excitatory band for N2 precedes the excitatory band for N1 by approximately 20 ms, which is exactly the

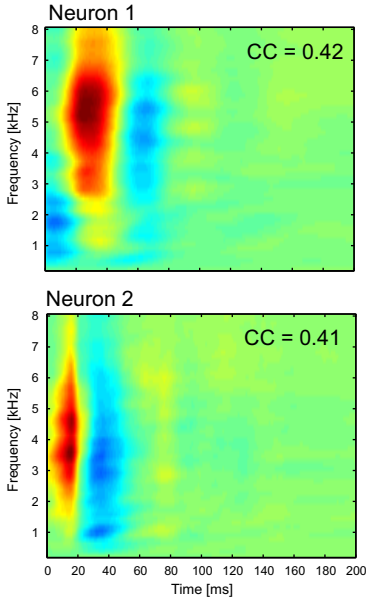

FIGURE 5.5: **STRFs for correlated onset neurons.** [g18r2 02] STRF estimated from CON songs for Neuron 1 (top) and Neuron 2 (bottom). Both neurons show large excitatory side bands (red areas) for a broad frequency range and inhibitory side bands (blue areas) that follow in time, indicating that both of these neurons would be good at detecting sound onsets. Both neurons had a high CC value greater than 0.4, indicating that the neurons respond predictably (and linearly) to auditory stimulation. SI_{STRF} = -0.25 for the two STRFs, indicating that the STRFS are slightly anti-correlated in time.

offset observed for the trial-shuffled CCV peak. By examining the STRFs for the two neurons, we see that much of the synchronous firing can be explained by the spectral and temporal feature tuning of the neurons.

1.1 Recapitulation

In order to provide an example for the correlation calculations that we use throughout the rest of the results section, we examined the correlated responses of two auditory Field L neurons in detail. Spontaneously the neurons share correlated pre-synaptic input, indicated by a significant peak in the noise-covariance (Fig. 5.2C) that is centered around zero time lag. During auditory stimulation, however, this correlated pre-synaptic input is modulated and reduced by the specific spectro-temporal tuning of the individual Field L neurons. By examining the STRFs of both neurons, we see that the excitatory side bands are anti-correlated in time, and that Neuron 2 begins integrating auditory information about 20 ms before Neuron 1 (Fig. 5.5). The effects of this auditory tuning are evident in both the trial-shuffled BOS_{Bout} and BOS_{Motif} CCV functions (Fig. 5.3C and Fig. 5.4C), which are characterized by an asymmetric peak that is centered around +20 ms. Additional secondary peaks at long time lags are also evident, and most

likely indicate responses to patterns of sound onsets in the auditory stimulus that drove synchronous firing.

Significant spike count correlations were also present for this neuron pair during BOS_{Bout} stimulation, indicating that the two neurons also share significant trial-to-trial variability during BOS stimulation. This spike count correlation was significant during the BOS_{Bout} stimulation, but not during spontaneous firing, suggesting that correlated trial-to-trial variability only occurs in the presence of auditory stimulation. However, because the BOS_{Motif} R_{SC} was not significant, it suggests that the correlated trial-to-trial variability is present only when spikes are summed over long time windows for this neuron pair.

2 Neural Correlation in Spatially Nearby Neuron Pairs

It has been widely reported that cortical neurons that are spatially near each other are generally highly correlated [Constantinidis and Goldman-Rakic, 2002; Lee et al., 1998; Smith and Kohn, 2008] and have similar tuning properties [Averbeck and Lee, 2003; Cohen and Newsome, 2008; Kohn and Smith, 2005; Zohary et al., 1994].

We recorded from 17 pairs of neurons with the same electrode. These spike were usually easy to sort due to large differences in spike waveform shape and height (Fig. 5.6).

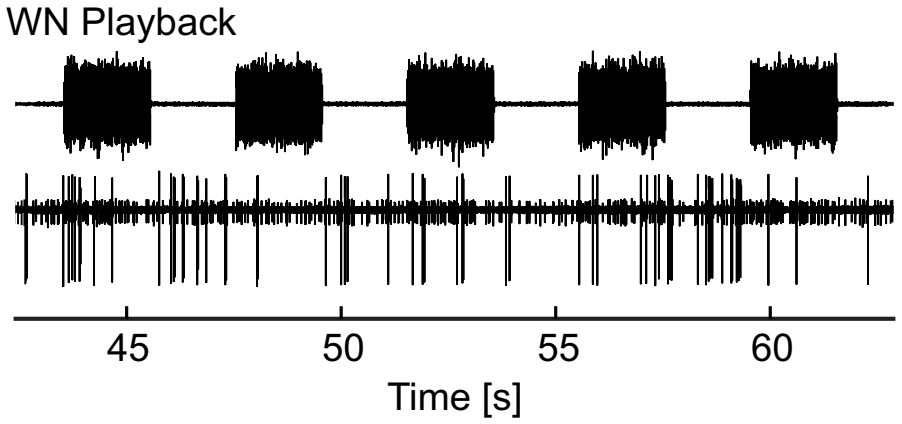

FIGURE 5.6: **Raw spiking responses of two neurons recorded on the same electrode-WN.** [y5y10 07] Top, WN playback; bottom, raw spiking responses from neuron pair. Spiking responses were easily sortable given the large difference in spike height for this neuron pair.

2.1 Few Nearby Neurons Have Significant R_{SC} Values ($p < 0.05$)

We probed the pairs of neurons with a wide range of auditory stimulation, including WN, Stacks, REV, CON, BOS, and qBOS. Of all the neuron pairs recorded on the same electrode, very few had significant spike count correlations ($p < 0.05$) in response to auditory stimulation (Fig. 5.7, open circles). Interestingly, the significant R_{SC} values were stimulus-dependent in the sense that while one neuron pair shared significant trial-to-trial variability for one stimulus, it was not necessarily shared for other stimuli, suggesting that different stimuli recruit different pre-synaptic pools of neurons and differentially affect the correlations that exist between neuron pairs.

In response to BOS playback, 4 neurons had significant spike count correlations (Fig. 5.7, BOS, open circles). Of these four neuron pairs, 2 were strongly positively correlated, whereas 2 were strongly negatively correlated. BOS was the only stimulus to evoke significantly negative correlations (qBOS also evoked more negative correlations

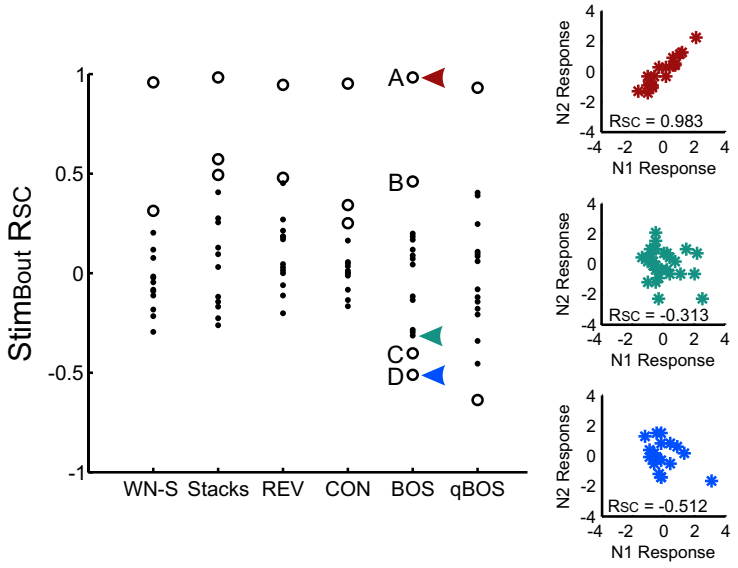

FIGURE 5.7: **Spike count correlations for nearby neurons during auditory stimulation.** Spike count correlations were calculated for spiking responses recorded on the same electrode in response to 1.5 s of stimulus playback. Significant R_{SC} values ($p < 0.05$) are depicted as open circles; other R_{SC} values ($p > 0.05$) are depicted as dots. Generally, few nearby neuron pairs were significantly correlated during auditory stimulation. In response to BOS playback, arrows indicate examples of spike count correlations that are significant and positive (red arrow), non-significant and negative (aqua arrow), and significant and negative (blue arrow). The z-normalized responses of these neuron pairs are indicated in color-coded panels on the right side of the figure. The letters, A, B, C, and D identify pairs that were significantly correlated during BOS playback; the CCV functions for each of these neuron pairs are depicted in Figure 5.8A-D.

from spatially near pairs). Significant negative spike count correlations for nearby neurons have not been reported for cortical neurons, which generally share weak - but positive - noise correlations [Cohen and Newsome, 2008; Kohn and Smith, 2005; Zohary et al., 1994].

We were curious whether the shared trial-to-trial variability observed in some neurons pairs was reflected in the synchronous firing of the neurons. We examined the cross-covariances for neurons that shared significant spike count correlations. Figure 5.8 depicts the BOS_{Bout} CCV functions for the 4 neuron pairs that had significant spike count correlations in response to BOS (Fig. 5.7), where each letter in Figure 5.7 refers to the correspondingly labeled panel in Figure 5.8.

Of the 4 neurons that had significant spike count correlations for BOS, only one pair shared pre-synaptic input (Fig. 5.8A; noise-covariance, green line). This neuron pair was significantly positively correlated ($R_{SC} = 0.983$, $p < 0.000001$) and will be discussed in greater detail in the following section. The other 3 neurons (Fig. 5.8B-D) did not fire synchronously to the BOS stimulus. For two neuron pairs that shared positive

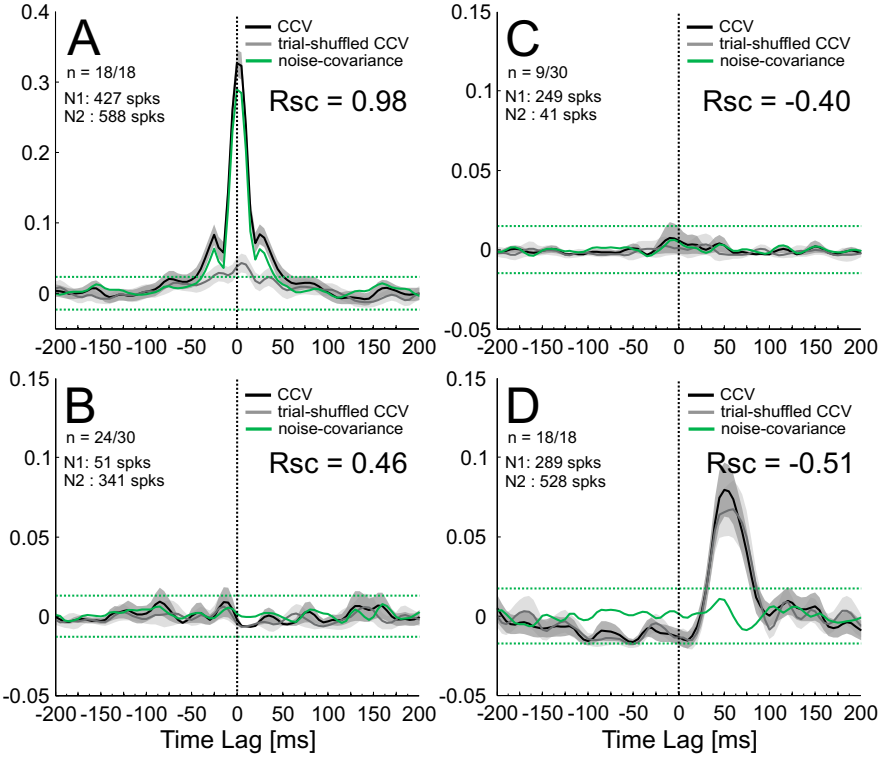

FIGURE 5.8: **Significant trial-to-trial variability does not reflect spiking synchrony.** Same figure conventions as described for Fig. 5.2. **A-B)** Each panel displays the BOS_{Bout} CCV functions for the correspondingly labeled neuron pair in Figure 5.7: Except for the neuron pair in **A**, nearby neuron pairs that had significant spike count correlations did not fire synchronously. See text for details.

(Fig. 5.8B) and negative (Fig. 5.8C) spike count correlations, there was no discernible correlations in spike times, indicated by the flat CCV functions. Another negatively correlated pair (Fig. 5.8D; R_{SC} = -0.51) had large asymmetrical peaks in the CCV and trial-shuffled CCV that were centered around +50 ms, suggesting that the auditory stimulus drove spiking in one neuron 50 ms before the other. Because the noise-covariance peak does not cross the confidence interval (Fig. 5.8D; green dotted line), the synchronous firing at long time lags is most likely not the result of indirect connections between the neurons.

2.2 Some Nearby Neurons Fire Synchronously

Although neurons that shared significant spike count correlations generally did not fire synchronously, several neurons that did not have significant spike count correlations ($p > 0.05$) did exhibit synchronous firing. The raw spiking responses of one such neuron pair are depicted in Figure 5.9. This neuron pair had a significant and positive spike count correlation in response to CON playback (R_{SC} = 0.25, $p < 0.05$), but not for any other stimulus.

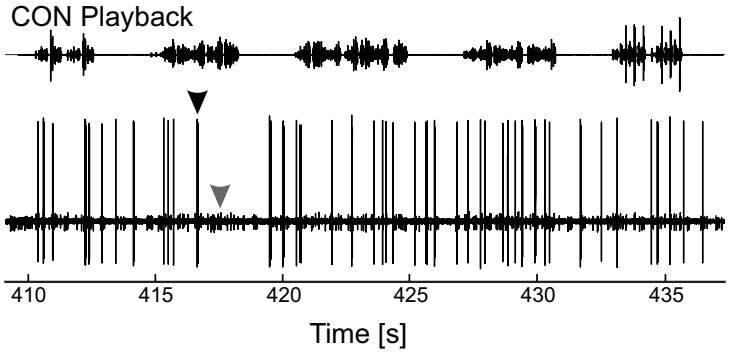

FIGURE 5.9: **Neural responses of two neurons recorded on the same electrode-CON.** [g4r4 05] Top, CON playback; bottom, raw spiking responses from neuron pair. Spiking responses for Neuron 1 (black arrow) and Neuron 2 (gray arrow) were easily sortable given the large difference in spike height.

We examined the spontaneous and stimulus-evoked CCV functions for this neuron pair in response to WN-S, Stacks, REV, BOS, qBOS, and WN-E. The spontaneous CCV functions are depicted in Figure 5.10, left column, and the stimulus-evoked CCV functions are depicted in Figure 5.10, right column.

During both spontaneous and stimulus-evoked firing, the noise-covariance functions for this neuron pair had a significant asymmetrical peak offset to the left of zero time lag (Fig. 5.10). The peaks were broad, ranging from almost -100 to zero time lag.

Asymmetrical peaks at long time lag indicate that the spike time patterns for the two neurons are not independent and suggest that the firing of one neuron indirectly influences the firing of the other neuron [Perkel et al., 1967b]. In this case, there is a high probability that a small spike from neuron 2 will occur after a large spike from neuron 1. Because the spontaneous and stimulus-evoked noise-covariance peaks are so broad and the coincident time lags so long, it is unlikely that these two neurons are directly connected. Rather, it suggests the involvement of at least one other (unrecorded) neuron that could modulate the spiking activity of neuron 2.

This neuron pair was also differentially modulated by the the auditory stimulus class. The noise-covariance peaks for WN-S and WN-E were similar, suggesting that these two neurons respond consistently to identical stimulus classes. Interestingly, the noise-covariance peak for the Stacks stimulus class was small, suggesting that Stacks did not elicit robust spiking responses from the neuron pair, and indeed, neuron 1 spiked much less in response to Stacks (48 spikes). Differences between the BOS noise-covariance and the qBOS noise-covariance are also apparent: whereas the BOS noise-covariance has a double-peak profile, the qBOS noise-covariance has a single broad peak. The difference in correlation structure between loud and quiet stimuli suggests that a member

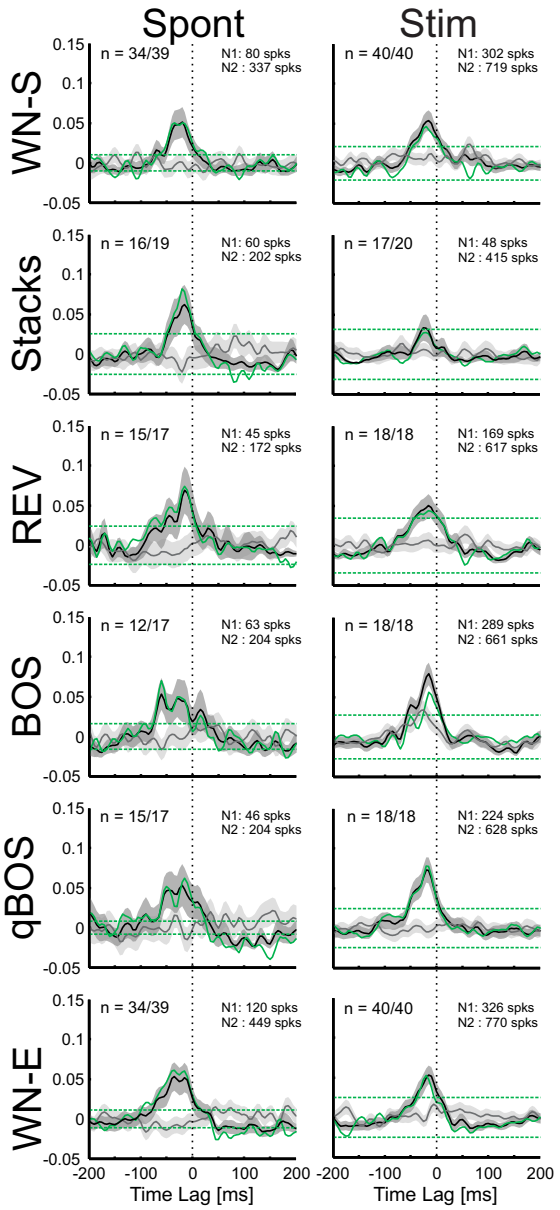

FIGURE 5.10: **CCV functions for neuron pair indicate broadly synchronous firing.** [g4r4 05] Panels display CCV functions calculated for different stimulus blocks: CCV (black line); trial-shuffled CCV (gray line); noise-covariance (green line). The CCV and noise covariance for this neuron pair is characterized by a broad asymmetrical peak, indicating that spikes from neuron 1 are followed by spikes from neuron 2, and furthermore, that this correlated pattern in spike timing most likely involves an indirect connection between the neurons. Dark and light gray shading around the black and gray lines indicate the SEM. Green dotted line indicate the 99% confidence interval for the noise-covariance. Dotted vertical line indicates 0 time lag. The number of trials for which spikes were present are indicated as a fraction of the total number of trials in the upper left corner of each panel. of the total number of trials. The number of spikes (spks) for both neurons are indicated in the upper right corner. See text for details.

of this neuron pair is sensitive to the intensity of the playback, since roughly similar numbers of spikes are elicited by both stimuli.

We examined the cross-covariance spike time correlations of another pair of neurons recorded on the same electrode. Raw spiking responses to Stacks playbacks are presented in Figure 5.11. This pair shared significant trial-to-trial variability for Stacks (R_{SC} = 0.571, $p < 0.01$) and for REV (R_{SC} = 0.479, $p < 0.05$), but not for any other stimulus class.

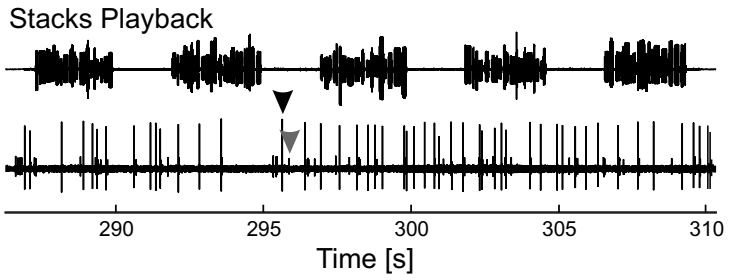

FIGURE 5.11: **Neural responses of two neurons recorded on the same electrode-Stacks** [o5r3 01] Top, Stacks playback; bottom, raw spiking responses from neuron pair. Spiking responses for Neuron 1 (black arrow) and Neuron 2 (gray arrow) were easily sortable given the large difference in spike height.

We examined the spontaneous and stimulus-evoked CCV functions for this neuron pair in response to WN-S, Stacks, REV, qBOS, and WN-E. The spontaneous CCV functions are depicted in Figure 5.12, left column, and the stimulus-evoked CCV functions are depicted in Figure 5.12, right column. Because there were generally fewer trials that contained spike from both neurons for this neuron pair, caution was taken in interpreting the noisier CCVs for this neuron pair. However, some characteristic features are obvious.

Unlike the previous neuron pair example, the noise-covariance for this neuron pair was characterized by a large and narrow central peak that was centered around zero time lag (Fig. 5.12). This peak indicates that the timing of the spiking responses of these two neurons are not independent, and most likely, both neurons receive shared pre-synaptic input.

The responses of this neuron pair are also differentially modulated by the stimulus class. WN-S and WN-E both evoked sharp peaks that are centered at zero time lag, suggesting that features of this stimulus elicited a response in the neuron whose input drove the neuron pair.

In contrast, when we examine the CCV functions for the two stimulus classes that evoked significant R_{SC} correlations (Stacks and REV) an interesting interaction arose. The Stacks stimulus elicited a broad, asymmetric CCV peak for the neuron pair, but also a substantial peak for the trial-shuffled CCV. This suggests that features of the Stacks stimulus were partially responsible for driving the responses of the neuron pair, and potentially interfered with any shared pre-synaptic input received by both neurons. The

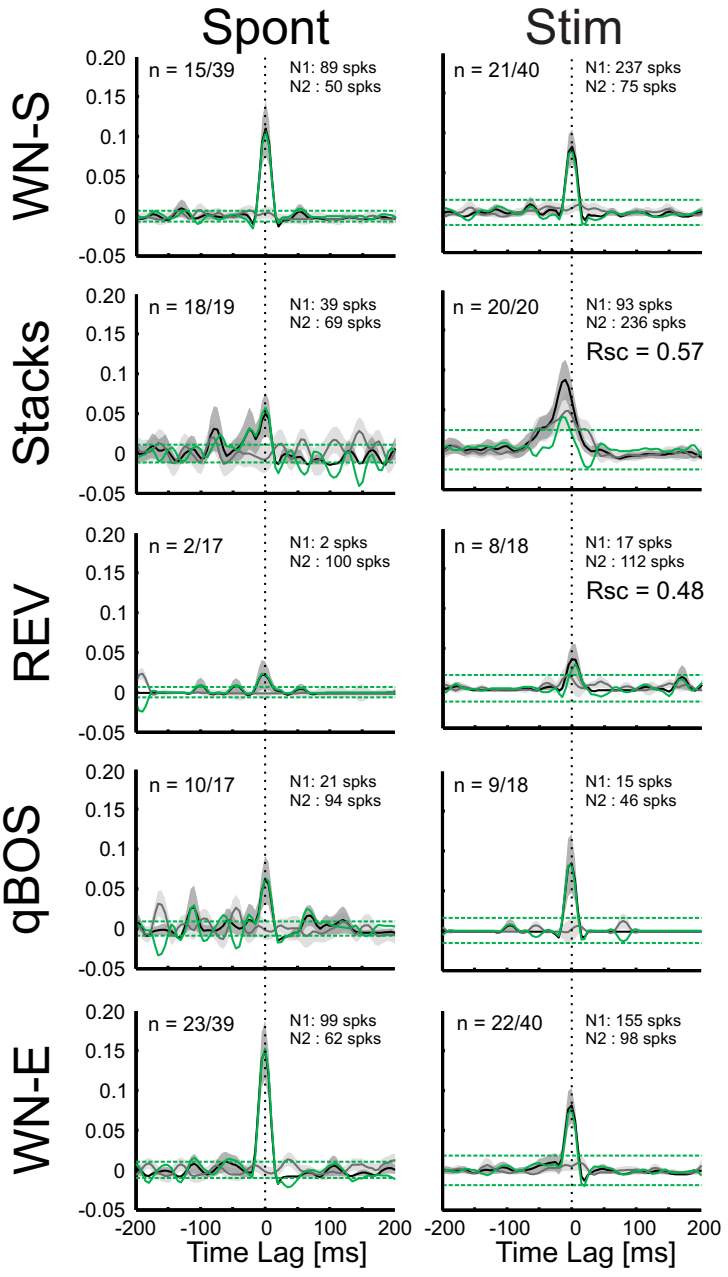

FIGURE 5.12: **CCV functions for a nearby neuron pair indicate tight synchronous firing.** [o5r3 01] Same figure conventions as described for Fig. 5.10. In contrast to the broadly synchronous firing present for the neuron pair in Fig. 5.10, the CCV and noise-covariance for this pair were characterized by a narrow peak centered around zero time lag, indicating that both neurons receive shared pre-synaptic input. Stacks and REV stimuli evoked significant spike count correlations for this neuron pair, and effects of stimulus tuning was also evident for correlations in spike timing. For both stimuli, the trial-shuffled CCV (gray line) had a peak that was absent for other stimulus classes, indicating that features of the stimuli drove spiking responses in addition to the shared pre-synaptic input.

peak for the Stacks noise-covariance is centered around a time lag of -10 ms, suggesting that the chance that a large spike from the Neuron 1 was followed by a small spike from Neuron 2 after about 10 ms was high. Indeed, when one examines the spiking responses

of the neuron pair during the Stacks playback (Fig. 5.11) there were several cases where this occurred.

A similar effect is also present for the REV stimulus. In this case, the noise-covariance peak is centered around a time lag of +10 ms, suggesting that the chance that spike from the neuron with the small spike waveform was followed after about 10 ms by a spike from the neuron with the larger waveform was large. As for the Stacks playback, the peak in the noise-covariance was small, suggesting that the spiking output due to stimulus tuning interfered with the spiking due to shared presynaptic input. Such differential effects of the stimulus highlight the neuronal interactions that shape the spiking output of these neurons.

A final example of a highly correlated neuron pair that was recorded on the same channel presents an interesting anomaly. The raw spiking responses of this neuron pair are displayed in Figure 5.13. The two neurons fire almost identically, except for cases when the neuron with the smaller spike waveform (neuron 2) fires alone (see Figure 5.13B,C; gray arrows). The fact that spikes from neuron 2 precede the larger spikes from neuron 1 with variable lengths of time and occasionally appears in isolation suggests that these two neurons are distinct.

Figure 5.14A displays the rasterplots of the two neurons in Figure 5.13 in response to BOS playback. Neuron 2 tends to fire more than neuron 1, as observed from the overlapping PSTHs in Figure 5.14B. We estimated the STRFs of these neurons using 75 CON stimuli and observed almost identical receptive fields from these neurons. The STRFS for both neurons were highly similar (STRF SI = 0.9732). However, neither neuron had a very high CC value, indicating that these neurons respond somewhat non-linearly to the CON auditory stimulus [Theunissen et al., 2000].

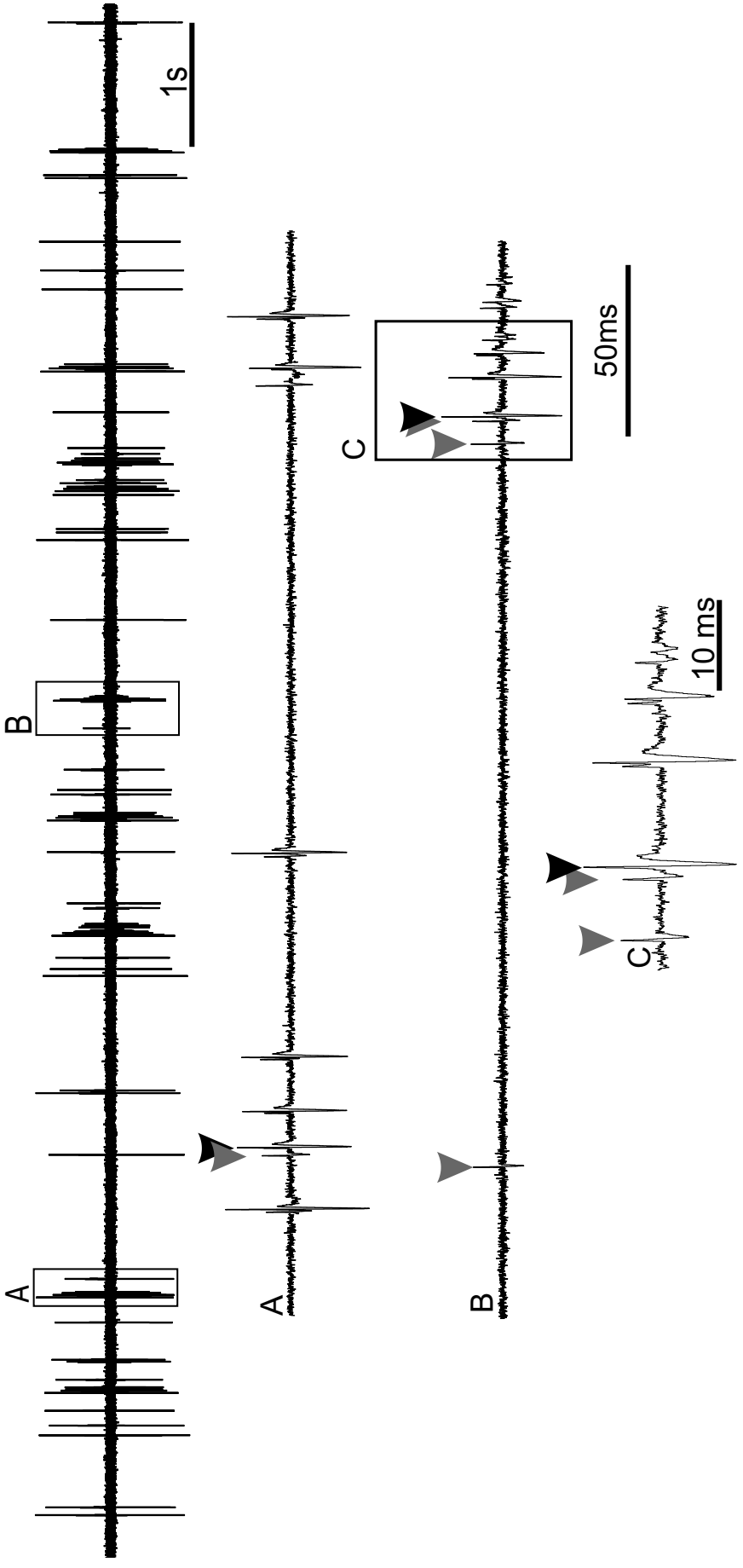

FIGURE 5.13: **Responses of a highly correlated neuron pair recorded on one electrode.** [k18r4 02] Upper panel depicts the spiking responses of the two neurons at a large time scale. The boxed areas **A** and **B** are magnified in the underlying panels. Black arrows highlight the large spike from one neurons, and gray arrows highlight the small spikes from the other neuron. Small spikes typically precede the larger spikes with variable time duration and occasionally appear in isolation. In **C**, once can see how the spikes often overlap to form a spike doublet.

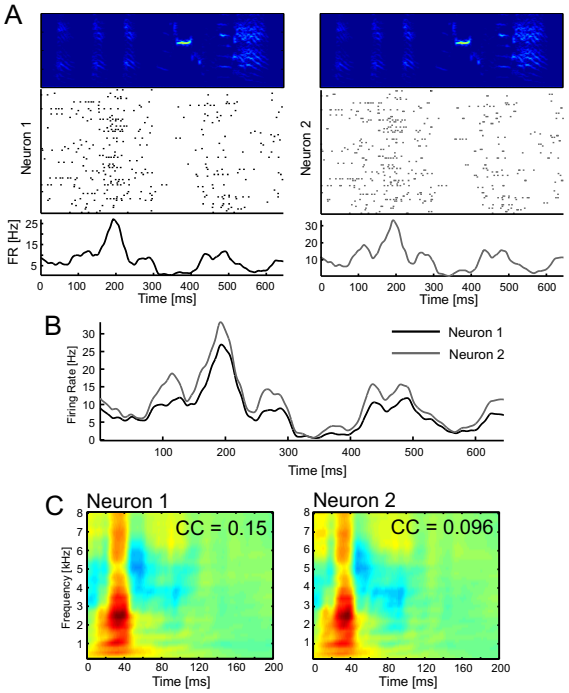

FIGURE 5.14: **Response dynamics and STRFs for a highly correlated neuron pair.** [k18r4 02] **A)** Rasterplot of neuron 1 (left panel) and neuron 2 (right panel) in response to BOS playback, aligned to the BOS motif (upper spectrograms). Lower panel displays the smoothed firing rate curve for both neurons. **B)** Overlaid firing rate curves for both neurons highlight the subtle differences in the firing patterns of the two neurons. Neuron 2 fires more robustly in response to BOS than neuron 1. **C)** STRFs estimated from CON stimuli. STRFs are remarkable similar for the two neurons ($SI_{STRF} = 0.9732$). However, low CC scores indicate that the neurons respond non-linearly to the auditory stimulus.

The almost identical firing patterns of these two neurons confers highly positive and significant spike count correlations (see Figure 5.7, top-most open circles) and highly synchronous firing (Fig. 5.15). We examined the spontaneous and stimulus-evoked CCV functions for this neuron pair in response to WN-S, Stacks, REV, BOS, qBOS, and WN-E. The spontaneous CCV functions are depicted in Figure 5.15, left column, and the stimulus-evoked CCV functions are depicted in Figure 5.15, right column. A CCV example for this neuron pair in response to BOS appeared in Fig. 5.8A.

Unlike the CCV functions that we have observed for other pairs of neurons recorded on the same electrode, this neuron pair is highly stable across conditions (Fig. 5.15.) Except for the cases of the spontaneous trials for BOS and qBOS, for which there were generally fewer spontaneous trials that contained spikes, the amplitude of the peaks are nearly identical for both the spontaneous and stimulus-evoked conditions. The noise-covariance peak for all conditions is centered around zero time lag, suggesting that these two neurons receive shared pre-synaptic input. Furthermore, this input drives the neurons almost identically during spontaneous firing and stimulus-evoked firing, suggesting a truly synchronous correlation between these two neurons [Eggermont et al., 1983].

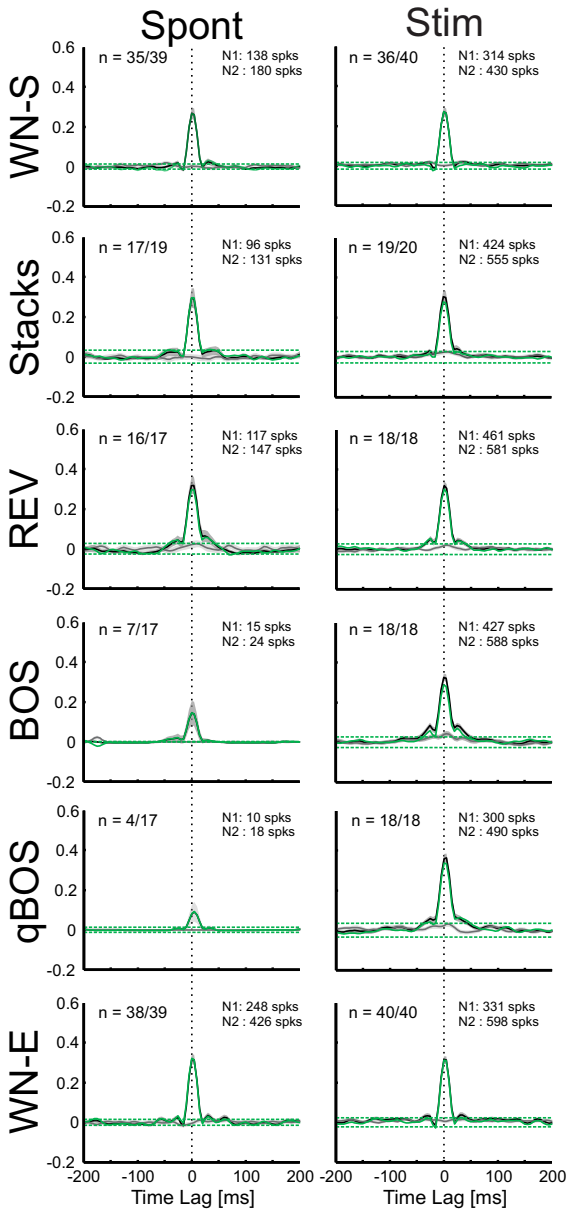

FIGURE 5.15: **CCV functions for neuron pair indicate stimulus-independent synchrony.** [k18r4 02] Same figure conventions as described for Fig. 5.10. CCV and noise-covariance functions are highly similar across spontaneous and stimulus conditions. The narrow peak centered around zero time lag indicate that both neurons receive shared presynaptic input. Because the peaks are nearly identical across stimulus conditions, it is likely that the shared presynaptic input is stimulus-independent.

2.3 Recapitulation

We examined spike count correlations and spike time correlations for neuron pairs that are spatially near each other and whose spiking responses were both recorded on the same electrode. In the mammalian cortex, nearby neurons typically share positive spike count correlations [Constantinidis and Goldman-Rakic, 2002; Lee et al., 1998; Smith and Kohn, 2008] and have similar tuning properties [Averbeck and Lee, 2003; Cohen

and Newsome, 2008; Kohn and Smith, 2005; Zohary et al., 1994]. We were curious whether the same responses could be found in the avian brain.

We recorded from 17 neuron pairs from the same electrode. Of these, very few had significant spike count correlation values ($p < 0.05$; Fig. 5.7). Furthermore, shared trial-to-trial variability was stimulus-specific. That is, several neurons showed significantly correlated trial-to-trial variability for a subset of the stimuli, but not for all of the stimulus classes. Additionally, some spatially nearby neuron pairs were significantly and strongly negatively correlated for BOS and qBOS stimuli. Negative R_{SC} values have not been reported for nearby neurons, and this finding may functional differences that exist between neural circuits in the mammalian cortex and the avian brain. Finally, neurons that shared significant trial-to-trial variability generally did not share synchronous firing patterns in time (Fig. 5.8).

Several neuron pairs that did not have significant spike count correlations did have significant peaks in the CCV functions. For one pair, these peaks indicated indirect activation of one neuron by another (Fig. 5.10), evidenced by broad and asymmetric noise-covariance peaks. For another pair, shared pre-synaptic input from a common source was modulated by the auditory stimulus and had differential effects for Stacks and REV (Fig. 5.12). A final neuron pair had almost identical firing patterns, and shared large trial-to-trial variability and fired synchronously in a stimulus-independent manner (Fig. 5.15). The response dynamics and STRFs for this neuron pair were highly similar (Fig. 5.14), and although spikes from one neuron did occur in isolation, we cannot rule out that these neurons may be somehow functionally coupled.

Overall, these results highlight the fact that, in the avian brain, nearby neurons do not necessarily have similar tuning features or share correlated input. While some nearby neurons pairs did share trial-to-trial variability, these neurons were in the minority. Similarly, while some neuron pairs did have correlated patterns in their spike times, many nearby neurons did not. These findings suggest that local neural circuits in the avian auditory forebrain may involve much more heterogeneously tuned neurons than observed in clusters of cortical circuits.

3 Neural Correlation in an Ensemble of Field L Onset Neurons

We examined the pairwise spike time cross correlations and the spike count correlations of an ensemble of three auditory neurons recorded in Field L. Since many neurons in Field L have broad spectral tuning [Woolley et al., 2009] we were curious to examine the correlated activity shared between neurons with similar broadband tuning. The putative reconstruction of the recording location of the neurons is depicted in Figure 5.16. Neuron 2 (N2) was recorded in L2b, and neurons 3 (N3) and 5 (N5) were recorded in L2a.

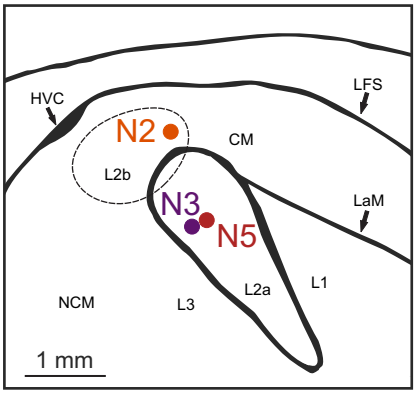

FIGURE 5.16: **Putative recording locations for an ensemble of Field L neurons.** [r6y12 02] Section represents the saggital plane 1.2 mm from the midline. Neuron 2 (N2; orange circle) was located in L2b. Neuron 3 (N3; purple circle) was located in L2a. Neuron 5 (N5; red circle) was located in L2a. The dotted line indicates the region that we considered to be L2b. Abbreviations: L2a, L2b, L1, L3: subregions of Field L; CM: caudal mesopallium; NCM: nidopallium caudal medial; HVC: Higher Vocal Center; LFS: Lamina frontalis superior; LaM: Lamina mesopallialis.

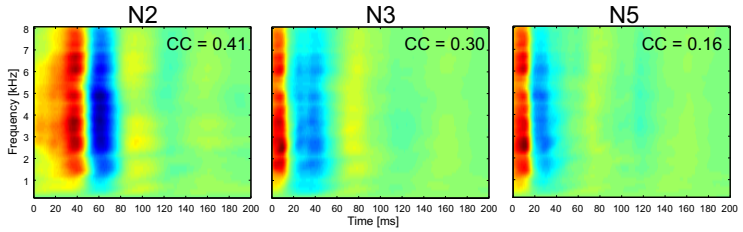

FIGURE 5.17: **STRFs indicate broadband tuning for 3 Field L neurons.** [r6y12 02] STRFs estimated from 100 CON songs for neuron 2 (N2; left panel), neuron 3 (N3; middle panel), and neuron 5 (N5; right panel). The STRFs indicate that all 3 neurons are similarly tuned to a broad range of frequencies, but differ terms of their response latencies. CC values indicate the correlation coefficient for the STRF prediction validation.

We estimated the STRFs of these neurons in response to 100 unique CON songs (see Methods for details). The STRFs for each neuron are depicted in Figure 5.17. All three neurons have similar STRFs that indicate that each of these neurons respond to a broad range of frequencies, but differ in terms of their response latencies. These STRFs match the "broadband" neurons commonly found in Field L [Woolley et al., 2009].

We examined the spiking activity of this ensemble of neurons in response to WN-S, REV, and BOS. The spiking dynamics of the three neurons are depicted in Figure 5.18A. The integrated firing rates for each neuron in response to 3 versions of REV and 3 versions of BOS are depicted in Figure 5.18B.

Each of the Field L neurons responds to the onsets of sounds with high reliability. N2 and N5 responded to the WN-S stimuli with a precisely-timed spike that marked the onset of the stimulus; N3 also responded to the onset of the WN-S stimulus, but with a tonic response that lasted for the duration of the stimulus. The neuron responses to REV and BOS stimuli were also characterized by highly-reliable spiking responses to certain features of the stimulus. Interestingly, all three neurons were well-driven by the introductory notes to the BOS song, in addition to syllables of the motif.

We began by examining the correlated spiking activity for N2 and N3 in response to WN, REV, and BOS (Fig. 5.19A). For this neuron pair, the trial-shuffled CCV was remarkable similar to the CCV, even across stimulus classes, indicating that correlations in spike times are largely the effect of stimulus drive. Indeed, the noise-covariance functions for this neuron pair were flat across stimulus conditions and did not surpass the confidence interval, suggesting that the effects of stimulus drive, and not intrinsic noise, contribute to the peaks of the CCV function.

The CCV and trial-shuffled CCV peaks functions were highly different across stimulus conditions. For WN-S, the CCV and trial-shuffled CCV functions were characterized by two peaks offset from zero time lag by about 25 ms, suggesting that after a spike from one neuron, the other neuron spikes at a delay of around 25 ms. By examining the spiking raster plots of these two neurons (Fig. 5.18A) we observe that N3 fires a spike at the onset of the WN stimulus, whereas N2 has a slight delay in its latency of response. These latencies are also evident in the STRFs for both neurons (Fig. 5.19B, C). This firing pattern explains the large peak at positive time lags, indicative of a spike from N3 preceding a spike from N2.

However, N3 does not often fire only a single spike at the onset of the WN stimulus, but several spikes separated by short time delays. These secondary spikes elicited by N3 follow the single spikes elicited by N2 at the onset of the WN stimulus, and account for the peak at negative time lags, indicative of a spike from N2 preceding a spike from N3.

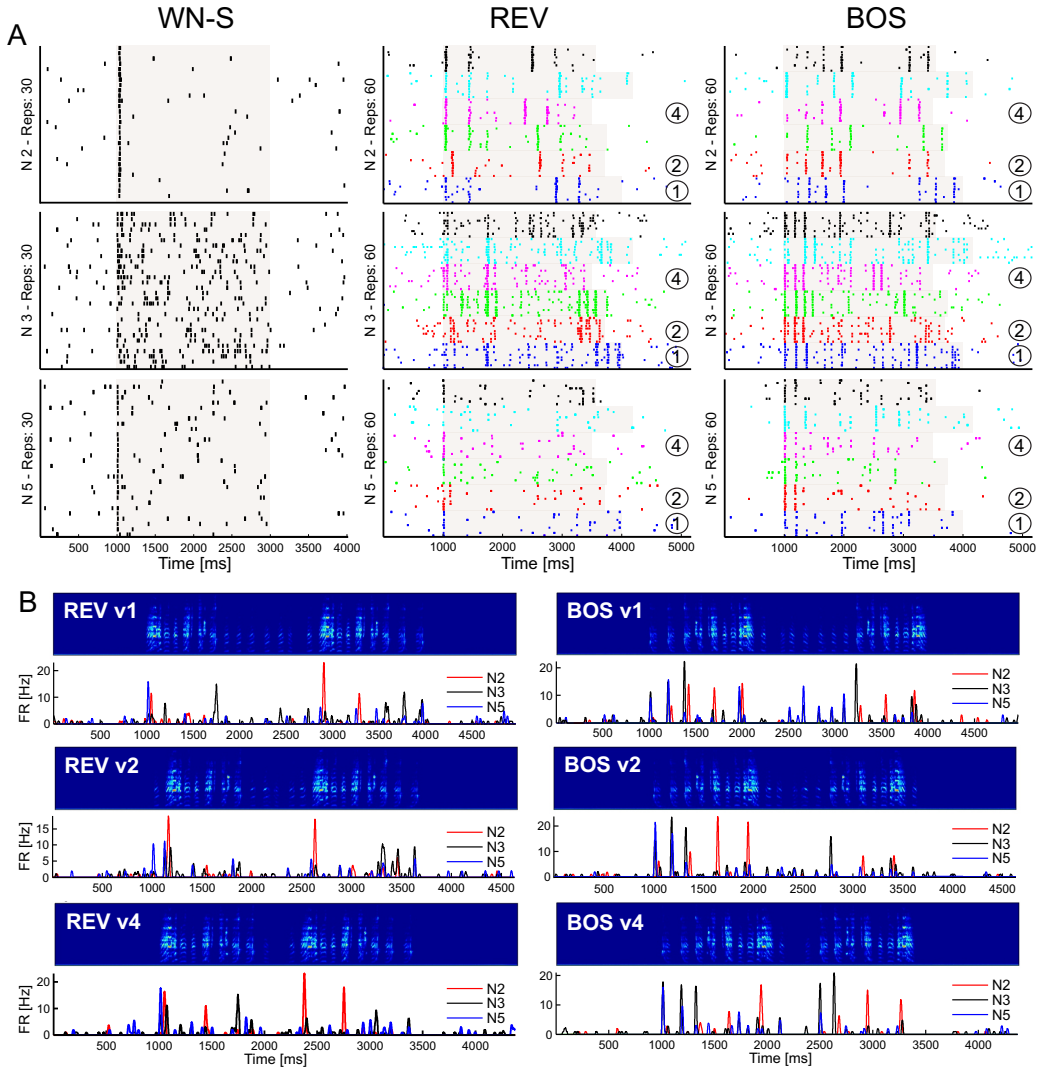

FIGURE 5.18: **Neural activity for 3 Field L neurons in response to auditory playback.** [r6y12 02] **A)** Raster plots depict the spiking responses of 3 neurons recorded simultaneously during WN-S (left column), REV (middle column), and BOS (right column). The top raster plot row depicts the responses to neuron 2 (N2), the middle row depicts the responses to neuron 3 (N3), and the bottom row depicts the responses to neuron 5 (N5). For raster plots, each spike is depicted as a line. For the BOS and REV stimuli, responses to each of the 6 unique song versions are represented as different colored lines. Shading indicates the overall stimulus duration for each stimulus versions. The circled numbers indicate the REV and BOS versions that are depicted in **B**. All three neurons display stereotyped responses that are locked to onset features of the stimuli. **B)** Integrated firing rate (FR) responses to 3 different versions of REV and BOS are overlaid for N2 (red line), N3 (black line) and N5 (blue line). Song versions correspond to the circled number in **A**. Neurons are robustly driven by the introductory notes present in the BOS songs, but not to reversed introductory notes that are present in REV. Firing rates are smoothed with the matlab "lowess" option using a smoothing window of 31 ms.

The probability of a spike from N3 preceding a spike from N2 was slightly higher than a spike from N2 leading a spike from N3, as indicated by the larger height of the peak located at positive time lags.

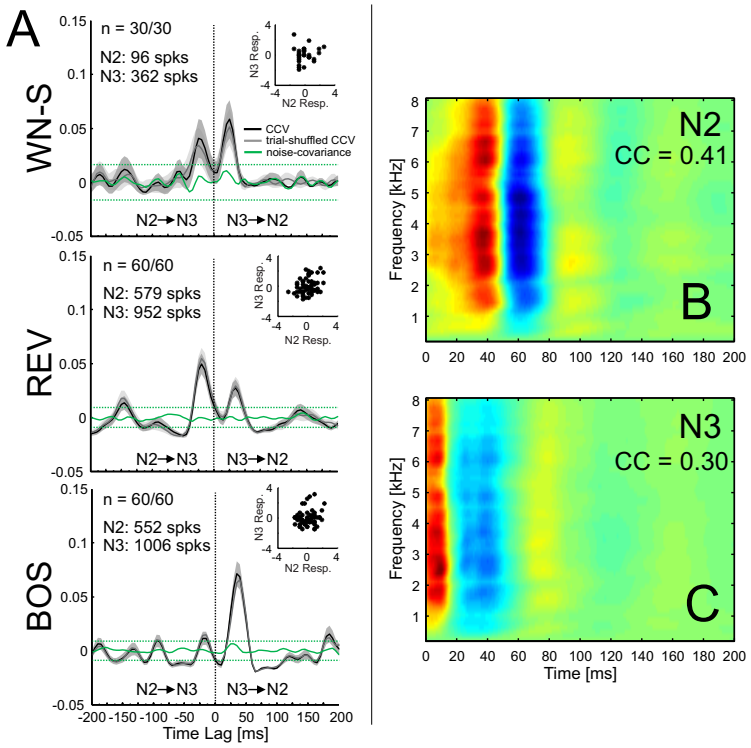

FIGURE 5.19: **Correlation dynamics for Field L neurons 2 and 3.** [r6y12 02] **A)** CCV (black line), trial-shuffled CCV (gray line), and noise-covariance (green line) functions calculated between N2 and N3 are depicted in response to WN-S, REV, and BOS. The CCV and trial-shuffled CCV were remarkably similar within a stimulus class, indicating that correlations in spike times arise from the stimulus tuning of the neuron pair. Across stimulus classes, the magnitude and timing of the peaks in the CCV and trial-shuffled CCV differed, highlighting the effect that temporally-modulated stimuli can have on correlated firing patterns of onset neurons. See text for further details. Dark and light gray shading indicate the SEM for the CCV and trial-shuffled CCV, respectively. Green dotted line indicates the 99% confidence interval for the noise-covariance. Dotted vertical line indicates 0 time lag. The number of trials for which spikes were present is indicated in the upper left corner of each panel. The number of spikes for both neurons are indicated in the upper right corner of each panel. Inset depicts the z-normalize spike count responses for each stimulus repetition for the neuron pair. **B, C)** STRFs calculated for N2 (**B**) and N3 (**C**) indicate that both neurons are tuned to a broad frequency range. CC values indicate the correlation coefficient for the STRF prediction validation.

Importantly, these neurons did not fire synchronously to the stimulus, as indicated by the trough at zero time lag. Although these two neurons tended to fire within 25 ms of each other, they did not have a significant spike count correlation for WN (R_{SC} = -0.0072, p = 0.97), most likely due to the fact that N2 fired relatively few spikes in response to the WN stimulus.

For REV and BOS, the CCV and trial-shuffled CCV functions were also characterized by two peaks offset from zero time lag, with one peak having a much larger magnitude than the other. In addition, several smaller secondary peaks were also present at long time lags, especially for the BOS stimulus. As observed for the WN stimulus, the CCV

and the trial-shuffled CCV functions for BOS and REV were nearly identical, indicating that the stimulus drives the correlated spike times.

The differences in CCV structure between REV and BOS stimuli highlight the importance that temporal patterns of frequency content can have on correlated firing patterns of onset neurons. REV and BOS have identical spectral content and only differ in terms of the temporal location of the energy spectra. However, we see that the temporal location of the spectral content can elicit widely different spike time correlation patterns. For example, the introductory notes of the BOS stimulus drive robust responses from both neurons N2 and N3, with N3 leading the firing of N2 (see Fig. 5.18B). This firing pattern accounts for the large and positive peak present at positive time lags in the BOS CCV and trial-shuffled CCV (Fig. 5.19A). However, the same introductory notes in the REV, located at the end of the stimulus instead of the beginning as in BOS, elicit greatly reduced responses, leading to a spike-time correlation patterns similar to what was observed for WN.

Spike count correlations calculated for 1.5 s of the REV and BOS stimuli were both significant, but were surprisingly different in terms of polarity. Spike counts were strongly positively correlated in response to REV (R_{SC} = 0.42, $p < 0.001$), but were significantly negatively correlated in response to BOS (R_{SC} = -0.26, $p < 0.05$).

For another neuron pair, N2 and N5, the spike time correlations across stimuli were remarkably similar (Fig. 5.20A). In response to WN-S, REV, and BOS, the CCV and trial-shuffled CCV were both dominated by a peak centered around 25 ms (WN-S) to 37 ms (REV and BOS). Because the trial-shuffled CCV was so similar to the CCV, it is highly likely that the correlated spike time patterns are the result of stimulus drive. Except for the case of the REV stimulus, for which the noise-covariance crossed the 99% confidence interval, intrinsic noise does not contribute significantly to the CCV.

The responses of this pair of neurons are similar to what we observed for another neuron pair discussed in section 5.1. Like that neuron pair, the large peak offset from zero observed in the CCV and trial-shuffled CCV for N2 and N5 can be explained by spectral-temporal tuning of the neurons. The STRFs of both neurons are depicted in Figure 5.20B, C). Although both neurons are tuned to approximately the same set of frequencies, N5 responds before N2. This shift in response latency between the two neurons explains why there is a high probability that a spike from N5 will be followed by a spike from N2: the stimulus drives both neurons, but N5 responds more quickly than N2.

One surprising result is the fact that the CCV function for the N2 and N5 neuron pair is so different than the CCV function for the N2 and N3 neuron pair, especially when one

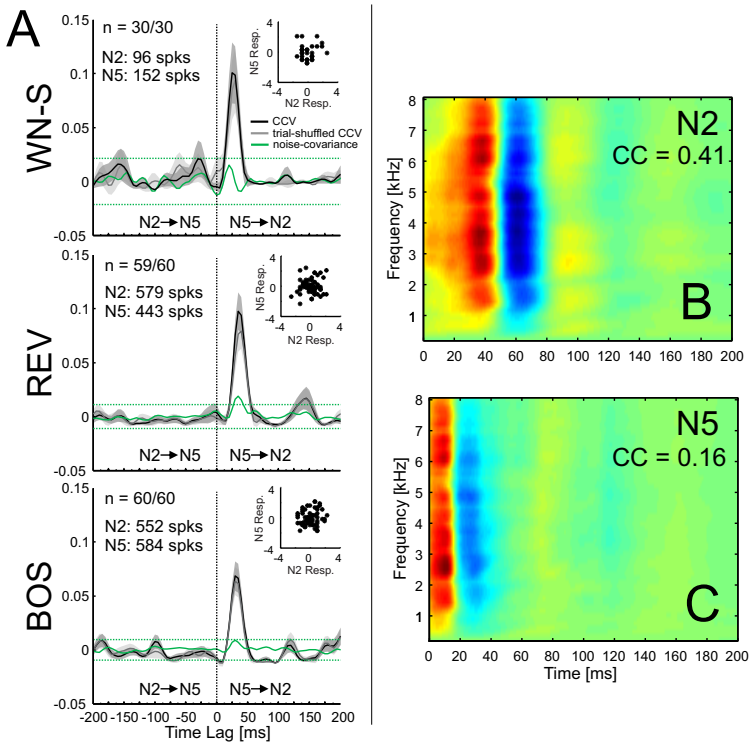

FIGURE 5.20: **Correlation dynamics for Field L neurons 2 and 5.** [r6y12 02] Same figure conventions as described for Fig. 5.19. **A)** The CCV and trial-shuffled CCV were remarkably similar within and across stimulus classes: both were dominated by an asymmetric peak centered around 25 ms (WN-S) to 37 ms (REV and BOS). The similarity between CCV and trial-shuffled CCV functions indicates that correlations in spike times arise from the stimulus tuning of the neuron pair, in which N5 responds to the stimulus before N2. A significant peak in the noise-covariance is present for REV, but not for the other stimuli. **B, C)** STRFs calculated for N2 (**B**) and N5 (**C**), indicate that both neurons are tuned to a broad range of frequencies, and that N5 responds to the stimulus before N2.

considers that the STRFs for N3 and N5 are so similar ($STRF_{SI}$ = 0.93). One explanation may be that although the tuning of N3 and N5 are similar, they are not identical, and the small differences in tuning properties of these neurons account for the widely different correlated spiking patterns observed.

Unlike N2 and N3, the neuron pair N2 and N5 did not have significant spike count correlations for any of the stimuli: R_{SC} = -0.028, p = 0.88 for WN-S; R_{SC} = 0.074, p = 0.57 for REV; R_{SC} = -0.13, p = 0.31 for BOS. It is interesting to note that in this neuron pair we once again observe that, although not significant, REV evoked a slightly positive correlation, whereas BOS evoked a negative correlation.

We examined the correlations between a final pair of neurons, N3 and N5. The spike train cross-covariance functions are displayed in Figure 5.21A. As previously mentioned, these neurons had highly similar tuning properties ($STRF_{SI}$ = 0.93; Fig. 5.21B, C). The CCV and trial-shuffled CCV functions were characterized by a peak centered

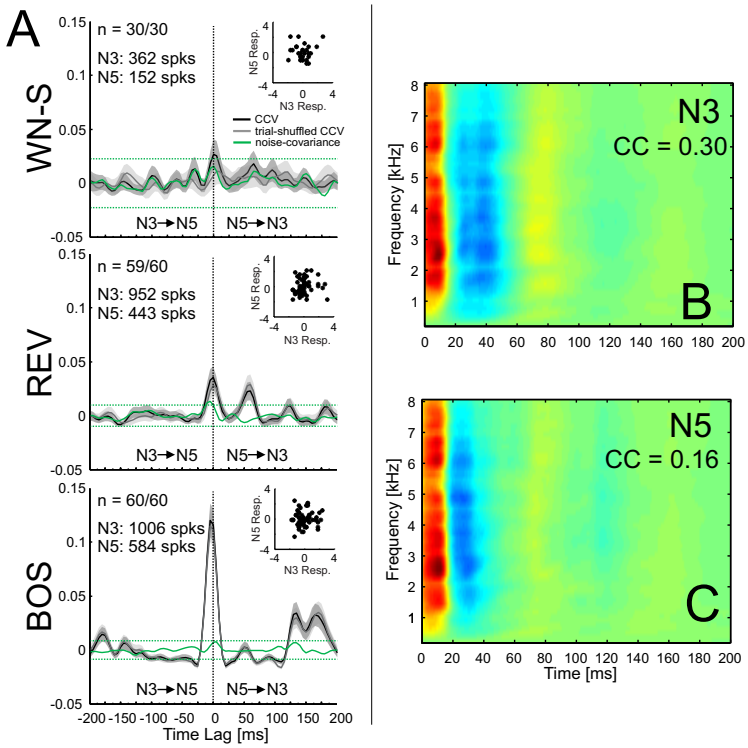

FIGURE 5.21: **Correlation dynamics for Field L neurons 3 and 5.** [r6y12 02] Same figure conventions as described for Fig. 5.19. **A)** The CCV and trial-shuffled CCV functions were characterized by a peak centered on zero time lag, indicative of spiking synchrony resulting from stimulus tuning. The peak height differed across stimuli, and was the largest for the BOS stimulus. A peak in the noise-covariance for REV was also present. **B, C)** STRFs calculated for N3 (**B**) and N5 (**C**) indicate highly similar tuning between the two neurons.

on zero time lag. The peak was present in response to WN, REV, and BOS, and only different across stimuli in terms of the peak height. Because the CCV and trial-shuffled CCV functions were so similar to each other it suggested that these two neurons tended to fire synchronously due to the effects of stimulus drive. The noise-covariance function for REV had a peak at slightly negative time lags that surpassed the 99% confidence interval, suggesting that N3 may directly influence the spiking of N5 in response to REV.

The CCV peaks centered at zero time lag for WN and REV were small in comparison to the CCV peak for BOS. Indeed, these two neurons seemed ideally equipped to fire in response to BOS introductory notes. Interestingly, N3 reliably responded to the first 3 introductory notes, whereas N5 typically responded robustly to only the first 2 introductory notes (Fig. 5.18). The only discernible difference between the 2 and 3rd introductory note is a slight increase in sound amplitude, highlighting the effect that subtle differences in the stimulus can have large effects on spiking responses. Despite the differential responses to sound amplitude, these two neurons had similar enough response properties that stimuli with energy in a broad range of frequencies elicits responses from both neurons simultaneously, and it just happens that the introductory

notes present during BOS drive robust responses from these neurons.

As would be anticipated from two neurons with such similar tuning, correlations in spike counts were positively correlated for all stimuli but only significant for BOS: R_{SC} = 0.017, p = 0.93 for WN-S; R_{SC} = 0.026, p = 0.84 for REV, and R_{SC} = 0.28, $p < 0.05$ for BOS.

3.1 Recapitulation

We examined the correlated responses of 3 neurons located in Field L (Fig. 5.16). STRF estimation indicated that these three neurons were tuned to a broad range of frequencies (Fig. 5.17) and spiking responses indicated that these neurons responded to sound onsets and were particularly driven by introductory notes present in the BOS (Fig. 5.18). Despite have similar receptive field tuning, the patterns of peaks and troughs in the spike time cross-covariance functions were surprisingly different for each pair of neurons.

For all neuron pairs, the CCV and trial-shuffled CCV functions were nearly identical, indicating that features of the auditory stimuli drive the patterns of correlated spike times. For two neurons pairs, the noise-covariance in response to REV crossed the 99% confidence interval, suggesting that there might be a direct interaction between neurons in response to REV.

The CCV and trial-shuffled CCV functions for N2 and N3 were characterized by two main peaks that were offset from zero time lag and multiple smaller secondary peaks (Fig. 5.19). The magnitude of the peak height was stimulus-specific, and the introductory notes present in BOS drove the most robust spiking responses from both neurons. The two neurons shared significantly positive spike count correlations in response to REV, but significant negative spike count correlations in response to BOS.

For N2 and N5, the CCV and trial-shuffled CCV functions were characterized by a large peak that was offset from zero time lag (Fig. 5.20). The magnitude of the peak height was similar for WN-S and REV, but reduced in response to BOS. Like the previous neuron pair, the two neurons share positive (although not significant) spike count correlations in response to REV, but negative (although also not significant) spike count correlations in response to BOS.

A final neuron pair, N3 and N5, had highly similar STRFs, and this similarity was apparent in the CCV and trial-shuffled CCV, which were dominated by a peak centered around zero time lag (Fig. 5.21). The magnitude of this peak was largest for BOS, and resulted from both neurons being robustly driven by the BOS introductory notes

present in the stimulus. This neuron pair was positively correlated for all stimuli, but significantly correlated only for BOS.

Overall, these results show that stimulus-evoked spike times of broadly-tuned onset neurons commonly found in Field L are strongly correlated with each other as a result of stimulus drive. These three neurons differed mainly in terms of their response latencies, and yet these slight differences were enough to elicit vastly different patterns of peaks and troughs in the CCV functions. Such correlation in spike times may allow a downstream neuron that received input from such onset neurons to determine specific features of the stimulus, such as tempo, by receiving near-coincident spikes from different input sources.

Furthermore, these results show that some neurons in Field L are robustly-driven by introductory notes found in the BOS, and furthermore, are sensitive to slight differences in the sound amplitude of these notes. While it is likely that these neurons would also respond to introductory notes or syllables in other conspecific songs that resemble the introductory notes of the BOS, it is nevertheless interesting to observe that a small population of neurons exist in Field L that are robustly driven by autogenous introductory notes.

CH 6

Chapter 6

Results: Neural Correlations in Auditory Populations

1 A Population of Auditory Forebrain Neurons

1.1 Population Inclusion Criteria

In order to examine correlated activity within a population of forebrain neurons, we defined a set of criteria such that a neuron was only included in the subsequent population analysis if it met the following criteria:

- The neuron responded during WN-S, REV, CON, BOS, and WN-E stimuli and displayed spikes to at least 1/3 of all stimulus repetitions in each stimulus block
- The evoked firing rate of the neuron to WN-S and WN-E were not significantly different from each other (paired t-test, $p > 0.01$).
- The putative location of the neuron was within the auditory forebrain (i.e., located in Field L, NCM, or CM).

These criteria provided a population of 106 neurons (out of 229) and 77 neuron pairs (out of 272) from 14 birds (median age = 235.5 dph). Figure 6.1 displays the reconstruction of the recording sites included in the population analysis. Neuron responses included in this analysis were characterized by large signal-to-noise (S/N) ratios (median S/N = 17.74, std = 10.82) and Gaussian-distributed waveform clusters. An example of 3 simultaneously recorded neurons included in the population analysis is presented in Figure 6.2.

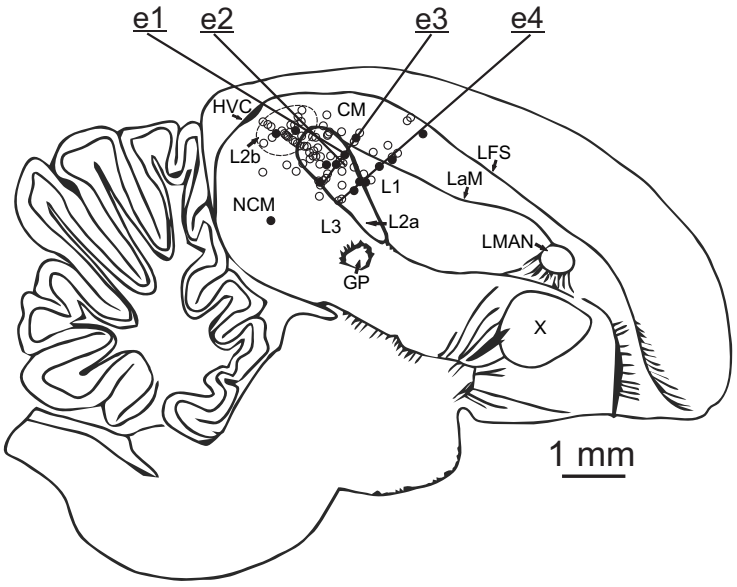

FIGURE 6.1: **Putative recording sites included in the auditory forebrain population.** Open circles represent the putative location of single neurons; filled circles indicate locations where multiple neurons were recorded. e1-e4: mean penetration tracks of the 4 electrodes. The dotted line indicates the region that we considered to be L2b. Abbreviations: L2a, L2b, L1, L3: subregions of Field L; CM: caudal mesopallium; NCM: nidopallium caudal medial; HVC: (used as a proper name); GP: globus pallidus; LMAN: lateral magnocellular nucleus of the anterior neostriatum; X: Area X; LFS: Lamina frontalis superior; LaM: Lamina mesopallialis.

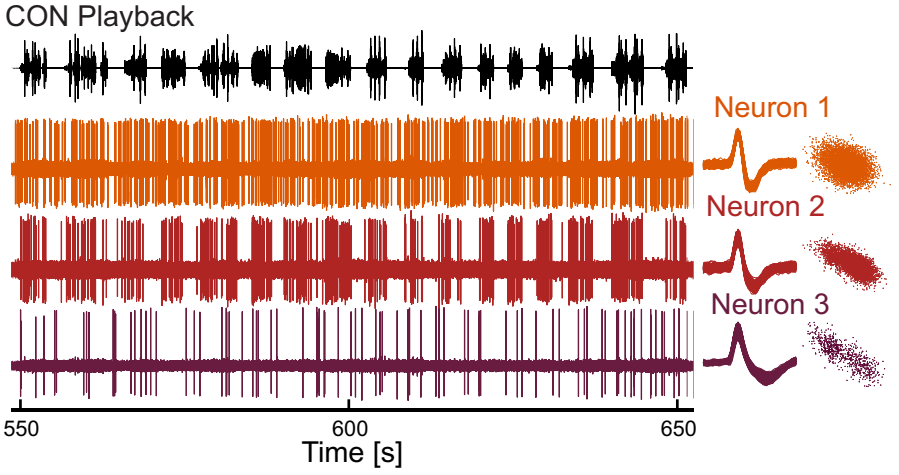

FIGURE 6.2: **Exemplary neural responses for 3 simultaneously recorded neurons.** [p11r4 03] Spiking dynamics in response to CON playback. Neural responses were characterized by large S/N ratios and Gaussian-distributed waveform clusters.

1.2 Notes on Population Analysis

1.2.1 Population Spike Count Correlations

The spontaneous and stimulus-evoked (z-normalized) R_{SC} values were pooled over all eligible pairs from all birds included in the population.

1.2.2 Population Spike Time Cross-Covariance

The mean CCV functions for each spontaneous and stimulus-evoked spike time cross-covariance calculation were pooled over all eligible pairs from all birds included in the population.

We were only interested in correlations that exist at short time lags indicative of spiking synchrony, since spike time correlations at short time lags usually indicate direct connectivity between neurons. Therefore, for our population analysis, we used the data from the -10 ms to +10 ms time bins, or a total of 5 data points per neuron pair (5 ms bins: -10 ms lag, -5 ms lag, 0 ms lag, +5 ms lag, +10 ms lag).

1.2.3 Population Averaging

For each stimulus block, we calculated R_{SC} and CCV for spontaneous and stimulus-evoked (bout and motif) periods separately. We calculated an average spontaneous correlation and an average stimulus-evoked correlation. Specifically, we calculated spontaneous correlations that occurred during the WN-S stimulus block separately from those that occurred during the BOS stimulus block. As a population, the spontaneous firing rates were not significantly different across stimulus blocks (RM Anova; $F(4, 420) = 2.35$, $p = 0.053$), so we pooled the spontaneous R_{SC} and CCV correlations over all stimulus blocks. We compared these pooled spontaneous correlations to stimulus-evoked correlations for each stimulus block.

1.2.4 Elimination of Outliers

In order to avoid contamination of our CCV estimates by outlier responses, we eliminated one neuron pair with a stimulus-evoked CCV value at 0 ms time lag larger than 5 standard deviations of the rest of the population. This was the highly correlated neuron pair discussed in Figure 5.13.

1.2.5 Population Statistics

Correlation Significance
A one-sample Student's t-test was used in order to determine whether the population mean (z-normalized) spike count correlation was significantly different from zero.

Stimulus Class Effect on Correlations
In order to analyze the effect of different stimulus classes on pairwise correlations, we used a Repeated Measures ANOVA (RM Anova). We determined significance using post-hoc Student's paired t-tests, Bonferroni corrected for the number of comparisons.

For four comparisons (e.g., comparisons between WN-S, REV, CON, BOS, and WN-E), the Bonferroni corrected alphas were

- α = 0.05; p = 0.0125
- α = 0.01; p = 0.0025
- α = 0.001; p = 0.00025

For three comparisons (e.g., comparisons between WN-S, REV, BOS, WN-E), the Bonferroni corrected alphas were

- α = 0.05; p = 0.0167
- α = 0.01; p = 0.0033
- α = 0.001; p = 0.00033

2 Neural Responses are Heterogeneous and Sparse in the Auditory Forebrain

Our population of auditory forebrain neurons included 106 neurons (Fig. 6.1) and 77 simultaneously recorded neuron pairs. We began by examining the basic response properties of this population of neurons.

2.1 Neural Responses to Auditory Stimuli are Heterogeneous

We probed our population with 4 different types of stimuli: WN, REV, CON, and BOS. Consistent with other studies of forebrain responses to natural and synthetic stimuli [Amin et al., 2004; Grace et al., 2003; Leppelsack, 1974; Woolley et al., 2009]), neurons in this population responded heterogeneously to the stimulus ensemble. An example of the diversity of neural responses to WN stimuli are shown in Figure 6.3. Many neurons responded to stimulus onsets with transient firing of a few spikes (Fig. 6.3A, 6.3B). Other neurons responded to the offsets of the stimulus with increases in phasic firing (Fig 6.3C, 6.3D). Yet other neurons responded with tonic excitation (Fig. 6.3E) or tonic suppression (Fig. 6.3F) for the duration of the WN stimulus. Other complex responses including delayed phasic responses to onsets were also observed (Fig. 6.3G-J). The tonic and phasic responses to WN highlight the ability of these neurons to encode the onsets and offsets of auditory stimuli.

For all stimulus classes, the neuron population showed a significant increase in median firing rate compared to spontaneous firing (Wilcoxon signed rank, WN; $p < 0.05$; REV, CON, BOS; $p < 0.001$; Fig. 6.4). Both spontaneous and evoked firing, however, were typically low. The median stimulus-evoked firing rate was 3.63 Hz. The mean stimulus-evoked firing rate was slightly higher (5.53 Hz) because it was influenced by a small set of high-firing neurons (see Fig. 6.4). The same pattern was observed for spontaneous firing: a low median (2.24 Hz) and a slightly higher mean (4.07 Hz).

These low firing rates are in line with previous reports: Grace et al. [2003] recorded from single units throughout Field L and CM in urethane-anesthetized zebra finches and reported a mean background firing rate of 2.8 ± 2.7 Hz for their single units. Our observed mean spontaneous response of 4.07 Hz $\pm$ 5.60 (std), albeit slightly higher, is within the range that they reported. Similarly, for the same neurons they reported a mean response to CON of 7.6 ± 9.1 Hz. Our mean population response to CON of 5.76 $\pm$ 6.03 Hz, is slightly lower, but still within their reported range.

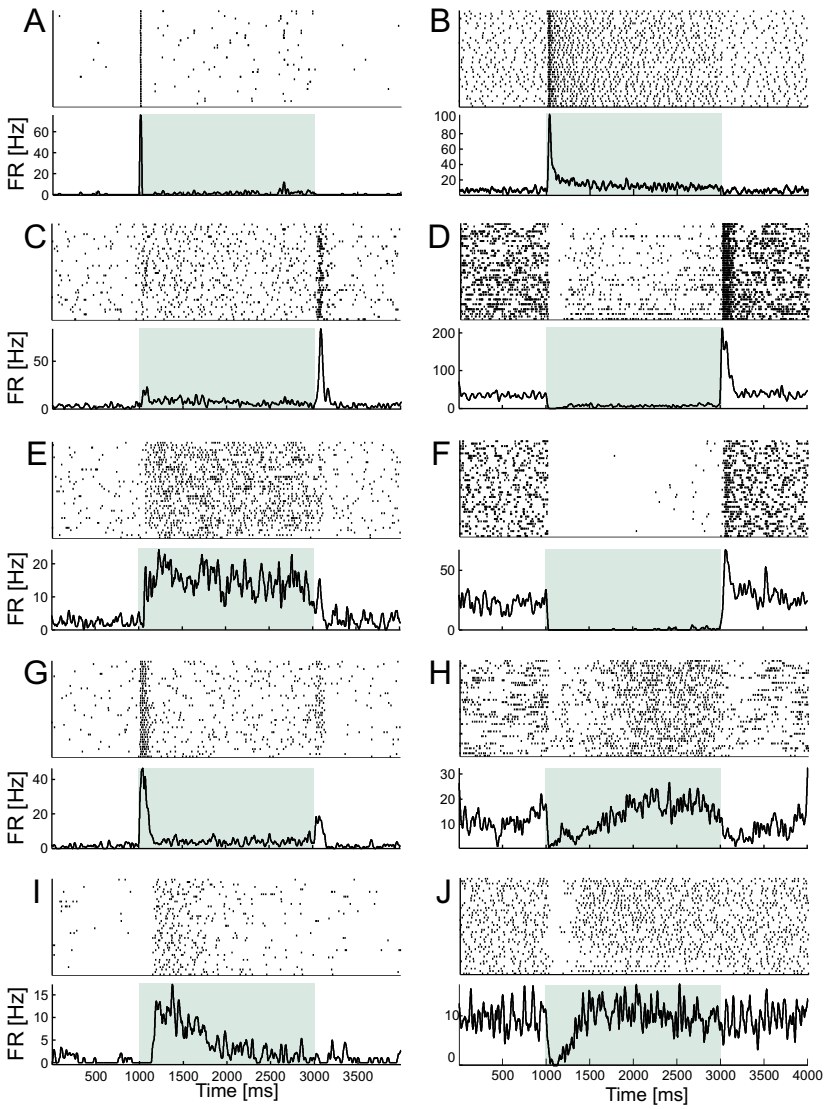

FIGURE 6.3: **Heterogeneous responses of auditory forebrain neurons to WN playback. A-J)** Panels display response dynamics of ten neurons included in the population analysis in response to WN. Top, raster plot shows spiking responses of a neuron to 40 trials of WN. Dots represent individual spikes. Bottom, firing rate curve integrated over all trials. Green shaded region indicates the WN duration (2 s). Many neurons showed transient and phasic responses to the onsets and the offsets of the stimuli. Neural responses included: **A)** Transient onset response. **B)** Transient onset response with tonic excitation. **C)** Tonic excitation with transient offset response. **D)** Tonic suppression with transient offset response. **E)** Tonic excitation. **F)** Tonic suppression. **G)** Phasic onset response followed by phasic offset response. **H)** Phasic onset suppression followed by tonic excitation. **I)** Delayed phasic onset response. **J)** Phasic onset suppression.

2.2 Neural Responses to Auditory Stimuli are Sparse

Given our low firing rates, we were curious as to what was the typical stimulus-evoked response of this population of neurons. Using previously established methods [Hromádka

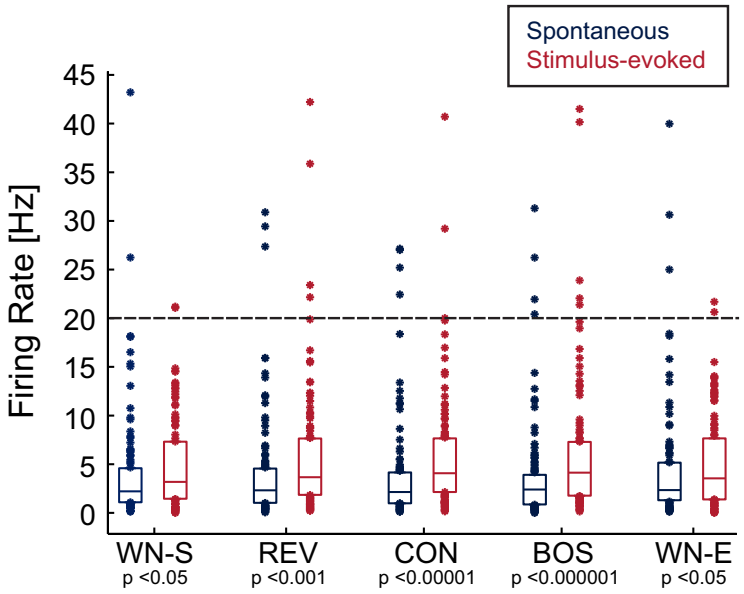

FIGURE 6.4: **Average firing rates for auditory forebrain population.** Spontaneous firing rates (dark blue) are significantly lower than stimulus-evoked firing rates (red; Wilcoxon signed rank; p-values listed below) for 106 neurons include in the population analysis. The line in the middle of the box represents the median value, the lower box edge represents the 25th percentile, and the upper box edge represents the 75th percentile. Responses outside of this range are plotted as dots. For each of the stimulus classes, only a few neurons were well-driven and displayed responses larger than 20 spikes per second (black dashed line).

et al., 2008], we defined an arbitrary threshold of 20 Hz (Fig. 6.4), beyond which a neuron was labeled as "well-driven". Using this metric, we determined the fraction of neurons per stimulus class with responses larger than 20 Hz. Only a small fraction of the population, less than 5%, showed well-driven responses. Interestingly, Hromádka et al. [2008] also reported that very few neurons (less than 5%) in unanesthetized rat auditory cortex were well-driven by tones, sweeps, white-noise burst, and natural stimuli, suggesting that sparse auditory responses may generalize across vertebrates.

2.3 Population is Driven by BOS and Birdsong Stimuli

In order to further understand the stimulus response in each neuron, we calculated z-scores. The z-score is the neuron's normalized auditory-evoked firing rate minus its normalized baseline firing rate. Positive z-scores indicate that the neuron's spiking activity increased during auditory playback, whereas negative scores indicate that activity was suppressed by the playback.

A significant effect of stimulus class was found for population z-scores (RM Anova; $F(4, 420) = 6.80$, $p < 0.0001$), and the mean population z-scores in response to BOS was significantly larger than the mean z-score to WN-S and WN-E (paired t-test, p

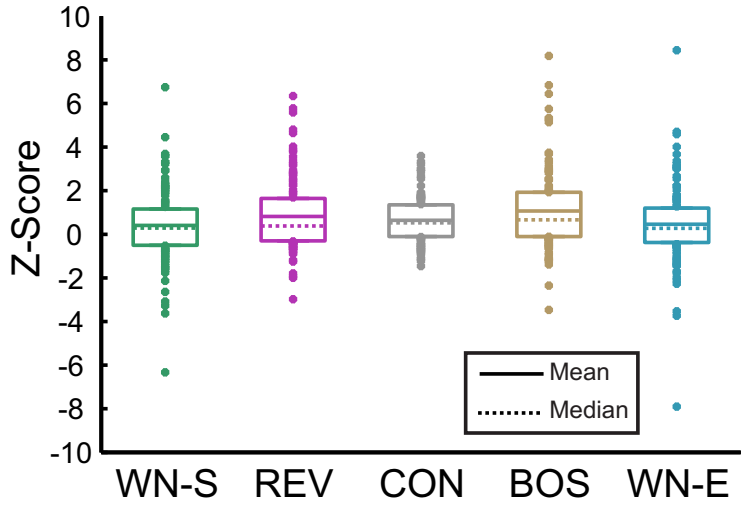

FIGURE 6.5: **Z-scores distributions for auditory forebrain population.** Box plots show the z-scores calculated for each of the 106 neurons included in the population for each stimulus class. Lower box edge represents the 25th percentile, and upper box edge represents the 75th percentile. All neural responses outside of this range are plotted as dots. The median z-scores (dotted line) tended to be lower than the mean values (full line).

< 0.01, Bonferroni corrected for four comparisons). The distributions of z-scores for different stimulus classes were diverse, with WN evoking the largest coefficient of variation (CV; $CV_{WN-S} = 4.33$; $CV_{WN-E} = 4.14$). REV ($CV_{REV} = 2.06$) also evoked slightly more variable z-scores than CON ($CV_{CON} = 1.65$) and BOS ($CV_{BOS} = 1.74$). CON was remarkable in that it evoked the smallest variance in z-scores (Fig. 6.5) despite the fact that CON stimulus block was presumably more variable than REV or BOS, because the CON stimulus block consisted of the songs of several birds, whereas BOS and REV stimulus blocks contained songs from a single bird. As observed with the firing rates, the mean z-scores were always higher than the medians, the result of a few responsive neurons dominating the mean.

Mean z-scores of our population were slightly lower than those reported elsewhere [Amin et al., 2004; Grace et al., 2003]. These differences are most likely due to the fact that these authors pooled together multi-unit and single-unit data [Amin et al., 2004] and calculated z-scores from separate neuron populations defined as having "excitatory" or "inhibitory" responses [Amin et al., 2004; Grace et al., 2003]. Unlike these authors, we used only single unit data, and pooled neurons regardless of response type. Grace et al. [2003] reported mean z-scores ranged from around 0.8 (WN) to around 2.3 (CON) for "excitatory" single units and from around -0.5 (WN) to around -0.3 (CON) for "inhibitory" single units. Amin et al. [2004] reported mean z-scores around 2 (BOS,

CON, REV) for pooled "stimulus-excited units" and around -0.5 for pooled "stimulus-inhibited units". Our mean z-scores are less than the values reported for "excitatory" /"stimulus-excited" neurons, and higher than those values reported for "inhibited" / "stimulus-inhibited" neurons.

In order to assess the selectivity of this population of neurons, we calculated d-prime scores for each neuron included in this analysis and averaged over the population. Averaged over the population, the neurons were slightly selective to the natural stimuli (REV, CON, BOS) versus WN (Fig. 6.6A). Individual neurons, however, displayed strong selectivity to one stimulus versus another. Individual d-prime scores for each of the 106 neurons included in the population are displayed in Figure 6.6B. While some neurons were strongly selective to BOS (d-prime > 2), other neurons were strongly non-selective (d-prime < 2).

Again, our mean population d-prime values are less than what has been reported elsewhere. Amin et al. [2004] reported a d-prime value of BOS versus CON for pooled "stimulus-excited units" of approximately -0.5, and for BOS versus REV of approximately 0.2. In contrast, our mean BOS versus CON d-prime score was close to 0 (d-prime = 0.08 ± 0.12 (SEM)), although our mean BOS versus REV d-prime score was similar to that reported in [Amin et al., 2004]: d-prime = 0.22 ± 0.18). Discrepancies between our d-prime values and those reported in [Amin et al., 2004] most likely arise because we did not subdivide our responses into "excitatory" or "inhibitory" responses.

2.4 Recapitulation

We characterized the basic response properties of our population of 106 auditory forebrain neurons. Generally, the population responded heterogeneously (Fig. 6.3) and robustly to auditory stimulation. As a population, the neurons showed an increase in firing rate in response to all of the stimulus classes (Fig. 6.4). Overall, few neurons responded with greater than 20 spikes/s, suggesting that this population of neurons responds sparsely to auditory stimulation. We used other firing rate-based metrics to characterize the responses of this neuron population. Z-scores for BOS were significantly different than those for WN-S or WN-E (Fig. 6.5). Similarly, d-prime scores indicated that this population of neurons is slightly selective for birdsong and birdsong-like stimuli (i.e., REV, CON, BOS) over WN stimuli (Fig. 6.6). Our z-scores and d-prime scores are slightly lower than previously reported [Amin et al., 2004; Grace et al., 2003], which is most likely an effect of including all neurons in our population analysis.

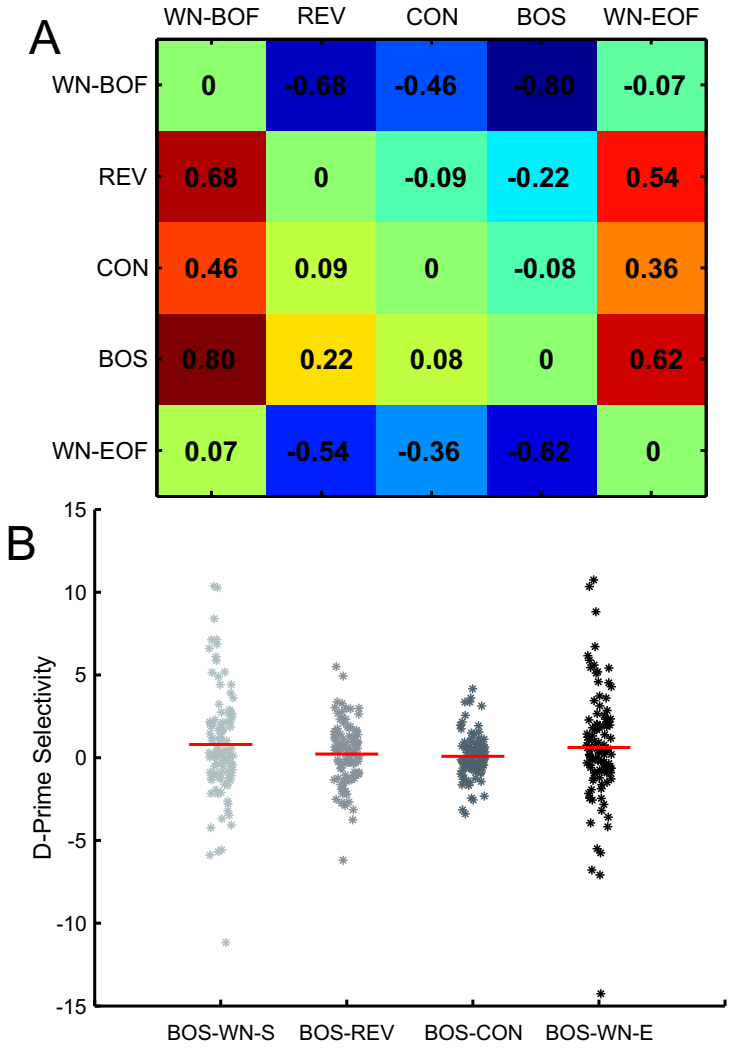

FIGURE 6.6: **D-prime scores for auditory forebrain population. A)** D-prime scores show that the population was slightly selective to natural stimuli versus WN. The mean population d-prime scores for each stimulus combination (row versus column) are shown numerically and depicted in a color-coded scheme where warm colors represent positive d-prime scores and cool colors represent negative d-prime scores. The population of neurons was slightly selective to REV, CON, and BOS versus WN-S and WN-E. **B)** D-prime scores calculated individually for each of the 106 neurons included in the population. BOS versus WN-S (BOS-WN-S); BOS versus REV (BOS-REV); BOS versus CON (BOS-CON); BOS versus WN-E (BOS-WN-E). Red horizontal lines indicate the means.

3 Auditory Forebrain Neuron Population: Spike Count Correlations

To examine spike count correlations R_{SC} in our population of 77 neuron pairs, for each stimulus block we compared the distribution of spontaneous R_{SC} values with the distribution of stimulus-evoked R_{SC} values.

3.1 Spontaneously, Few Neurons Have Significant R_{SC} Values

Because the mean spontaneous firing rates were not significantly different across stimulus blocks, we pooled spontaneous spike count correlations across all stimulus blocks.

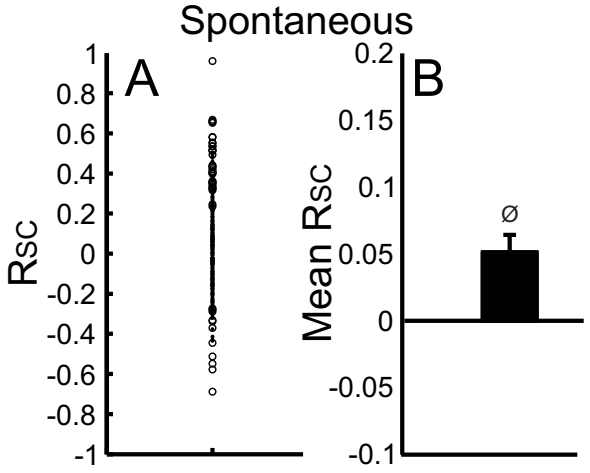

FIGURE 6.7: **Population spike count correlations during spontaneous firing. A)** Spontaneous R_{SC} values (n = 385 pairs; = 77 pairs x 5 stimulus blocks). Open circles indicate R_{SC} values that are statistically significant ($p < 0.05$), and dots indicate R_{SC} values that were not ($p > 0.05$). **B)** Mean R_{SC} value pooled over all pairs. Error bar indicates the SEM, and Ø indicates rejection of the zero mean correlation hypothesis (t-test; $p < 0.05$).

Overall, only a few of the neurons pairs ($\sim$12%; 48 out of 385 pairs) showed significant spike count correlations ($p < 0.05$) during spontaneous firing within one of the stimulus blocks (Fig. 6.7A, open black circles). Of these pairs, the majority was positively correlated ($\sim$77%; 37 out of 48 correlated pairs).

The average spontaneous R_{SC} over all pairs was positive ($R_{SC} = 0.051 \pm 0.013$; mean $\pm$ SEM) and significantly different from zero (t-test; $p < 0.001$; Fig. 6.7B, black bar).

3.2 No Difference Between Spontaneous and Stimulated R_{SC} Values

Next we examined spike count correlations for the first 1.5 s of auditory stimulation with WN and bouts of birdsong (REV, CON, and BOS). Stimulus-evoked R_{SC} pooled over all stimulus classes were not significantly different from pooled spontaneous R_{SC} (paired t-test, p = 0.27).

3.3 Mean Population R_{SC} Values Not Modulated by Stimulus Class

Although we found no significant difference between spontaneous and stimulus-evoked spike count correlations, we were curious to see whether pairwise spike count correlations during auditory stimulation might be stronger for one stimulus class versus another, so we analyzed the stimulus-evoked R_{SC} separately for each stimulus class.

We found that stimulus class did not have a significant effect on the R_{SC} (RM Anova; F(4, 304) = 1.47, p = 0.21). None of the spike count correlations for different stimulus classes were significantly different from each other (paired t-test; $p > 0.05$; Bonferroni corrected for 4 comparisons).

Approximately the same percentage of pairs showed significant stimulus-evoked correlations as during spontaneous firing ($\sim$14%, or 55 out of 385 pairs for stimulus-evoked firing versus $\sim$12%, or 48 out of 385 pairs for spontaneous firing).

The fractions and percentages of pairs associated with significant spike count correlations ($p < 0.05$) are summarized in Figure 6.9. WN-S evoked the largest number of significantly correlated pairs, and $\sim$19% of neuron pairs (15 out of 77) showed significant WN-S R_{SC} (Fig. 6.8A; open green circles). Of these pairs, about half were positively correlated ($\sim$46%; 7 out of 15 pairs). The average R_{SC} for WN-S over all pairs was low ($R_{SC} = 0.015 \pm 0.031$), and not significantly different from zero (t-test, p = 0.63; Fig. 6.8B; green bar).

The same pattern was true for WN-E: $\sim$16% of neuron pairs (12 out of 77) were significantly correlated (Fig. 6.8A; open blue circles). Similar to the WN-S stimulus, half of the correlated neuron pairs showed significant positive spike count correlations (50%, or 6 of 12 pairs).

The mean WN-E R_{SC} was slightly more positive than the mean R_{SC} for WN-S ($R_{SC} = 0.030 \pm 0.029$ for WN-E versus $R_{SC} = 0.015 \pm 0.031$ for WN-S), but the means were not significantly different from each other (paired t-test, p = 0.65). When we pooled over all pairs, the mean WN-E R_{SC}, like the mean WN-S R_{SC}, was not significantly different from zero (t-test, p = 0.30; Fig. 6.8B; blue bar).

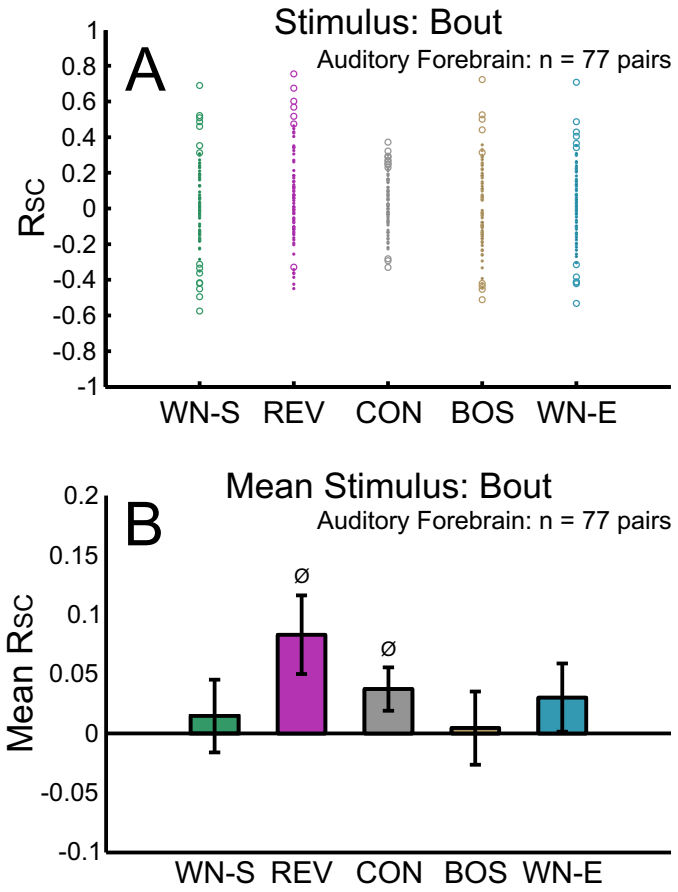

FIGURE 6.8: **Spike count correlations during playback of WN and song bouts.** Same figure conventions as described for Fig. 6.7. **A)** R_{SC} values reported separately for each stimulus block (n = 77 neuron pairs per block). CON playback elicited the smallest variance in R_{SC} values compared to the other stimuli, while BOS elicited the largest coefficient of variation. See text for details. **B)** Mean R_{SC} values per stimulus block. REV and CON were the only stimuli that elicited mean spike count correlations that were significantly different from zero.

Of the birdsong stimuli, bouts of both REV and BOS were associated with similar numbers of significantly correlated pairs, but with different proportions of positive and negative R_{SC} values. For REV, only ∼9% of pairs (7 out of 77) were significantly correlated (Fig. 6.8A; open pink circles). Of these, all but one pair were positively correlated (∼86 %, or 6 out of 7).

During bouts of BOS, ∼12% of pairs (9 out of 77) were associated with significant R_{SC} values (Fig. 6.8A; open brown circles). In contrast to REV, the ratio of positive to negative correlations for BOS was more evenly split: ∼56% of pairs (5 out of 9) were positively correlated, and ∼44% (or 4 out of 9) of pairs were negatively correlated.

The mean spike count correlation for REV was significantly different from zero (t-test, $p = 0.014$; mean $R_{SC} = 0.083 \pm 0.033$; Fig. 6.8B; pink bar). In contrast, the mean spike

	Bouts					Motifs		
	Fraction of Pairs					Fraction of Pairs		
	WN-S	**REV**	**CON**	**BOS**	**WN-E**	**REV**	**CON**	**BOS**
Significantly correlated pairs (p < 0.05)	15/77	7/77	12/77	9/77	12/77	9/77	20/77	12/77
Positively correlated pairs (RSC > 0)	7/15	6/7	9/12	5/9	6/12	9/9	14/20	8/12
Negatively correlated pairs (RSC < 0)	8/15	1/7	3/12	4/9	6/12	0/9	6/20	4/12
	Percentage of Pairs					Percentage of Pairs		
	WN-S	**REV**	**CON**	**BOS**	**WN-E**	**REV**	**CON**	**BOS**
Significantly correlated pairs (p < 0.05)	19%	9%	16%	12%	16%	12%	26%	16%
Positively correlated pairs (RSC > 0)	47%	86%	75%	56%	50%	100%	70%	67%
Negatively correlated pairs (RSC < 0)	53%	14%	25%	44%	50%	0%	30%	33%

FIGURE 6.9: **Significant R_{SC} values as a function of stimulus class.** Table summarizes the total number of significant R_{SC} values ($p < 0.05$) per stimulus class, as well as the fraction of significant pairs that had positive ($R_{SC} > 0$) or negative ($R_{SC} < 0$) correlations. Numbers are summarized for bouts (left) and motifs (right). Fractions are converted to percentages and summarized in the lower half of the table.

count correlation for BOS was not significantly different from zero (t-test, p = 0.88; R_{SC} = 0.0045 $\pm$ 0.031; Fig. 6.8B; brown bar).

In response to bouts of CON, $\sim$16% of neuron pairs (12 out of 77) were associated with significant spike count correlations (Fig. 6.8A; open gray circles). Similar to the REV stimulus, the number of positively correlated pairs was larger than the number of negatively correlated pairs: $\sim$75% of pairs (9 out of 12) were associated with positive R_{SC}s. The mean spike count correlation for CON was positive and significantly different from zero (t-test, p = 0.044; mean R_{SC} = 0.037 $\pm$ 0.018; Fig. 6.8B; gray bar).

The spike count correlations in response to CON were interesting in the sense that, like the z-score distributions (see Fig. 6.5), the CON R_{SC} distribution showed the smallest variance of all stimulus classes: σ^2_{CON} = 0.026 versus σ^2_{WN-S} = 0.072, σ^2_{REV} = 0.084, σ^2_{BOS} = 0.073, and σ^2_{WN-E} = 0.063). For each stimulus class we also examined the coefficient of variation of R_{SC} values within stimulus classes. BOS was associated with the largest variability in spike count correlations, with CV_{BOS} = 60.16, versus CV_{REV} = 3.49,CV_{CON} = 4.28, $CV_{WN\text{-}S}$ = 18.27, and $CV_{WN\text{-}E}$ = 8.34.

In order to investigate whether spike count correlations might be stronger for smaller stimulus windows, we also calculated R_{SC} for spikes that occurred during the first 200 ms of the song motif. In line with the results from our bout analysis, there was not a significant effect of the stimulus class on spike count correlations at the motif level for REV, CON, or BOS (RM Anova, F(2, 224) = 0.68, p = 0.51).

Generally more spike train pairs were significantly correlated for the 200 ms of a stimulus than for 1.5 seconds of the stimulus. This most likely reflects the fact that many

more trials (> 100 motifs) were used for the calculation of the R_{SC} at the motif level rather than were used for the bout level calculations.

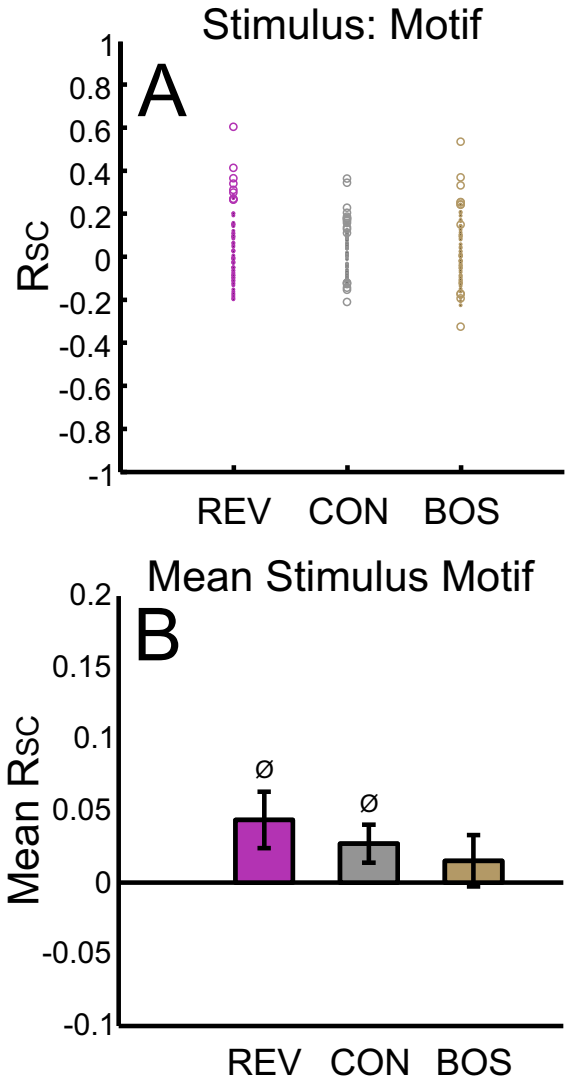

FIGURE 6.10: **Spike count correlations during motif-evoked responses.** Same figure conventions as described for Fig. 6.7. **A)** R_{SC} values were analyzed separately for each stimulus block (n = 77 pairs per block). Matching our findings for the song bout analysis, the smallest variance of R_{SC} values were associated with CON motif playback, whereas the largest coefficient of variance was associated with BOS motif playback. **B)** Mean R_{SC} values per stimulus block. Mean spike count correlations calculated for 200 ms of the motif were generally smaller than those calculated on 1.5 s of the song bout. Matching our findings for the song bout analysis, REV and CON (but not BOS) motifs elicited mean spike count correlations that were significantly different from zero.

For REV motif spike count correlations, ∼11% of pairs (9 out of 77) were significantly correlated, and all were positively correlated (Fig. 6.10A; open pink circles). When we averaged over all pairs, the mean REV motif R_{SC} was significantly different from zero (R_{SC} = 0.044 ±0.020, p = 0.033; Fig. 6.10B; pink bar). Interestingly, the mean R_{SC}

for the REV motif was almost half of the value of the mean R_{SC} for REV bouts (R_{SC} = 0.083 $\pm$ 0.033 for REV bouts versus R_{SC} = 0.037 $\pm$ 0.018 for REV motifs).

For CON motif spike count correlations, almost twice the number of neuron pairs showed significant R_{SC} values compared to CON bouts. $\sim$30% of neuron pairs (20 out of 77) were significantly correlated during CON motifs (Fig. 6.10A; open gray circles) compared to $\sim$16% that were correlated during CON bouts. Of the 20 pairs significantly correlated during CON motifs, 70% (14 out of 20) were positively correlated, and only 30% (6 out of 20) were negatively correlated. These proportions are roughly the same as we observed for CON bout spike count correlations. When we averaged over all pairs, the mean CON motif R_{SC} was slightly lower than the mean CON bout R_{SC}, but still significantly different from zero (R_{SC} = 0.027 $\pm$ 0.013; p = 0.043; Fig. 6.10B; gray bar).

As we observed for the bout level analysis, the variance of CON motif R_{SC} values was smaller than that of the other stimulus classes (σ^2_{CON} = 0.013 compared to σ^2_{REV} = 0.030 and σ^2_{BOS} = 0.025). Also similar to what we observed for song bouts, the BOS motif elicited the largest variability in spike count correlations relative to the mean: CV_{BOS} = 10.35 compared to CV_{REV} = 3.95 and CV_{CON} = 4.25.

Slightly more neurons were correlated during the BOS motif than during the BOS bout: $\sim$16% of pairs (12 out of 77) showed significant R_{SC} values ($p < 0.05$; Fig. 6.10A; open brown circles). Of these pairs, $\sim$67% (8 out of 12) were positively correlated. When we averaged over all pairs, the mean spike count correlation for the BOS motifs was larger than the mean R_{SC} value obtained for the BOS bouts (R_{SC} = 0.0152 $\pm$ 0.018 for BOS motifs compared to R_{SC} = 0.0045 $\pm$ 0.0308 for BOS bouts), although like the spike count correlation for BOS bout, it was not significantly different from zero (t-test, p = 0.4024; (Fig. 6.10B; brown bar).

3.4 Tuning Similarity Partially Explains Significant R_{SC} Values

We examined whether significant spike count correlations resulted from overlapping regions present in STRFs. We estimated STRFs for each neuron of the 12 pairs that showed significant spike count correlations during CON bouts (Fig. 6.8A; open gray circles). We then compared the spike count correlation for each neuron pair to the similarity index of the estimated STRFs.

For 5 of the 12 neuron pairs, significant positive spike count correlations could be explained by overlapping excitatory and inhibitory regions of the STRFs, indicated by positive similarity index (SI_{STRF}) values (see Methods for details; Fig. 6.11; upper right

STRFs). However, for an additional 4 pairs, positive spike count correlations for CON bouts could not be explained by overlapping regions of STRFs, as evidenced by negative SI_{STRF} values (Fig. 6.11; lower right STRFs).

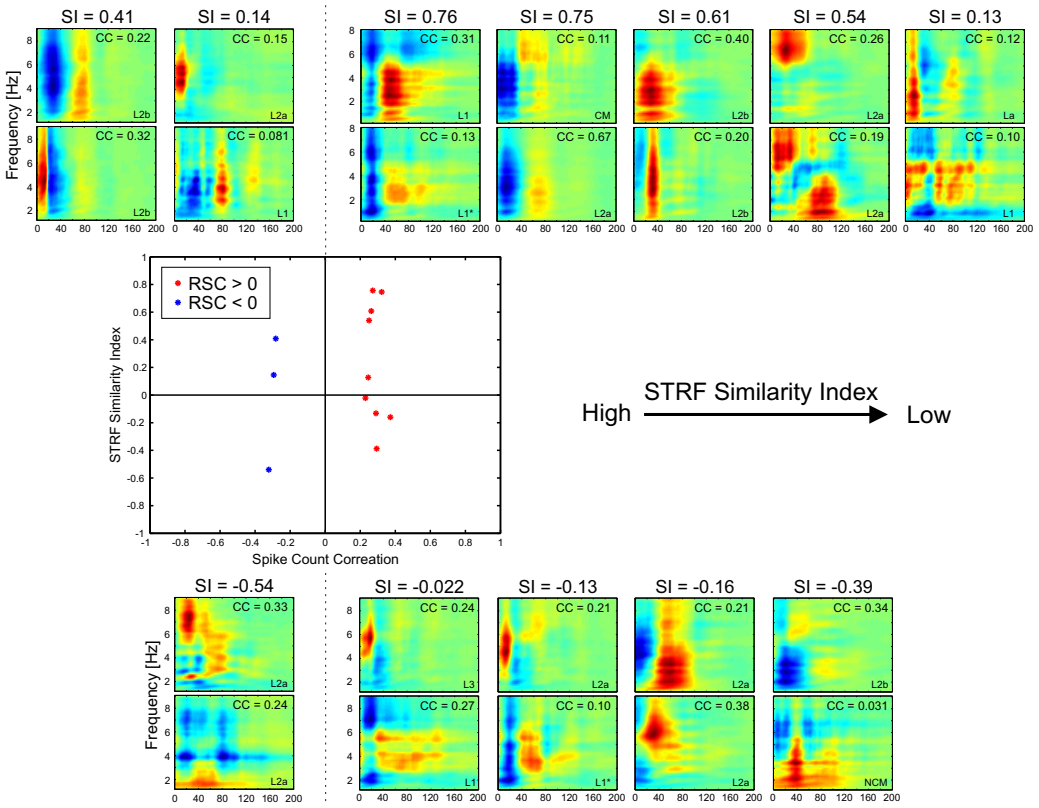

FIGURE 6.11: **Spike count correlations compared to STRF similarity.** STRFs were estimated for neuron pairs with significant spike count correlations in response to CON bouts, and the STRF similarity index (SI_{STRF}) was calculated for the resulting pair of STRFs. SI_{STRF} are plotted as a function of spike count correlation (center plot). Positive spike count correlations are indicated in red, negative spike count correlations are indicated in blue. STRFs for each neuron pair are also displayed. Left, STRFs for pairs with $R_{SC} < 0$; right, STRFs for pairs with $R_{SC} > 0$. Top, STRFs for pairs with $SI_{STRF} > 0$; bottom, STRFs for pairs with $SI_{STRF} < 0$. For each STRF, the predictive CC value is depicted in the upper right corner, and the putative location of the neuron is located in the lower right corner. * indicates an L1 neuron that was recorded twice in combination with different neurons. See text for details.

Although it is easy to understand why overlapping regions of the STRFs could confer significant positive spike count correlations, since neurons with similar auditory responses will tend to fire together and are likely to have correlations in their spike counts, it is less clear how anti-correlated tuning could still confer significant positive spike count correlations. One explanation could be that the features to which neurons with anti-correlated STRFs respond often appear together in CON songs. Thus, although the individual tuning of the neurons is anti-correlated, the overall effect of CON stimulation is to drive a response in both neurons, leading to positive correlations in spike counts.

Among neuron pairs that showed significant negative spike count correlations in response to CON bouts, 2 pairs were associated with positive SI_{STRF}s (Fig. 6.11; upper left STRFs), and one pair with a negative SI_{STRF} (Fig. 6.11; lower left STRFs). Negative spike count correlations typically mean that while one neuron was driven by the stimulus, the other neuron was not. We can see how this interplay might work for the neuron pair consisting of two L2b neurons (Fig. 6.11; top left STRF pair). Both of their STRFs have large inhibitory side bands between 20 and 40 ms, which gives the neuron pair a positive SI_{STRF} = 0.41. However, one of the neurons is a typical "offset" neuron (top STRF) in the sense that its responds to sounds in a broad range of frequencies followed by silence. In contrast, the other neuron of this pair is a typical "onset" neuron (lower STRF) that fires preferentially to a broad range of frequencies preceded by silence. In this sense, although the two neurons have similar STRFs, the overall function of their tuning properties anti-correlates the spike counts evoked by CON stimuli.

One final neuron pair had both a significant negative spike count correlation (R_{SC} = -0.33) and a negative STRF similarity index (SI_{STRF} = -0.54; Fig. 6.11; lower left STRFs). How do two neurons with anti-correlated tuning also respond to CON bouts with anti-correlated spike counts? We hypothesize that like the neuron with strong positive spike count correlations but anti-correlated tuning, these two neurons must respond to features of CON that are likely to appear together in time, only in this case, one tends to be excited by the stimulus whereas the other neuron is suppressed.

3.5 Recapitulation

We examined spike count correlations in a population of neurons during spontaneous firing and during auditory stimulation. During spontaneous firing, several pairs shared significant spike count correlations (Fig. 6.7A). When averaged over the population of neurons, the mean spike count correlation during spontaneous firing was non-zero and positive (Fig. 6.7B), suggesting that as a population, these neurons share significant variability during silence.

However, when we compared spike count correlations during spontaneous firing to spike count correlations during stimulus-evoked firing, there was no significant difference between the two conditions. Indeed, when we examined the effect of auditory stimulation on spike count correlations, we found that stimulus class did not have a significant effect on correlations in spike counts that were summed either over long time periods (1.5 s; Fig. 6.8) or over short time periods (200 ms; Fig. 6.10).

Although we did not see a significant effect of stimulus class on the mean spike count correlation values (RM Anova), the mean R_{SC} for WN and BOS were small and not

significantly different from zero. WN-S, BOS, and WN-E were the only stimuli that evoked almost equal numbers of significant negatively and positively correlated pairs. This stands in contrast to CON and REV, which were associated with primarily positively correlated pairs.

We also examined the STRFs for pairs with significant spike count correlations for CON bouts (Fig. 6.11). Although for many neuron pairs positive spike count correlations could be explained by overlapping regions of the STRFs, the correlations in other neuron pairs could only be explained by considering the specific features of the auditory stimulus that drove the neurons in the context of the entire CON song.

4 Auditory Forebrain Neuron Population: Spike Train Cross-Covariances

In this chapter we are interested in whether spike-train correlations differ for spontaneous and stimulus-evoked conditions. To examine the structure of spike train correlations in our population of neurons, we examined CCV functions (CCV, trial-shuffled CCV, and noise-covariance) for all neuron pairs included in our analysis. For stimulus-evoked correlations, we examined the CCV functions for each stimulus block separately, such that for each stimulus-block, we had a total of 77 stimulus-evoked CCV functions and a distribution of 77 spontaneous CCV functions.

We calculated CCV function on spike train discretized into 5 ms binds and used the "coef" option for the Matlab *xcov* function (see Methods for details). Although we used the spiking response to the entire spontaneous or stimulus-evoked period in our CCV calculation, we were only interested in correlations that occurred at short time lags, indicative of spiking synchrony. Therefore, for the subsequent statistical analysis, we used the data from the -10 ms to +10 ms time lags, or a total of 5 data points per neuron pair (5 ms bins: -10 ms lag, -5 ms lag, 0 ms lag, +5 ms lag, +10 ms lag).

4.1 Significant Synchrony Present During Spontaneous Firing

In order to understand the baseline spike train correlation structure during silence, we examined spiking responses during spontaneous firing during silent inter-stimulus intervals. During spontaneous firing, the mean CCV over all pairs was dominated by a large peak centered at zero time lag (Fig. 6.12; black line). As expected, shuffling over trials flattened the CCV peak, (Fig. 6.12; gray line), indicating that spike train correlations were not due to residual influences of the preceding auditory stimulus but due to intrinsic noise that caused the neuron pairs to fire synchronously during silence. The mean noise-covariance (Fig. 6.12; green line), calculated by averaging over the mean pairwise differences between the CCV and the trial-shuffled CCV, almost perfectly matches the CCV.

When we pooled data from the -10 ms to +10 ms time lags for each CCV (Fig. 6.12; blue shading), the resulting mean was large and positive (mean = 0.011 $\pm$ 0.00086; Fig. 6.13; black bar) and almost identical to the mean value at short time lags of the noise-covariance (mean = 0.011 $\pm$ 0.00085; Fig. 6.13; green bar). In contrast, the mean trial-shuffled CCV values at short time lags was essentially zero (trial-shuffled CCV = 0.000017 $\pm$ 0.00016; Fig. 6.13; gray bar).

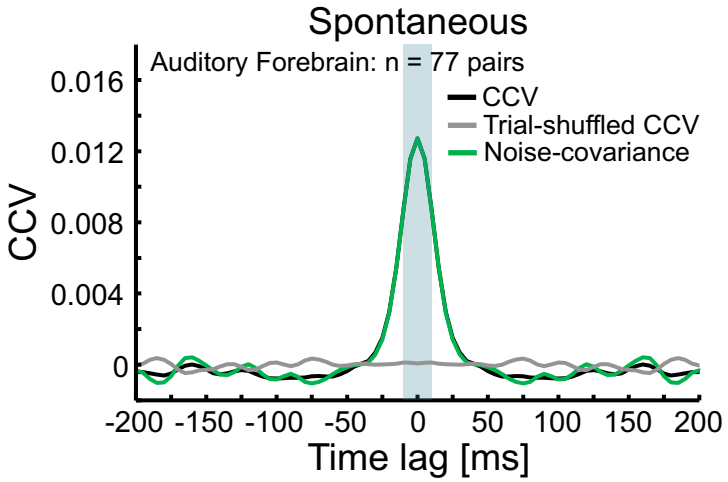

FIGURE 6.12: **Mean population CCV functions for spontaneous activity.** CCV (black line), trial-shuffled CCV (gray line), and noise-covariance (green line) functions for spontaneous activity. CCV and noise-covariance functions are nearly identical and have a broad peak centered at zero time lag. In contrast, the trial-shuffled CCV, representing the contribution from stimulus drive, is flat. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

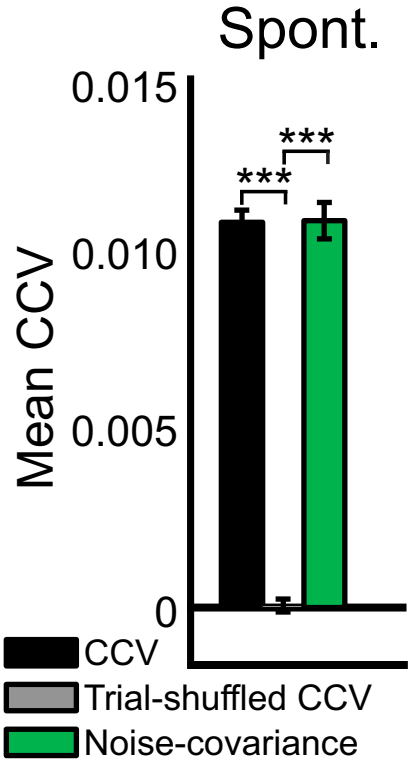

FIGURE 6.13: **Mean CCV values at short time lags for spontaneous activity.** Mean CCV values are pooled from -10 to +10 ms time lags (Fig. 6.12; blue shading). CCV (black bar) and noise-covariance (green bar) values at short time lags were significantly different than the trial-shuffled CCV values (paired t-test, p = 1.03e-076 and 1.96e-059, respectively). Error bar indicates the SEM.

The noise-covariance was significantly more positive than the trial-shuffled CCV at short time lags (paired t-test, p = 1.96e-059), but was not significantly different from the CCV (paired t-test, p = 0.799). This suggests that during spontaneous firing, the large and positive CCV peak at short time lags results from correlated input rather than stimulus-induced reverberations.

4.2 Auditory Stimulation Greatly Reduces Noise-Covariance

In order to evaluate the effect that auditory stimulation has on the network, we pooled data from short time lags for the CCV, the trial-shuffled CCV, and the noise-covariance functions for all the stimulus classes.

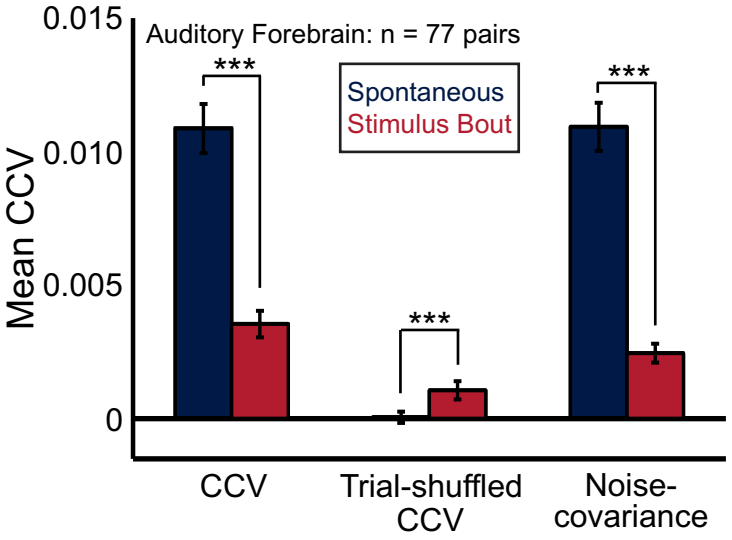

FIGURE 6.14: **Spontaneous vs. stimulus-evoked CCV values at short time lags.** Comparison of mean CCV calculated from pooled spontaneous trials (dark blue bars) compared to mean pooled stimulus trials (red bars). Noise-covariance values at short time lags during spontaneous trials was significantly higher than noise-covariance at short time lags during auditory stimulation.

We pooled data from the -10 ms to +10 ms time lags and found that stimulus-evoked CCVs were dramatically smaller than spontaneously evoked CCVs (paired t-test, p = 1.14e-47; Fig. 6.14). Specifically, auditory stimulation drastically and significantly reduced the spike-time synchrony at short time lags (paired t-test, p = 9.74e-045; Fig. 6.14). Unsurprisingly, the effect of stimulus drive, represented by trial-shuffled CCV, was significantly more positive at short time lags during auditory stimulation than during spontaneous firing (paired t-test, $p < 0.001$; Fig. 6.14).

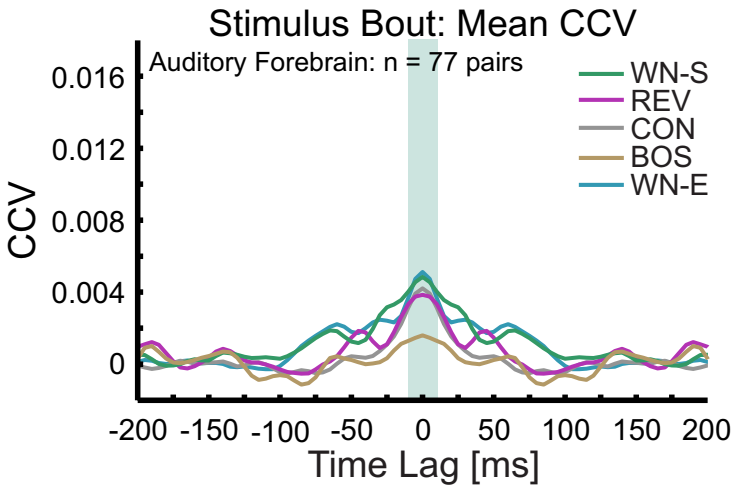

FIGURE 6.15: **Mean CCV functions for stimulus-evoked (bout) responses.** CCV functions were calculated for WN-S, REV, CON, BOS, and WN-E separately. CCV function for BOS has the lowest values at short time lags. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

4.3 Stimulus Class Modulates Cross-Covariance

We were curious whether CCVs during auditory stimulation might be stronger for one stimulus class versus another, so we analyzed the stimulus-evoked CCVs separately for each stimulus class. Smoothed CCV functions for WN-S, REV, CON, BOS, and WN-E are depicted in Figure 6.15.

Unlike what was observed for the stimulus-evoked spike count correlations, a significant effect of the stimulus class was found for the stimulus-evoked spike train cross-covariance at short time lags (RM Anova, $F(4, 1516) = 6.53$, $p < 0.0001$; Fig. 6.16).

The CCV values at short time lags for WN-S and WN-E were not significantly different from each other (paired t-tests, $p = 0.92$), but both were slightly positive and had mean values of 0.0045 ± 0.00058 and 0.0045 ± 0.00057, respectively.

CCV values at short time lags evoked by REV bouts (mean = 0.0035 ± 0.00074) were slightly lower than WN, but not significantly different than WN-S or WN-E (paired t-tests, $p > 0.05$). The same was also true for CCV values evoked by CON bouts (mean = 0.0037 ± 0.00057), which were not significantly different than WN-S or WN-E (paired t-tests, $p > 0.05$).

Unlike all of the other stimuli, the spike train cross-covariance evoked by BOS bouts were low and had a mean of 0.0016 ± 0.00073. The BOS CCV values at short time lags were significantly smaller than all of the other stimulus classes (paired t-tests, $p < 0.01$ for WN-S, REV, and WN-E; $p < 0.001$ for CON; Fig. 6.16).

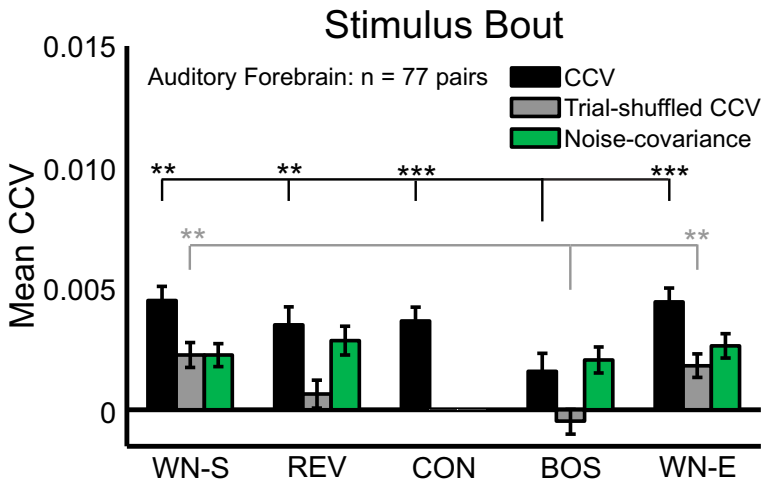

FIGURE 6.16: **Short time lag mean CCV values for stimulus-evoked (bout) responses.** Mean CCV values (black bar), mean trial-shuffled CCV values (gray bar) and mean noise-covariance values (green bar). BOS CCV values at short time lags were significantly lower than all of the other stimuli, and BOS trial-shuffled CCV values were significantly more negative than WN-S and WN-E.

Although we used data from the -10 ms to +10 ms time lags for our statistical analysis (Fig. 6.15; blue shaded area) the CCV results were similar for other time lags (e.g., -20 ms to +20 ms or -5 ms to +5 ms; Fig 6.17).

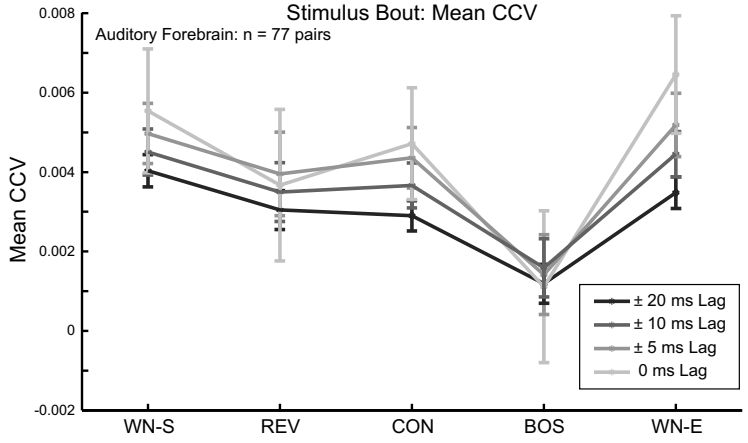

FIGURE 6.17: **Mean CCV values as a function of time lag.** Although we used data from the -10 ms to +10 ms time lags, mean CCVs were similar for larger (-20 ms to +20 ms) and smaller (-5 ms to +5 ms) time lags. Error bars indicate the SEM.

4.4 Stimulus Class Modulates Trial-Shuffled Correlations

We shuffled identical trials to determine the effect that extrinsic stimulus correlations contributed to the CCV correlations. Smoothed trial-shuffled CCV functions are depicted in Figure 6.18. A significant stimulus effect was also found for the trial-shuffled

CCV values at short time lags (RM Anova, F(3, 1137) = 6.91, $p < 0.001$; Fig. 6.16; gray bars).

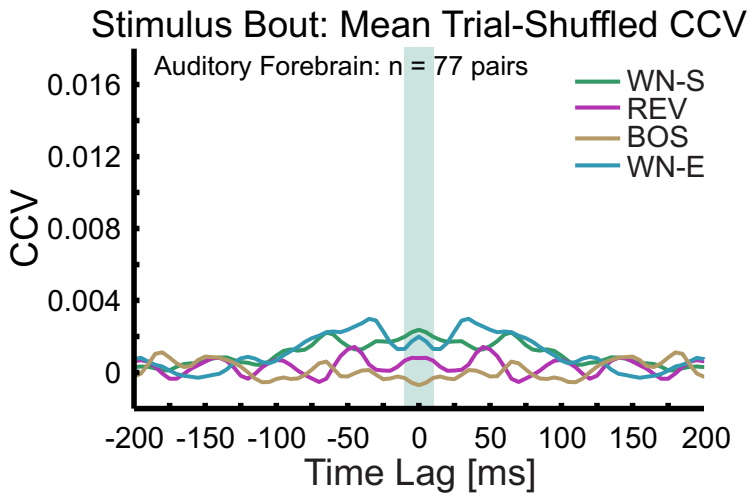

FIGURE 6.18: **Mean trial-shuffled CCV functions for stimulus-evoked (bout) responses.** Trial-shuffled CCV functions were calculated individually for WN-S, REV, BOS, and WN-E, but not for CON, because CON bouts were not repeatedly presented. The mean trial-shuffled CCV function for BOS had the lowest values at short time lag. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

Similar to the CCV, the mean trial-shuffled CCVs at short time lags for WN-S (0.0023 $\pm$ 0.00051) and WN-E (0.0018 $\pm$ 0.00043) were not significantly different from each other (paired t-test, p = 0.37).

The mean trial-shuffled CCV for REV bouts was close to zero (mean = 0.0006 $\pm$ 0.57) and not significantly different than WN-S or WN-E (paired t-test, $p > 0.05$). The trial-shuffled CCV mean for values at short time lags for BOS bouts was unique among all the stimuli in that the mean was negative (mean = -0.0005 $\pm$ 0.00054). It was significantly different from trial-shuffled CCV values at short time lags for both WN-S and WN-E ($p < 0.01$) and almost significantly less than REV bouts (paired t-test, p = 0.048; Fig. 6.16; gray bars).

We were unable to examine the effects of trial-shuffled CCV for CON bouts, but see section 4.6.

4.5 Stimulus-Independent Noise-Covariance

The mean noise-covariance (calculated by subtracting the mean trial-shuffled CCV from the mean CCV for each pair for each stimulus) was computed to determine if there are stimulus-specific effects of noise-covariance that contribute to CCV during auditory stimulation.

Stimulus class was not found to have a significant effect on the noise-covariance (RM Anova; F(3, 1137) = 0.62, p = 0.60; Fig. 6.16C, green bars). All of the stimuli evoked noise-covariance that had positive values at short time lags, but none of the stimuli evoked noise-covariance that was significantly different from another stimulus (paired t-test, $p > 0.05$).

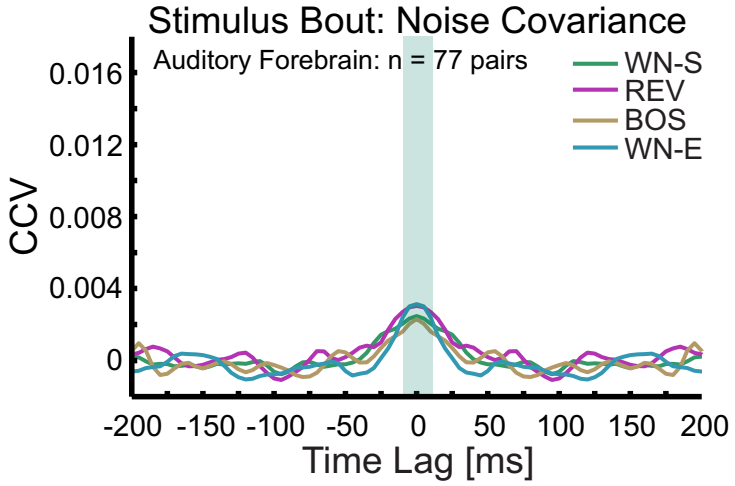

FIGURE 6.19: **Mean noise-covariance functions for stimulus-evoked (bout) responses.** The mean noise-covariance functions were similar for all stimuli. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

4.6 Motif-Level Spike Train Cross-Covariance

In order to investigate more fully the trial-shuffled CCV and noise-covariance which resulted from CON stimuli, we recalculated the CCV, trial-shuffled CCV, and noise-covariance for spikes that occurred during the song motif.

Unlike our bout level findings, no significant effect of stimulus class was found to influence motif CCVs for REV, CON, or BOS (RM Anova, F(2, 758) = 0.10, p = 0.90; Fig. 6.20; black bars). None of the mean CCVs calculated for motifs were significantly different from each other (paired t-test, $p > 0.05$). The mean CCVs at short time lags for CON motifs was slightly more positive than the mean CCVs for REV and BOS motifs.

When we shuffled the motifs to examine whether the stimulus contributes to the spike time correlations, we found a slightly significant effect of stimulus class (RM Anova, F(2, 758) = 3.17, p = 0.042; Fig. 6.20; gray bars). The mean trial-shuffled CCVs at short time lags for BOS motifs was slightly significantly more negative than for CON motifs (paired t-test, p = 0.025). The mean trial-shuffled CCVs at short time lags was not significantly different for BOS and REV motifs, matching our findings for BOS and REV bouts.

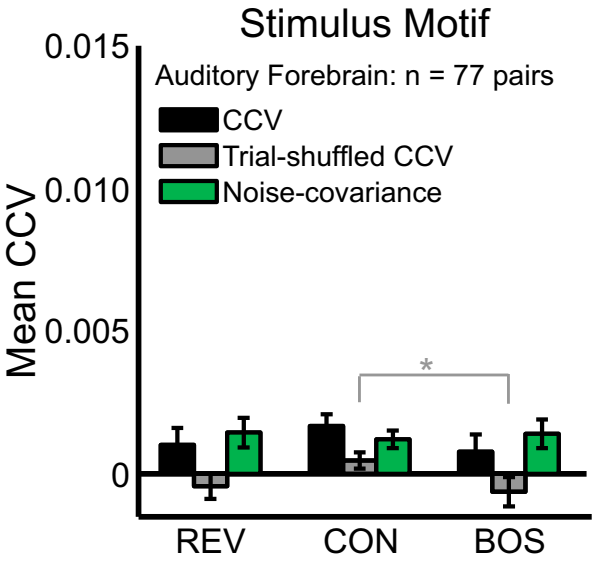

FIGURE 6.20: **Short time lag mean CCV values for stimulus-evoked (motif) responses.** Mean CCV values (black bar), mean trial-shuffled CCV values (gray bar) and mean noise-covariance values (green bar). The mean trial-shuffled CCV for BOS motifs was significantly more negative than the mean trial-shuffled CCV for CON motifs (paired t-test, p = 0.025).

The contribution of noise-covariance to the CCV was not influenced by the stimulus class (RM Anova, F(2, 758) = 0.10, p = 0.90; Fig. 6.20; green bars). Mean noise-covariances for all of the stimuli were positive, but not significantly different from each other (paired t-test, $p > 0.05$), also matching our findings at the bout level.

We were surprised to find such a significant effect of correlations for BOS bouts but not for BOS motifs. We suspected that this had to do with variability of motif lengths. While the BOS motif duration of some birds was very stereotyped (Fig. 6.21A), other birds sang BOS motifs of more variable durations and syllable placements (Fig. 6.21B). In these cases, although the first few syllables are generally very stereotyped, the timing of later syllables shows some jitter. We suspect that this syllable jitter in the motif is responsible for the differences between CCVs calculated on BOS bout and CCVs calculated on the BOS motif.

4.7 Recapitulation

We examined spike time cross-covariance in a population of 106 auditory forebrain neurons. During spontaneous firing, the population showed a large peak centered at zero time lag for the noise covariance (Fig. 6.12) indicating that as a population, the neurons receive correlated input that causes synchronous firing at short time lags (Fig. 6.13). The noise covariance was significantly reduced during auditory stimulation (Fig. 6.14).

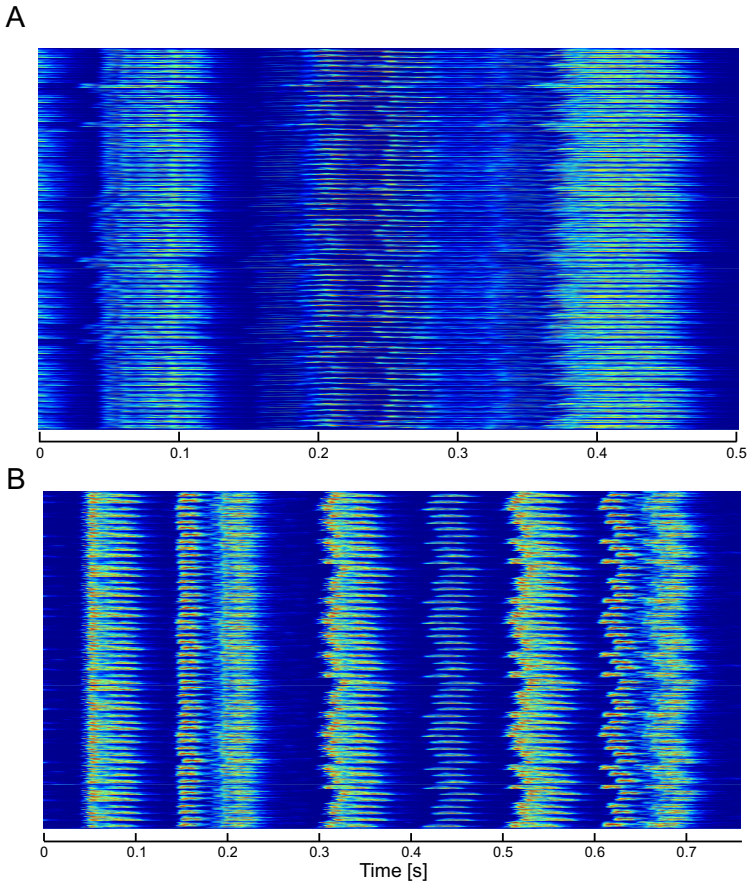

FIGURE 6.21: **Syllable timing jitter present for some BOS motifs.** Repetitions of motifs are depicted as spectrograms and are aligned to the onset of the motif. In **A**, the timing of the syllables is stereotyped, and syllable durations are approximately the same over motif renditions. In **B**, the timing of the later syllables is jittered for this BOS motif, as indicated by the non-aligned syllable onsets and the end of the song.

Stimulus class identity was found to have a significant effect on the spike time cross-covariances. Specifically, the CCV values at short time lags of -10 to +10 ms were significantly lower for BOS bouts than for all other stimuli (Fig. 6.16). These results were similar for longer and shorter time lags (Fig 6.17). Furthermore, the trial-shuffled CCV values at short time lags for BOS bouts were also significantly lower than WN-S and WN-E (Fig. 6.16), and the trial-shuffled CCV values at short time lags for BOS motifs were significantly lower than CON motifs (Fig. 6.20).

Overall these results indicate that during spontaneous firing, the population of neurons receives correlated input from an intrinsic source in the brain. The influence of this source is greatly reduced during stimulus-evoked firing. Furthermore, the effects of stimulus drive during BOS stimulation anti-correlates spike timing at short time lags, causing significantly lower CCV values at short time lags for BOS than for any of the other stimuli.

5 Field L Neuron Population: Correlated Responses

We were curious whether our spike count correlation and spike train cross-covariance results were a property of Field L neurons. Although our population of auditory forebrain neurons consisted primarily of Field L neurons, we wondered whether our results depended on the inclusion of higher-order neurons from CM and NCM.

We used the same selection criteria that we used for the population of 106 auditory forebrain neurons, except that for this analysis we excluded all neurons that were not located in Field L. This left us with a population of 62 neurons that were located in Field L and a total of 45 neuron pairs.

We probed this population of neurons with WN-S, REV, CON, BOS, and WN-E and examined the correlations that occurred spontaneously during silent periods in between the stimulus repetitions and for stimulus-evoked responses.

5.1 Spike Count Correlations Not Modulated By Stimulus Class

We examined the spike count correlations that exist during the first 1.5 s of auditory stimulation with WN and bouts of natural bird song (REV, CON, and BOS) for the population of Field L neurons.

When we compared the population spike count correlations across stimulus blocks, we found that stimulus class did not have a significant effect on the spike count correlation (RM Anova, $F(4, 176) = 0.079$, $p = 0.99$). None of the spike count correlations for any of the stimulus classes were significantly different from each other (paired t-test, $p > 0.05$; Bonferroni corrected for 4 comparisons).

Although this result matched our observations of the larger population of auditory forebrain neurons, some aspects of the spike count correlations we observed in Field L were different from the results we obtained for spike count correlations in the larger population of auditory forebrain neurons.

The fractions and percentages of pairs associated with significant R_{SC} are summarized in Figure 6.9. WN-S evoked the second-largest number of significantly correlated pairs after BOS. $\sim$22% of neuron pairs (10 out of 45) had significant WN-S R_{SC} values (Fig. 6.22A; open green circles). Of these pairs, the majority were positively correlated ($\sim$90%; 9 out of 10 pairs); only one pair had a significant negative spike count correlation. The mean R_{SC} value for WN-S averaged over all pairs was larger than observed for the larger population of auditory forebrain neurons ($R_{SC} = 0.086 \pm 0.040$) and significantly different from zero (t-test, $p = 0.035$; Fig. 6.22B; green bar), unlike what was

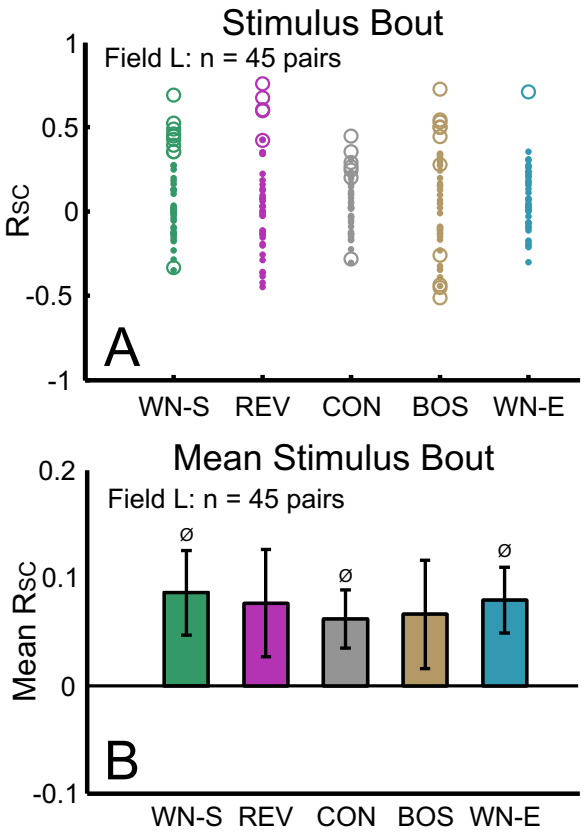

FIGURE 6.22: **Spike count correlations during auditory stimulation for Field L neurons. A)** R_{SC} values were analyzed separately for each stimulus block (n = 45 pairs per block). Open circles indicated R_{SC} values that were statistically significant ($p < 0.05$), and dots indicated values R_{SC} values that were not significant ($p > 0.05$). Matching our results of the larger population of auditory forebrain neurons, CON was associated with the smallest variance in R_{SC} values, whereas BOS was associated with the the largest coefficient of variation. **B)** Mean R_{SC} values per stimulus block. Spike count correlations in response to WN-S, CON, and WN-E were all significantly more positive than zero. This contrasts with what was observed for the larger population of auditory forebrain neurons, for which mean spike count correlations in response to only REV and CON were significantly different from zero. Error bars indicate the SEM. Ø indicates rejection of the zero mean correlation hypothesis (t-test, $p < 0.05$).

observed for the larger population of auditory forebrain neurons, for which the mean WN-S R_{SC} was not significantly different from zero.

Far fewer pairs had significant spike count correlations for WN-E: ∼2% of the neuron pairs (1 out of 45) were significantly correlated (Fig. 6.22A; open blue circle). This reduction in spike count correlations could not be explained by a reduction in overall spiking activity, since the individual evoked firing rates in response to WN-S and WN-E of neurons included in this population were not significantly different. The mean spike count correlation for all 45 neuron pairs in response to WN-E was slightly lower than WN-S ($R_{SC} = 0.079 \pm 0.031$; Fig. 6.22B; blue bar) but not significantly different (paired t-test; $p = 0.88$). Like WN-S for this population of Field L neurons, but unlike the larger

Bouts

Fraction of Pairs

	WN-S	REV	CON	BOS	WN-E
Significantly correlated pairs (p < 0.05)	10/45	5/45	7/45	11/45	1/45
Positively correlated pairs (RSC > 0)	9/10	5/5	6/6	7/11	1/1
Negatively correlated pairs (RSC < 0)	1/10	0/5	1/6	4//11	0/1

Percentage of Pairs

	WN-S	REV	CON	BOS	WN-E
Significantly correlated pairs (p < 0.05)	22%	11%	16%	24%	2%
Positively correlated pairs (RSC > 0)	90%	100%	86%	64%	100%
Negatively correlated pairs (RSC < 0)	10%	0%	14%	36%	0%

FIGURE 6.23: **Significant R_{SC} values as a function of stimulus class.** Table summarizes the total number of significant R_{SC} values ($p < 0.05$) per stimulus class, as well as the fraction of significant pairs that had positive ($R_{SC} > 0$) or negative ($R_{SC} < 0$) correlations. Fractions are converted to percentages and summarized in the lower half of the table.

population of auditory forebrain neurons, the mean WN-E spike count correlation was significantly different from zero (t-test, p = 0.013).

REV and CON bouts both evoked similar numbers of significant spike count correlations. For REV bouts, ∼11% (5 out of 45 pairs) had significant spike count correlations, and all were positive (Fig. 6.22A; open pink circles). Similarly, ∼15.5% (7 out of 45 pairs) of neuron pairs had significant spike count correlations in response to CON bouts. In this case, all but one pair were positively correlated (Fig. 6.22A; open gray circles). The mean population R_{SC} value for REV was 0.0764 ± 0.050, which was not significantly different from zero (t-test, p = 0.13; Fig. 6.22B; pink bar). In contrast, the mean population R_{SC} value for CON, which was slightly less than REV (R_{SC} = 0.062 ± 0.027) was significantly different from zero (t-test, p = 0.027; Fig. 6.22B; gray bar).

As noted for the larger auditory forebrain population, the spike count correlations in response to CON had the smallest variance when compared to the R_{SC} distributions for the other stimuli: σ^2_{CON} = 0.033 versus σ^2_{WN-S} = 0.07, σ^2_{REV} = 0.11, σ^2_{BOS} = 0.11, and σ^2_{WN-E} = 0.042).

Starkly different from what we observed for the auditory forebrain population, BOS evoked the highest number of significantly correlated neuron pairs in Field L: ∼24% of the neuron pairs (11 out of 45) were significantly correlated for BOS bouts (Fig. 6.22A; open brown circles). Furthermore, BOS evoked the highest percentage of significant negatively correlated neuron pairs (∼36%, or 4 out of 11). When we pooled over all

Field L pairs included in the population, the mean BOS spike count correlation was 0.066 ± 0.050, which was not significantly different from zero (t-test, p = 0.19). As observed for the larger population of auditory forebrain neurons, BOS evoked the largest variability in spike count correlations relative to the mean, as measured by the coefficient of variation: CV_{BOS} = 5.09 compared to $CV_{WN\text{-}S}$ = 3.08; CV_{REV} = 4.36; CV_{CON} = 2.93; $CV_{WN\text{-}E}$ = 2.60.

5.2 Spiking Synchrony During Spontaneous Firing

In addition to exploring the spike count correlations that exist between Field L neurons, we were also interested in exploring the correlations in spike times that exist. We began by examining the spike train cross-covariances that exist during spontaneous firing to see whether Field L neurons have the same spontaneous properties observed in the larger population of auditory forebrain neurons.

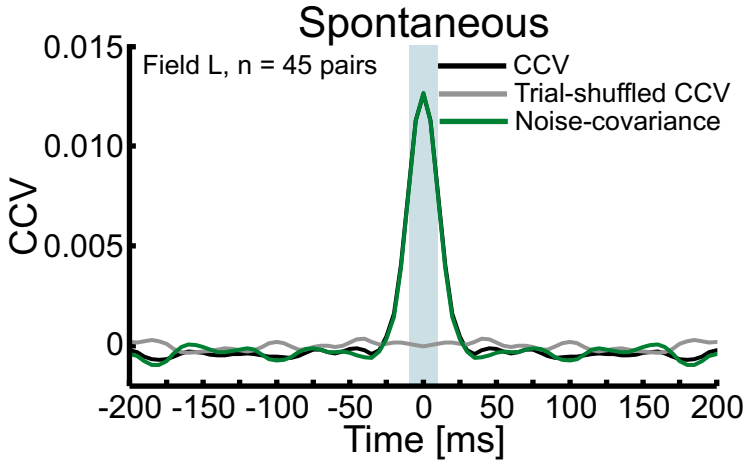

FIGURE 6.24: **Mean CCV functions for spontaneous responses in Field L.** CCV (black line), trial-shuffled CCV (gray line), and noise-covariance (green line) functions for spontaneous responses. CCV and noise-covariance functions are nearly identical and have a broad peak centered at zero time lag. In contrast, the trial-shuffled CCV, representing the contribution from stimulus drive, is flat. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

As was true for the larger population of 106 auditory forebrain neurons, we found that during spontaneous firing of Field L neurons, the CCV and noise covariance functions were both dominated by a large and positive peak that was centered around zero time lag (Fig. 6.24). Also as observed for our larger population of auditory forebrain neurons, the trial-shuffled CCV, which represents the contribution of stimulus drive, was flat (Fig. 6.24). These results suggest that Field L neurons receive shared correlated input and that stimulus drive does not contributes to the mean spontaneous CCV.

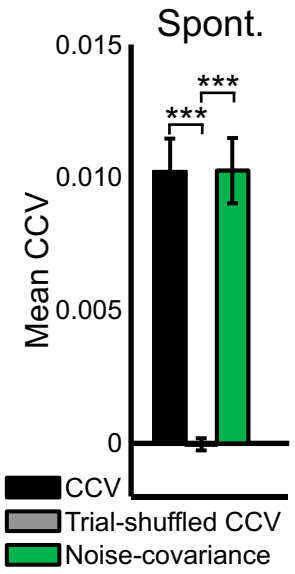

FIGURE 6.25: **Mean CCVs values for spontaneous Field L responses.** Mean CCVs are pooled from -10 to +10 ms time lags (Fig. 6.24; blue shading). CCV (black bar) and noise-covariance (green bar) values at short time lags were significantly different from the trial-shuffled CCVs at short time lags (paired t-test, p = 3.71e-50 and 1.92e-41, respectively). Error bar indicates the SEM.

When we pooled data from the -10 ms to +10 ms time lags for each CCV (Fig. 6.24; blue shading), the mean spike time correlations at short time lags was large and positive (mean = 0.010 $\pm$ 0.0013; Fig. 6.25; black bar) and identical to mean value at short time lags of the noise-covariance (mean = 0.010 $\pm$ 0.0012; Fig. 6.25; green bar). In contrast, the mean trial-shuffled CCVs at short time lags was essentially zero (trial-shuffled CCV = -0.000051 $\pm$ 0.00023; Fig. 6.25; gray bar).

The mean CCV and noise covariance values at short time lags for the Field L population were only slightly less than the mean values at short time lags for the larger auditory forebrain population: forebrain CCV mean = 0.011 $\pm$ 0.00086; forebrain noise covariance mean = = 0.011 $\pm$ 0.00085; Fig. 6.13).

5.3 Auditory Stimulation Reduces Spiking Synchrony

In order to evaluate the effect that auditory stimulation has on the population of Field L neurons, we pooled data from short time lags for the CCV, the trial-shuffled CCV, and the noise-covariance functions for all the stimulus classes.

When we pooled data from the -10 ms to +10 ms time lags and compared stimulus-evoked CCVs to the spontaneously evoked CCVs, we found that the spike time correlations were drastically reduced during auditory stimulation compared to spontaneous

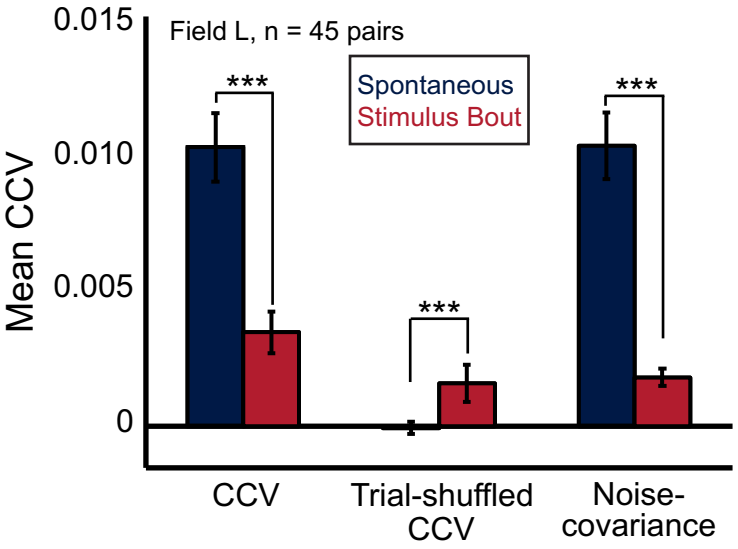

FIGURE 6.26: **Auditory stimulation reduces noise covariance in Field L neurons.** Comparison of CCV calculated from pooled spontaneous trials (dark blue bars) compared to pooled stimulus trials (red bars). Noise-covariance during spontaneous firing was significantly greater than noise-covariance during auditory stimulation.

periods (paired t-test, p = 4.49e-16; Fig. 6.26). Specifically, auditory stimulation drastically and significantly reduced the noise-covariance at short time lags (paired t-test, p = 2.69e-23; Fig. 6.26). Unsurprisingly, the effect of stimulus drive, represented by trial-shuffled CCV, was significantly more positive at short time lags during auditory stimulation than during spontaneous firing (paired t-test, $p < 0.001$; Fig. 6.26).

These stimulus-induced reduction in CCV and noise-covariance, combined with the significant increase in trial-shuffled CCVs during auditory stimulation, closely match those observed for the larger population of auditory forebrain neurons (see Fig. 6.14 for comparison).

5.4 Stimulus Class Modulates Cross-Covariance

We were curious whether CCV correlations during auditory stimulation might be stronger for one stimulus class versus another in Field L neurons, so we analyzed the stimulus-evoked CCV functions separately for each stimulus class. Smoothed CCV functions for WN-S, REV, CON, BOS, and WN-E are depicted in Figure 6.27. Unlike the population of auditory forebrain neurons, the CCV functions for the population of Field L neurons had many prominent secondary peaks at long time lags, which most likely result from stimulus-induced correlations in spike times that occur at longs times.

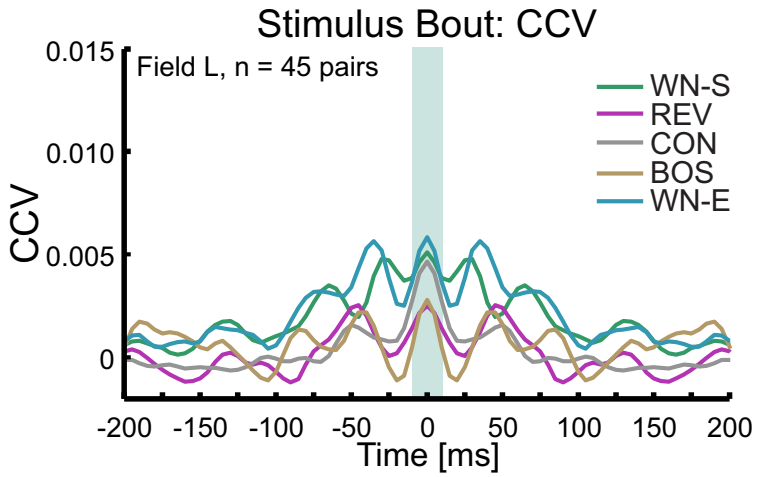

FIGURE 6.27: **Mean CCV functions for stimulus-evoked (bout) responses in Field L.** CCV functions were calculated for WN-S, REV, CON, BOS, and WN-E separately. CCV functions for both REV and BOS had low values at short time lags. Prominent secondary peaks at long time lags are also present for all stimuli. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

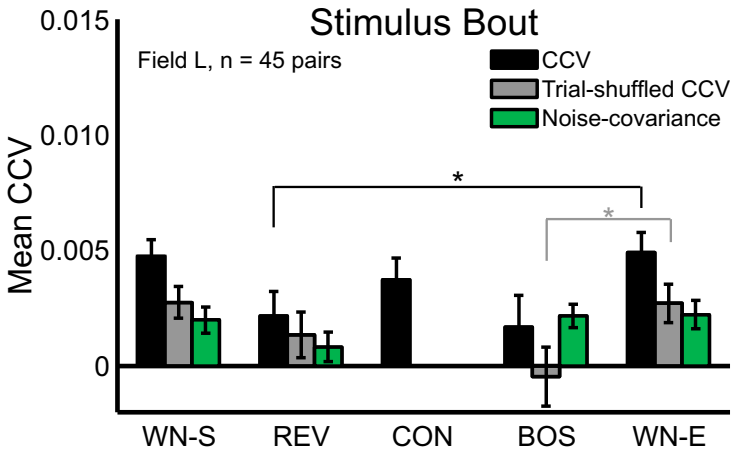

FIGURE 6.28: **Mean CCV values for stimulus-evoked (bout) responses in Field L.** Mean CCVs (black bar), mean trial-shuffled CCVs (gray bar) and mean noise-covariance values (green bar). REV CCVs at short time lags were significantly lower WN-E values (paired t-test, $p < 0.05$), and BOS trial-shuffled CCV values were significantly greater than WN-E values (paired t-test, $p < 0.05$).

As observed for the larger population of auditory forebrain neurons, a significant effect of the stimulus class was found for the stimulus-evoked spike train cross-covariance at short time lags (RM Anova; $F(4, 876) = 3.73$, $p < 0.0052$; Fig. 6.28).

The CCVs at short time lags for WN-S and WN-E were not significantly different from each other (paired t-test, $p = 0.78$), but both were slightly positive and had mean values of 0.0047 ± 0.0007 and 0.0049 ± 0.0009, respectively.

In contrast to the CCVs at short time lags of the larger population of auditory forebrain neurons, spike time correlations for both BOS and REV seem to be similar for Field

L neurons. CCVs at short time lags evoked by REV bouts (mean = 0.0022 $\pm$ 0.0011) were smaller than WN (paired t-test, p = 0.025) and significantly smaller than WN-E (paired t-test, p = 0.010; Fig. 6.28).

CCVs at short time lags for BOS bouts were lower than for REV (mean = 0.0017 $\pm$ 0.0014), but did not meet our stringent criteria for significance (paired t-test, p = 0.031 for WN-S; p = 0.051 for CON; p = 0.014 for WN-E).

CCVs evoked by CON bouts (mean = 0.0037 $\pm$ 0.00057), were not significantly different from any of the other stimuli (paired t-test, $p > 0.05$).

5.5 Stimulus Class Modulates Trial-Shuffled Correlations

We shuffled identical trials to determine the effect that extrinsic stimulus correlations contributed to the CCVs. Smoothed trial-shuffled CCV functions are depicted in Figure 6.29. As observed for the CCV functions (Fig. 6.27) several prominent secondary peaks are present for the mean trial-shuffled CCV functions, highlighting the effect of stimulus-drive on the Field L neuron pairs. A significant stimulus effect was also found for the trial-shuffled CCVs at short time lags (RM Anova, F(3, 657) = 3.59, p = 0.014; Fig. 6.28; gray bars).

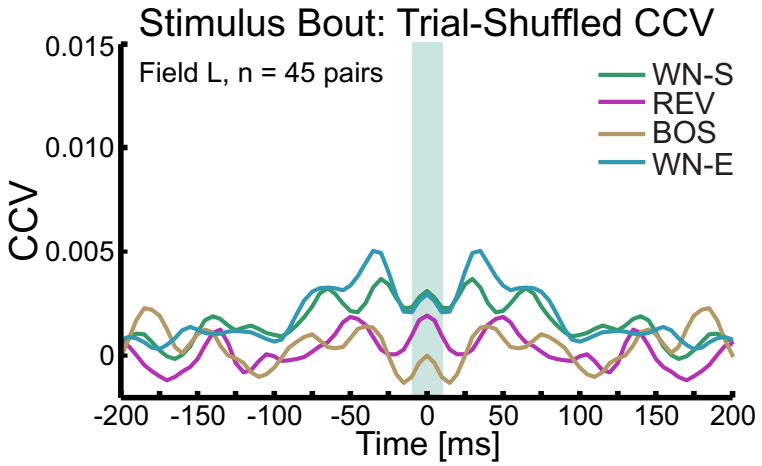

FIGURE 6.29: **Mean trial-shuffled CCV functions for stimulus-evoked (bout) responses.** CCV functions were calculated for WN-S, REV, BOS, and WN-E separately. CCV functions for both REV and BOS had low values at short time lags. Prominent secondary peaks at long time lags are also present for all stimuli. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

Similar to the CCV, the mean trial-shuffled CCVs at short time lags for WN-S (0.003 $\pm$ 0.0007) and WN-E (0.003 $\pm$ 0.0008) were not significantly different from each other (p = 0.96).

The mean trial-shuffled CCV for REV bouts was low (mean = 0.0013 $\pm$ 0.001) and not significantly different from WN-S, BOS, or WN-E (paired t-test, $p > 0.05$).

The trial-shuffled CCV mean for values at short time lags for BOS bouts was unique among all the stimuli in that, as was also observed for the larger population of auditory forebrain neurons, the mean was negative (mean = -0.0005 $\pm$ 0.0013). BOS trial-shuffled CCVs at short time lags were significantly different from trial-shuffled CCVs at short time lags for WN-E (paired t-test, p = 0.011), but did not meet the stringent criteria for significance for the other stimuli (paired t-test, p = 0.023 for WN-S and p = 0.108 for REV).

We were unable to examine the effects of trial-shuffled CCV for CON bouts in this population of Field L neurons, but see the following section.

As observed for the larger population of auditory forebrain neurons, there was not a significant effect of stimulus on the noise-covariance at short time lags (RM Anova, F(3, 657) = 1.25, p = 0.29; Fig. 6.28).

5.6 Motif-Level Spike Train Cross-Covariance

In order to investigate more fully the trial-shuffled CCV and noise-covariance which resulted from the CON stimuli in Field L neurons, we recalculated the CCV, trial-shuffled CCV, and noise- covariance for spikes that occurred during the song motif.

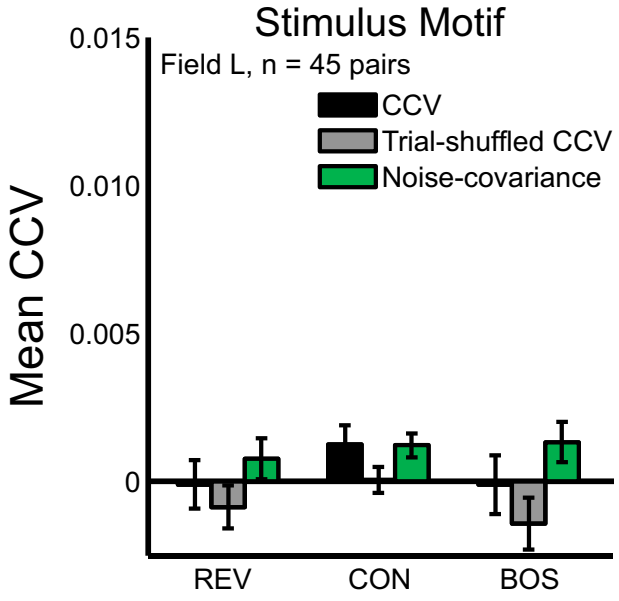

FIGURE 6.30: **Mean CCV values for stimulus-evoked (motif) responses in Field L.** Mean CCVs (black bar), mean trial-shuffled CCVs (gray bar) and mean noise-covariance values (green bar).

Although a significant effect of the stimulus class on CCVs at short time lags was not present for song motifs (RM Anova, F(2, 438) = 0.30, p = 0.74), several interesting differences were present when we compared the Field L population to the larger auditory forebrain neuron population. Most notably, whereas CCVs at short time lags were positive for REV and BOS motifs in the larger population of auditory neurons (see Fig. 6.20) the mean values at short time lags for REV and BOS motifs in the Field L population were near zero (mean = -0.0001 $\pm$ 0.00082 for REV and mean = -0.0001 $\pm$ 0.00098 for BOS; Fig. 6.30). In contrast, the CCVs at short time lags for CON motifs were positive and large (mean = 0.0013 $\pm$ 0.00064; Fig. 6.30), as was also observed for the larger population of auditory forebrain neurons. However none of the CCVs at short time lags were significantly different for one stimulus versus another (paired t-test, $p > 0.05$).

Although a significant effect of the stimulus class on trial-shuffled CCVs at short time lags was not present for song motifs (RM Anova, F(2, 438) = 1.68, p = 0.19), mean trial shuffled CCVs at short time lags suggest that population of Field L neurons could have different stimulus-induced spike time correlations for BOS and REV versus CON. Specifically, the mean trial-shuffled CCVs at short time lags for both REV and BOS motifs were negative (mean = -0.0009 $\pm$ 0.00072 for REV motifs and mean = -0.0014 $\pm$ 0.00088 for BOS motifs) compared to CON motifs, which was near zero (mean = 3.98e-5 $\pm$ 0.00044). However none of the CCVs at short time lags were significantly different for one stimulus versus another (paired t-test, $p > 0.05$), although trial-shuffled CCVs at short time lags were almost significantly different from those evoked by CON motifs (paired t-test, p = 0.054).

The noise-covariance values at short time lags for the song motifs were not significantly different from each other (RM Anova, F(2, 438) = 0.30, p = 0.74).

5.7 Recapitulation

We examined a smaller population of 62 Field L neurons (45 neuron pairs) to see whether the effects that we observed in a larger population of auditory forebrain neurons that included higher-order neurons in CM and NCM were also true of Field L neurons.

As observed for the larger population of auditory forebrain neurons, we found that there was not a significant effect of stimulus class on the spike count correlations for Field L neurons. However, unlike the population of auditory forebrain neurons, spike count correlations for WN stimuli were significantly different from zero for Field L neurons (Fig. 6.22B), which was not the case for the larger population of auditory forebrain neurons.

Similarly, whereas spike count correlations for the larger population of auditory forebrain neurons was significantly different from zero in response to REV, this was not the case for Field L neurons, for which the mean spike count correlation for REV was not significantly different from zero (Fig. 6.22B). Mean spike count correlation responses to CON and BOS were consistent between populations, with CON evoking mean spike count correlations that were significantly different from zero in both populations, and BOS evoking mean spike count correlations that were not significantly different from zero (Fig. 6.22B).

The spike train cross-covariance values were largely similar between the two populations of neurons. As was observed for the larger population of auditory forebrain neurons, both the CCV and the noise-covariance during spontaneous firing for the population of Field L neurons were characterized by a large and positive peak centered around zero time lag (Fig. 6.24). The CCV and noise covariance values at short time lags were significantly different from the trial-shuffled CCVs at short time lags (Fig. 6.24), indicating that during spontaneous firing, this population of Field L neurons received correlate input that cause pairs of Field L neuron to fire synchronously during silence. Furthermore, as was also observed in the larger population of auditory forebrain neurons, auditory stimulation significantly reduced the CCV and noise covariance values at short time lags for Field L neurons (Fig. 6.26).

As was also true for the larger population of auditory neurons, the stimulus class significantly modulated both the CCVs and the trial-shuffled CCVs at short time lags for Field L neurons (Fig. 6.28). Unlike what was observed for the larger class of auditory forebrain neurons, neuron in Field L seem to have similar spike time correlations at short time lags in response to both REV and BOS bouts. While CCVs at short time lags were significantly more negative for BOS compared to any other stimuli for the larger population of auditory forebrain neurons, mean CCVs at short time lags for both BOS and REV bouts were more negative that any of the other stimuli for the population of Field L neurons. Although these values did not meet our stringent criteria for significance, we expect that in a larger population of Field L neurons, both BOS and REV would have CCVs at short time lags that would be significantly different from WN and CON bouts.

We also examined the spike count correlations aligned to song motifs. Once again, the spike time correlations between Field L neurons for BOS and REV motifs were remarkably similar. Mean CCVs at short time lags for both BOS and REV motifs were near zero for Field L neurons, which was not the case for the neuron pairs in the larger population of auditory forebrain neurons, which had positive mean CCVs at short time lags(Fig. 6.30). Furthermore, the mean trial-shuffled CCVs at short time lags were

negative for both BOS and REV motifs, which is similar to the results obtained from the larger population of auditory forebrain neurons. Mean CCVs at short time lags for CON motifs were positive for both populations of neurons, and mean trial-shuffled CCVs across populations were also similar, although the mean trial shuffled values were slightly lower for the population of Field L neurons (Fig. 6.30).

6 Auditory Forebrain Neuron Population: BOS and qBOS

We were curious as to whether correlations between neurons in the auditory forebrain were similar for BOS and qBOS (Fig. 6.31).

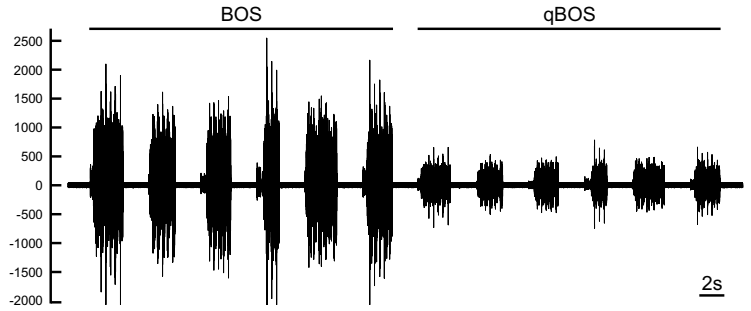

FIGURE 6.31: **Amplitude difference between BOS and qBOS playback.** Oscillogram of BOS (left) versus qBOS (right) amplitudes during auditory playback. Panel displays 6 BOS versions and the identical qBOS versions. qBOS was approximately 1/3 the loudness of BOS.

We used the same criteria that were used to select the larger population of auditory forebrain neurons except that we only included neurons which had responses to WN-S, BOS, qBOS and WN-E. As for the larger population of neurons, we only included neurons if the evoked firing rate of the neuron to WN-S was not significantly different than the WN-E evoked firing rate responses (paired t-test, $p > 0.01$). Because qBOS was not a stimulus used for all neurons, these criteria left us with a sub-population of 82 auditory forebrain neurons and 63 neurons pairs.

Like the larger population of auditory forebrain neurons, this sub-population of neurons was well driven by the stimuli and showed a significant increase in median firing rate compared to spontaneous firing for all of the stimulus classes (Wilcoxon signed rank, WN; $p < 0.01$; BOS and qBOS; $p < 0.001$; Fig. 6.32). For the population, the stimulus-evoked firing rate was significantly higher for BOS versus qBOS (paired t-test, $p = 0.0012$), but was otherwise not significantly different for the other stimuli.

In order to further evaluate the stimulus responsiveness of this population of neurons, we calculated the z-scores for each neuron. The distributions of z-scores are depicted in Figure 6.33. A significant effect of stimulus class was found for the distributions of z-scores (RM Anova; $F(3, 243) = 4.25$, $p < 0.0060$), and z-scores calculated in response to BOS were significantly greater than those calculated for WN-S (paired t-test, $p = 0.015$) and qBOS (paired t-test, $p = 0.016$), and almost significantly different for WN-E (paired t-test, $p = 0.0189$).

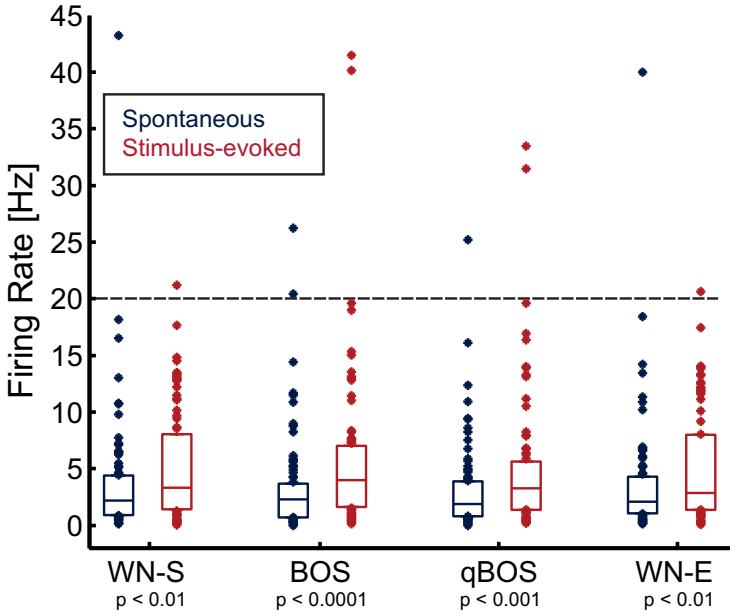

FIGURE 6.32: **Average firing rates for neurons included in the population analysis.** Box plots show the firing rate responses for the 82 neurons included in the population analysis. Spontaneous firing rates (dark blue) are significantly lower than the stimulus-evoked firing rates (red; Wilcoxon signed rank, p-values listed below). The line in the middle of the box represents the median value, the lower box edge represents the 25th percentile, and the upper box edge represents the 75th percentile. Responses outside of this range are plotted as dots. For each of the stimuli, only a few neurons were well-driven and had firing rate responses greater than 20 Hz (black dashed line).

6.1 Mean BOS and qBOS R_{SC} Values Not Significantly Different

We began by examining the stimulus-evoked spike count correlations for this population of neurons. Stimulus class did not have a significant effect on spike count correlations for this population of neurons (RM Anova, $F(3, 186) = 0.29$, $p = 0.84$), and none of the spike count correlations for any of the stimulus classes were significantly different from each other (paired t-test, $p > 0.05$; Fig. 6.34). Furthermore, none of the the mean spike count correlations for any of the stimuli were significantly different from zero (t-test, $p > 0.5$). This was also true for the larger population of auditory forebrain neurons, for which the mean spike count correlations for WN-S, BOS bouts, and WN-E were also not significantly different from zero (see Fig. 6.8).

6.2 Synchrony is Differentially Modulated by BOS and qBOS

In our larger population of auditory forebrain neurons, we observed that spike time cross-covariance values at short time lags were significantly lower for BOS bouts compared to other stimulus classes. We were curious whether spike-time correlations would

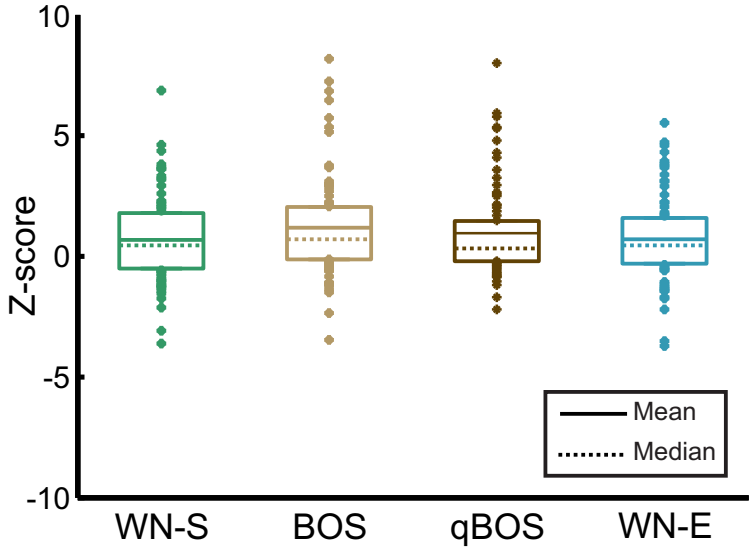

FIGURE 6.33: **Z-scores distributions for 82 auditory forebrain neurons.** Box plots show the z-scores calculated for each neuron for each stimulus type. Z-scores in response to BOS were significantly greater than WN-S and qBOS (paired t-test, p = 0.015 and 0.016, respectively). Lower box edge represents the 25th percentile, and upper box edge represents the 75th percentile. All neuron responses outside of this range are plotted as dots. The median z-scores (dotted line) tended to be lower than the mean values (full line).

also be lower for qBOS, so we analyzed the stimulus-evoked CCV functions separately for each stimulus class. Smoothed CCV functions for WN-S, BOS, qBOS, and WN-E are depicted in Figure 6.35.

A significant effect of stimulus class on CCV values at short time lags was found for this population of neurons (RM Anova, $F(3, 912) = 8.01$, $p < 0.0001$). Interestingly, we found that while mean CCV values at short time lags were still low in response to BOS bouts, this was not true of qBOS. The spike time cross-covariance evoked by BOS bouts were low and had a mean of 0.0012 ± 0.0008. The BOS CCV values at short time lags were significantly lower than WN (paired t-test, $p < 0.05$ for WN-S and WN-E) and strongly significantly lower than qBOS (paired t-test, $p < 0.0001$; Fig. 6.36). WN-S CCV values were not significantly different than WN-E CCV values (paired t-test, $p = 0.499$), and qBOS values were not significantly different than WN-S or WN-E CCV values (paired t-test, $p > 0.05$).

We shuffled identical trials to determine the effect that extrinsic stimulus correlations contributed to the CCV correlations. Smoothed trial-shuffled CCV functions are depicted in Figure 6.37. A significant stimulus effect was also found for the trial-shuffled CCV values at short time lags (RM Anova, $F(3, 912) = 6.42$, $p < 0.001$; Fig. 6.36; gray bars).

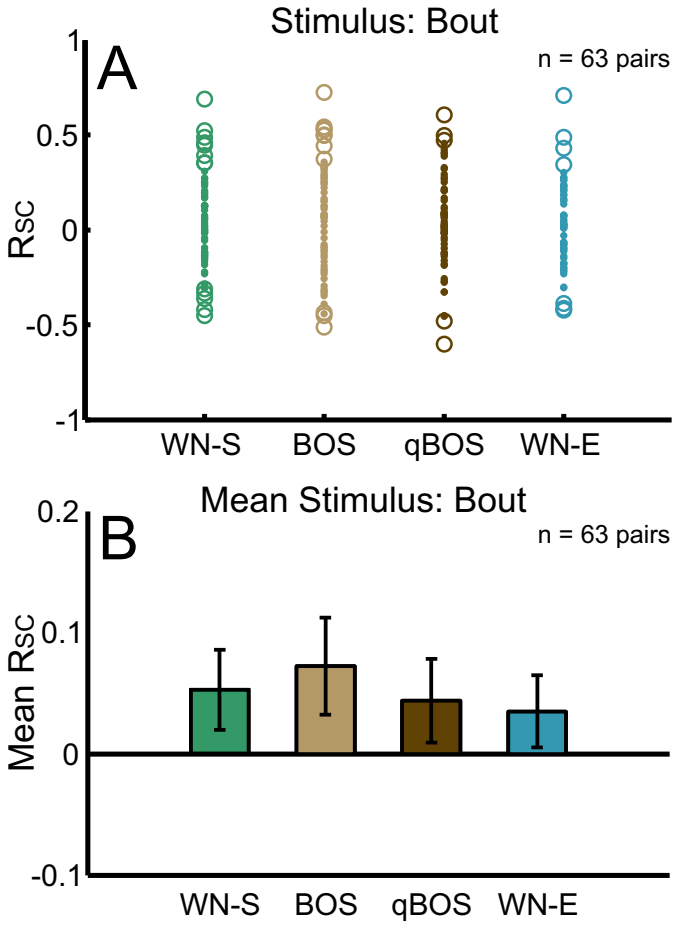

FIGURE 6.34: **Population spike count correlations during WN, BOS, and qBOS. A)** R_{SC} values were analyzed separately for each stimulus block (n = 63 pairs per block). Open circles indicated R_{SC} values that were statistically significant ($p < 0.05$), and dots indicated values R_{SC} values that were not significant ($p > 0.05$). **B)** Mean R_{SC} values per stimulus block. None of the mean spike count correlations for any of the stimuli were significantly different from zero. Error bars indicate the SEM. Ø indicates values that are significantly different than zero.

The trial-shuffled CCV mean for values at short time lags for BOS bouts was unique among all the stimuli in that, like observed for the larger population of auditory forebrain neurons, the mean was negative (mean = -0.0013 $\pm$ 0.000663). It was significantly lower than trial-shuffled CCV values at short time lags for both WN-S (paired t-test, $p < 0.01$) and WN-E (paired t-test, $p < 0.05$) and also significantly lower than trial-shuffled CCV values at short time lags for qBOS (paired t-test, $p < 0.05$; Fig. 6.36; gray bars). WN-S trial-shuffled CCV values were not significantly different than WN-E trial-shuffled CCV values (paired t-test, $p = 0.0686$), and qBOS values were not significantly different than WN-S or WN-E trial-shuffled CCV values (paired t-test, $p > 0.05$).

The mean noise-covariance, calculated by subtracting the mean trial-shuffled CCV from the mean CCV for each pair for each stimulus, was computed to determine if there are stimulus-specific effects of noise-covariance that contribute to CCV during auditory

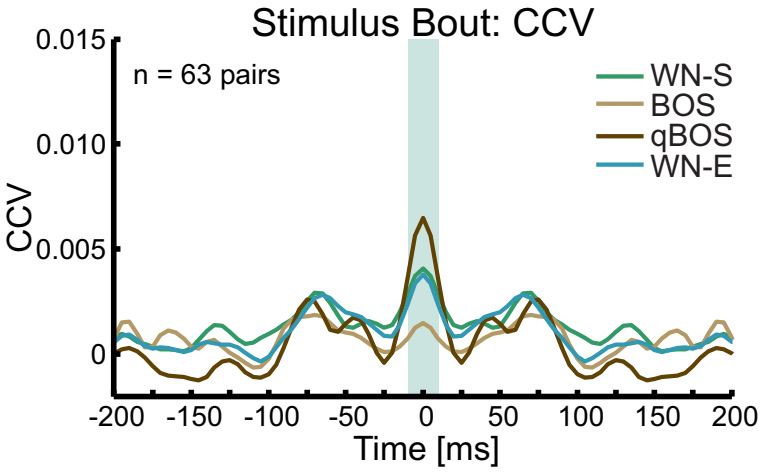

FIGURE 6.35: **Mean CCV functions for stimulus-evoked (bout) responses in Field L.** CCV functions were calculated for WN-S, BOS, qBOS, and WN-E separately. CCV functions for BOS had low values at short time lags, whereas CCV values for qBOS were high. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

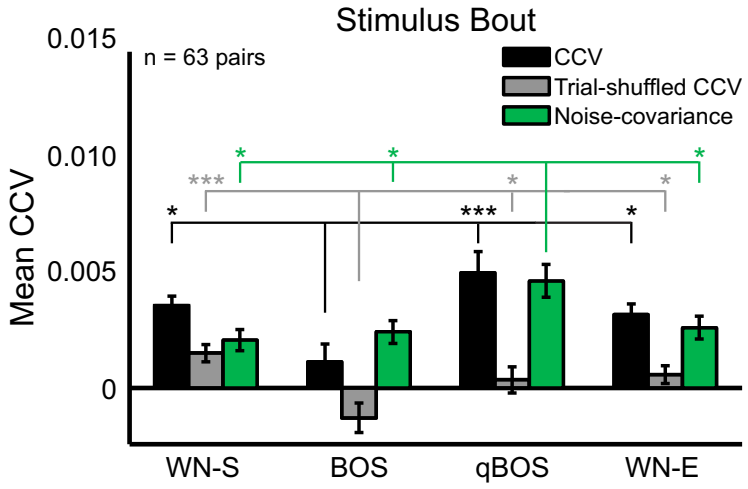

FIGURE 6.36: **Mean CCV values for stimulus-evoked (bout) responses in Field L.** Mean CCV (black bar), trial-shuffled CCV (gray bar) and noise-covariance (green bar). Mean CCVs and trial-shuffled CCVs in response to BOS were significantly lower than WN-S, qBOS, and WN-E. Mean noise-covariance in response to qBOS were significantly larger than WN-S, BOS, and WN-E. See text for details.

stimulation. Interestingly, stimulus class was found to have a significant effect on the noise-covariance (RM Anova, $F(3, 912) = 4.94$, $p = 0.0021$; Fig. 6.36; green bars).

Although all of the stimuli had noise-covariance values that were positive at short time lags, qBOS had significantly higher noise-covariance values compared to all of the other stimuli (paired t-test, $p < 0.05$; Fig. 6.36). In contrast, the noise-covariance values at short time lags for the other stimuli were not significantly different than each other (paired t-test, $p > 0.05$).

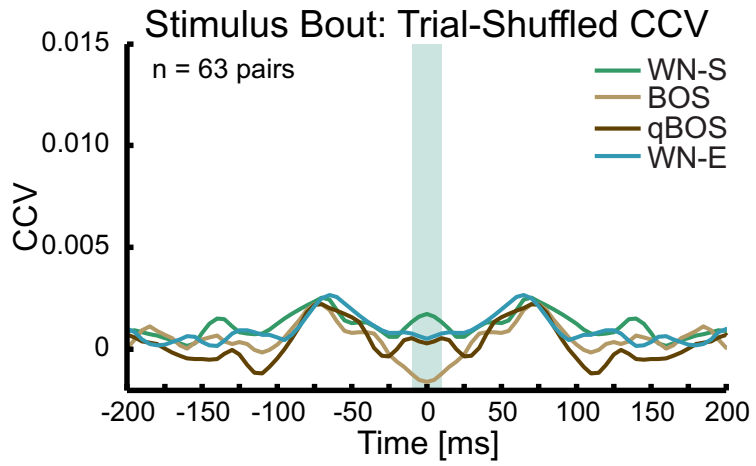

FIGURE 6.37: **Mean trial-shuffled CCV functions for stimulus-evoked (bout) responses.** CCV functions were calculated for WN-S, BOS, qBOS, and WN-E separately. CCV functions for BOS had low values at short time lags. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

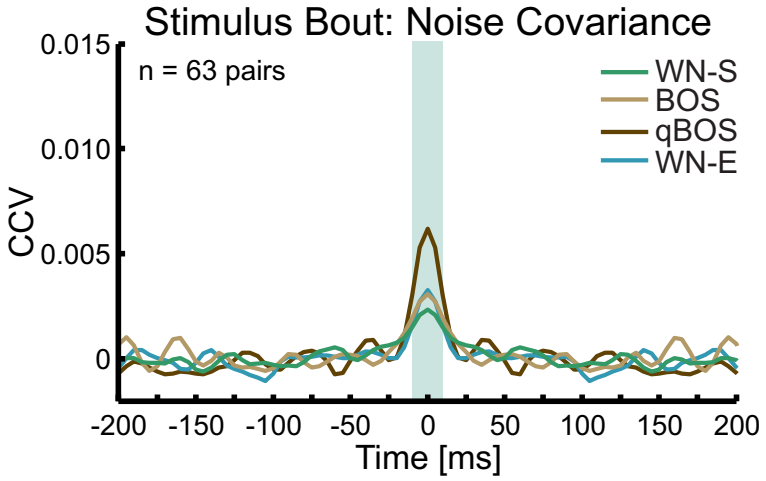

FIGURE 6.38: **Mean noise-covariance functions for stimulus-evoked (bout)responses.** CCV functions were calculated for WN-S, BOS, qBOS, and WN-E separately. Blue shading indicates the -10 ms to +10 ms of data that was used in the analysis.

6.3 Recapitulation

In a sub-population of auditory forebrain neurons, we examined the correlated responses to WN and bouts of BOS and qBOS. This population had a stimulus-evoked firing rate that was significantly higher for BOS versus qBOS (Fig. 6.32) and had significantly higher z-scores for BOS versus WN and qBOS (Fig. 6.33). While there was no significant difference in spike count correlations for any of the stimuli in this population of neurons, we observed differential spike time correlations for BOS versus qBOS.

Although the population of neurons had higher stimulus-evoked firing rates in response to BOS versus qBOS, CCV values at short time lags were significantly lower for BOS

than for WN and qBOS (Fig. 6.36; black bars). This trend was also true of trial-shuffled CCV values, for which BOS was again significantly lower than WN and qBOS (Fig. 6.36; gray bars).

Interestingly, we observed a significant effect of stimulus on the noise-covariance values at short time lags. Specifically, qBOS had significantly larger noise-covariance values at short time lags compared to any of the other stimuli (Fig. 6.38; Fig. 6.36; green bars).

Overall, these results suggest that although correlations in spike counts are not different between BOS and qBOS, the network responds differently to BOS and qBOS in terms of correlations in spike timing. Whereas loud versions of BOS evoked low spike time correlations at short time lags, quiet versions of BOS evoked larger spike time correlations at short time lags. These results indicate that neurons that are sensitive to the loudness of a stimulus could encode this information in the timing of spikes, rather than the number of spikes.

Chapter 7

Results: Efficient Coding Simulations

1 Simulations Using an Efficient Coding Strategy

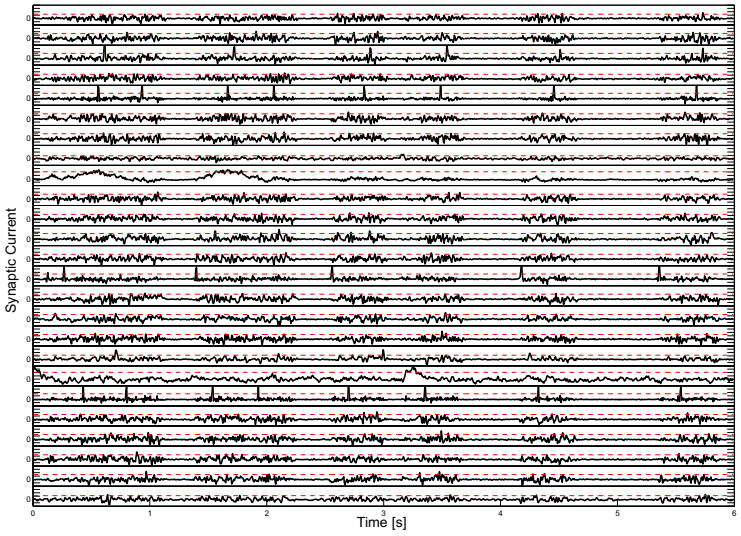

FIGURE 7.1: **Mean-subtracted synaptic currents for 25 simulated neurons.** Simulated synaptic currents for 25 neurons in response to a BOS song. Red dashed line indicates $\Theta = 3$.

We were curious whether the spike train cross covariance results we observed in our population of auditory forebrain neurons could be confirmed by a recently proposed efficient coding strategy [Blättler and Hahnloser, 2011]. In order to explore this, we simulated 11 different networks of 1000 neurons each using songs from 11 different birds. In each case we trained the network on a BOS to CON ratio of 5:1. See Methods section for details.

Once we had trained the network, we analyzed synaptic current responses from 1000 neurons in response to 10 different stimulus versions for 5 different stimulus classes: BOS, REV, CON, qBOS, and WN. The synaptic current from each neuron in the network was thresholded to create spiking responses. We used thresholds of 0, 0.5, 1, 2, 3, 4, 5, and 6 times the standard deviation of each neuron's mean-subtracted response. The mean-subtracted synaptic currents from 25 neurons in response to a BOS song are depicted in Figure 7.1.

Thresholding the synaptic currents had the effect of sparsifying the spiking responses as the threshold level increased. The spiking responses of the b1r10 network in response to one version of each of the different stimulus classes are displayed in Figure 7.2 for each of the thresholds.

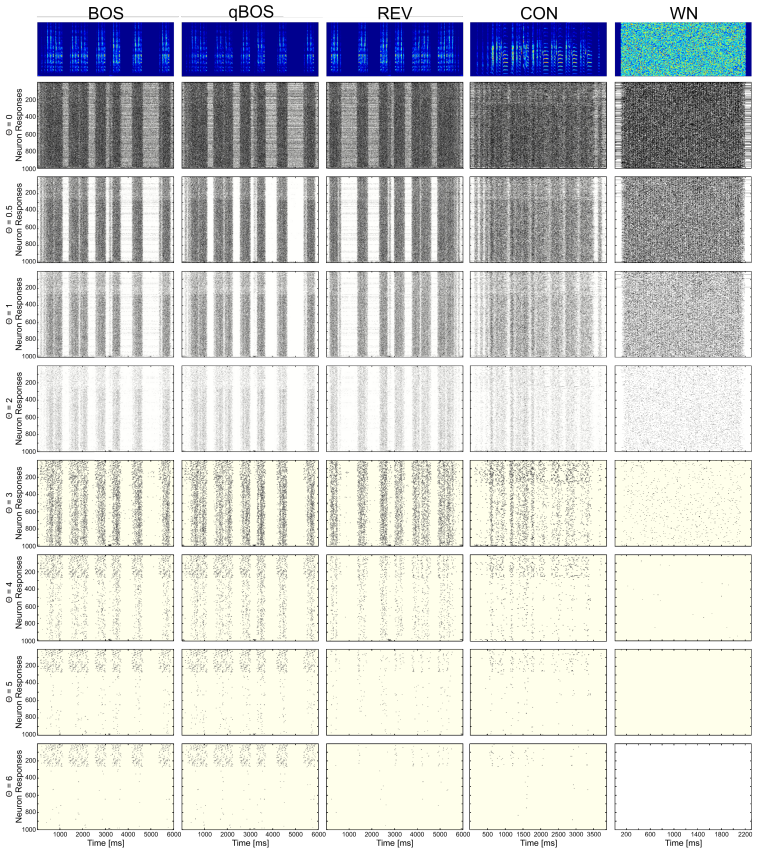

FIGURE 7.2: **Network spiking responses for b1r10.** Upper panels depict the spectrograms (from left to right) of a BOS song, qBOS, REV, CON, and WN. The thresholded responses of the network are discretized into 10 ms bins and depicted in the panels beneath the spectrograms as raster plots, where each black dot indicates a bin that contained a spike, and each row represents a different neuron in the network (n = 1000 neurons). Spiking responses become sparser at high threshold levels (Θ = 5, 6).

The contrast of the raster plots was increased for $\Theta \geq 3$ to improve visibility.

Qualitatively, spiking responses across the network were more homogeneous that what we observed *in vivo*. For example, in response to WN stimuli, the network typically

responded with strong tonic firing that did not capture the variety of heterogeneous spiking activity that we observed *in vivo* (see Fig. 6.3 for comparison). Similarly, network responses to BOS and qBOS responses were very similar to each other and did not seem to capture the differential responses of loud and quiet songs reported in [Nagel and Doupe, 2006, 2008], although the median network firing rate was lower in response to qBOS than BOS at low thresholds ($\Theta \leq 1$; see Figure 7.3). Considering that the network responses to qBOS and WN do not completely capture the characteristic responses we observed *in vivo*, we will largely focus our correlation analysis on the network responses to BOS, REV, and CON.

At high thresholds ($\Theta \geq 5$) most stimulus classes elicited very few spikes. For example, in response to a WN version, only 1 % of neurons (10 out of 1000) in the b1r10 network responded to WN at $\Theta = 5$, and even fewer (3 out of 1000) responded at $\Theta = 6$. For all stimulus classes, network firing rates decreased sharply at $\Theta s > 4$. Median network firing rates for b1r10, calculated on pooled spiking responses to each of the stimulus version per stimulus class, are presented in Figure 7.3.

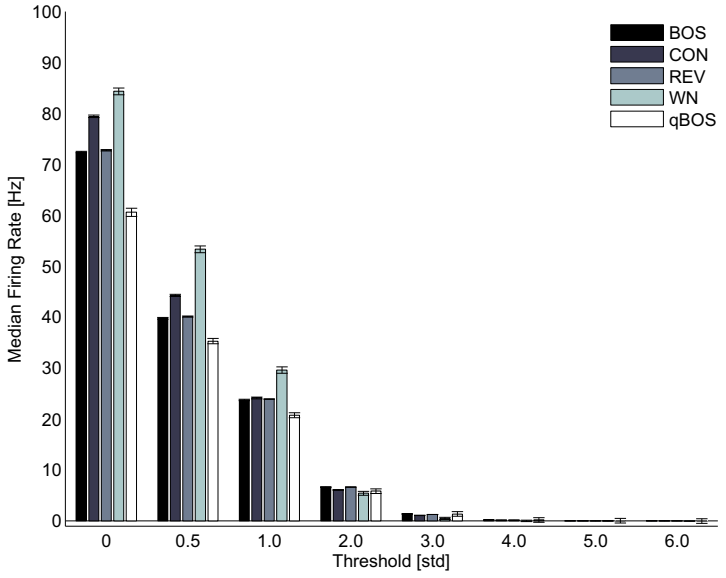

FIGURE 7.3: **Median network firing rates for b1r10.** Median network firing rates for all stimulus classes decrease sharply for $\Theta s > 4$. Error bars indicate the SEM.

We also calculated d-prime scores for each neuron in the network. D-prime scores for BOS-related comparisons of neurons in network b1r10 are displayed in Figure 7.4. Generally, the b1r10 network became more selective to BOS than other comparison stimuli as the firing rate threshold increased. The network became more selective to BOS than CON at $\Theta s \geq 2$, more selective to BOS than REV at $\Theta s \geq 3$, and more selective to BOS than WN at $\Theta s \geq 2$. At low firing rate thresholds, this network was more selective to BOS than qBOS at $\Theta s \leq 2$, but as the firing rate threshold increased,

the median d-prime score BOS and qBOS stabilized at around 0 indicating no network selectivity between the two stimuli.

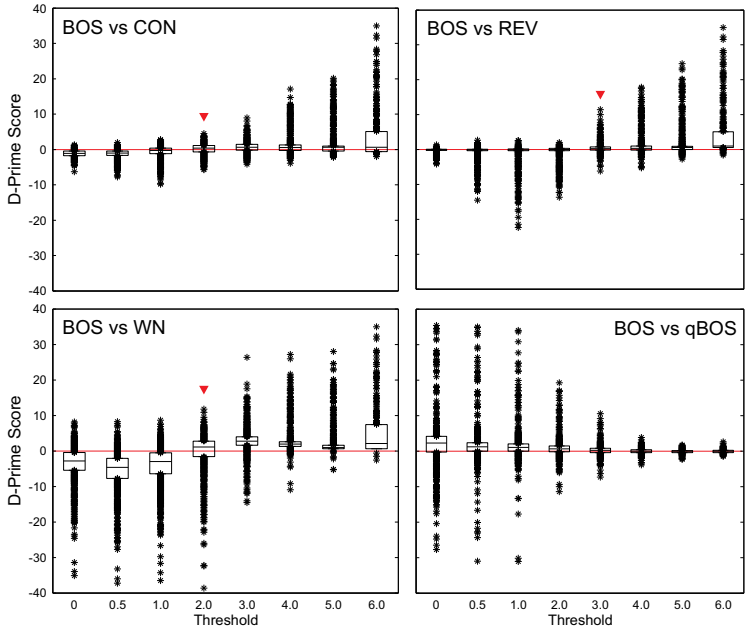

FIGURE 7.4: **D-prime scores for network b1r10.** Box plots indicate d-prime scores for each neuron in the b1r10 network per threshold. Lower box edge represents the 25th percentile, and upper box edge represents the 75th percentile. All d-prime scores outside of this range are indicated as *. The central box line indicates the median d-prime score, and the red line indicates d-prime = 0 (no stimulus selectivity). The red arrowhead indicates network selectivity switching (see text for details). No arrowhead is plotted for the BOS vs qBOS comparison.

2 Efficient Coding Strategy Captures Spike Time Correlations

We calculated pairwise CCVs between every neuron in the network and 10 randomly selected, non-identical members of the network. We first discretized the spike times into 10 ms bins and then then calculated pairwise CCVs for each of the 10 stimulus version per stimulus class and at each firing rate threshold, using the same methods as for our auditory forebrain responses. We then calculated an average CCV function per pairwise comparison by averaging the CCVs over the stimulus versions within a stimulus class. These calculations left us with 10,000 mean CCV functions per stimulus class per firing rate threshold. We pooled over these mean CCV functions to calculate mean network response to each of the stimulus classes for each firing rate threshold. An example of the mean network CCVs for b1r10 are displayed in Figure 7.5.

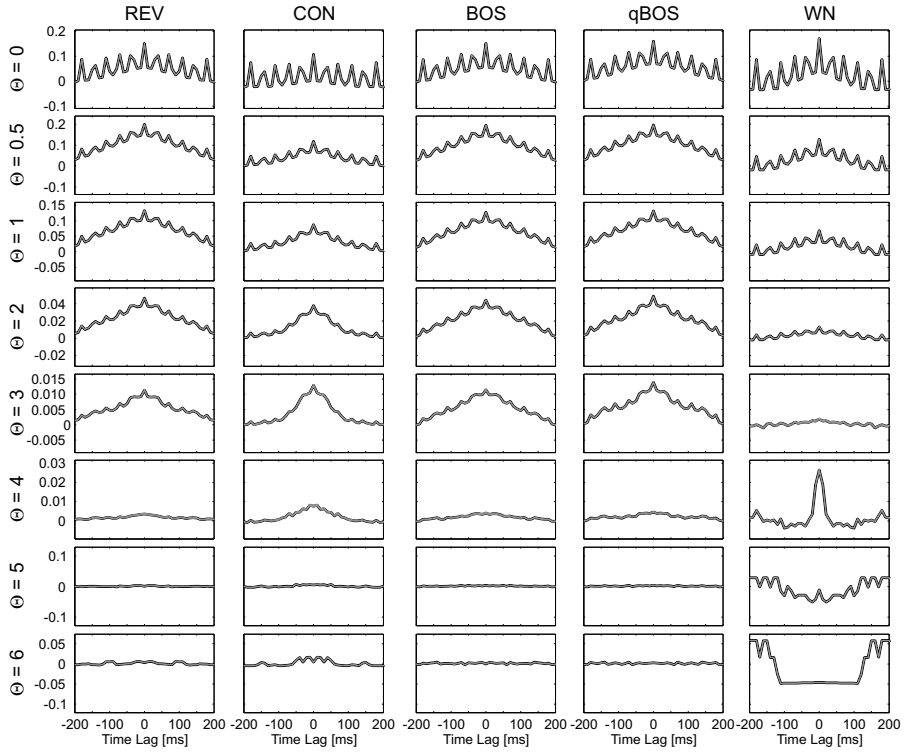

FIGURE 7.5: **b1r10: Mean CCV functions for REV, CON, BOS, qBOS, and WN.** At low thresholds ($\Theta = 0$), mean CCV functions show marked periodicity, whereas at higher thresholds ($\Theta > 2$), CCV functions can be characterized by a large, central peak. CCV functions are smoothed with the matlab "loess" option using a span of 4 data points (equivalent to 40 ms).

In order to determine which threshold was associated with correlation behavior most similar to that in our own population of auditory forebrain neurons, we compared the firing rates and the d-prime selectivity scores for each of the 1000 neurons in the networks. We calculated mean stimulus class firing rate responses for each neuron in each of the 11 networks, and then calculated a median population firing rate for the 11,000 neurons included in the simulation analysis. The median firing rates for each stimulus class per firing rate threshold are displayed in Figure 7.6A. As expected, the network firing rates decreased as the firing rate threshold increased. There was not a simulated firing rate threshold that perfectly matched the median firing rates from our population of auditory forebrain neurons. Based solely on median firing rates, our firing rate data aligns with network activity between the simulated firing rate thresholds of 2 and 3.

In addition to population firing rates, we also compared the mean network selectivity using the d-prime score. We calculated mean d-prime scores for each neuron in each network, and then pooled over all cells in all networks to calculate a mean d-prime score per firing rate threshold. The mean d-prime scores for BOS versus CON and BOS versus REV are depicted in Figure 7.7. As the firing rate threshold increases (and the median firing rates decrease, see Fig. 7.6), the mean BOS selectivity of the network

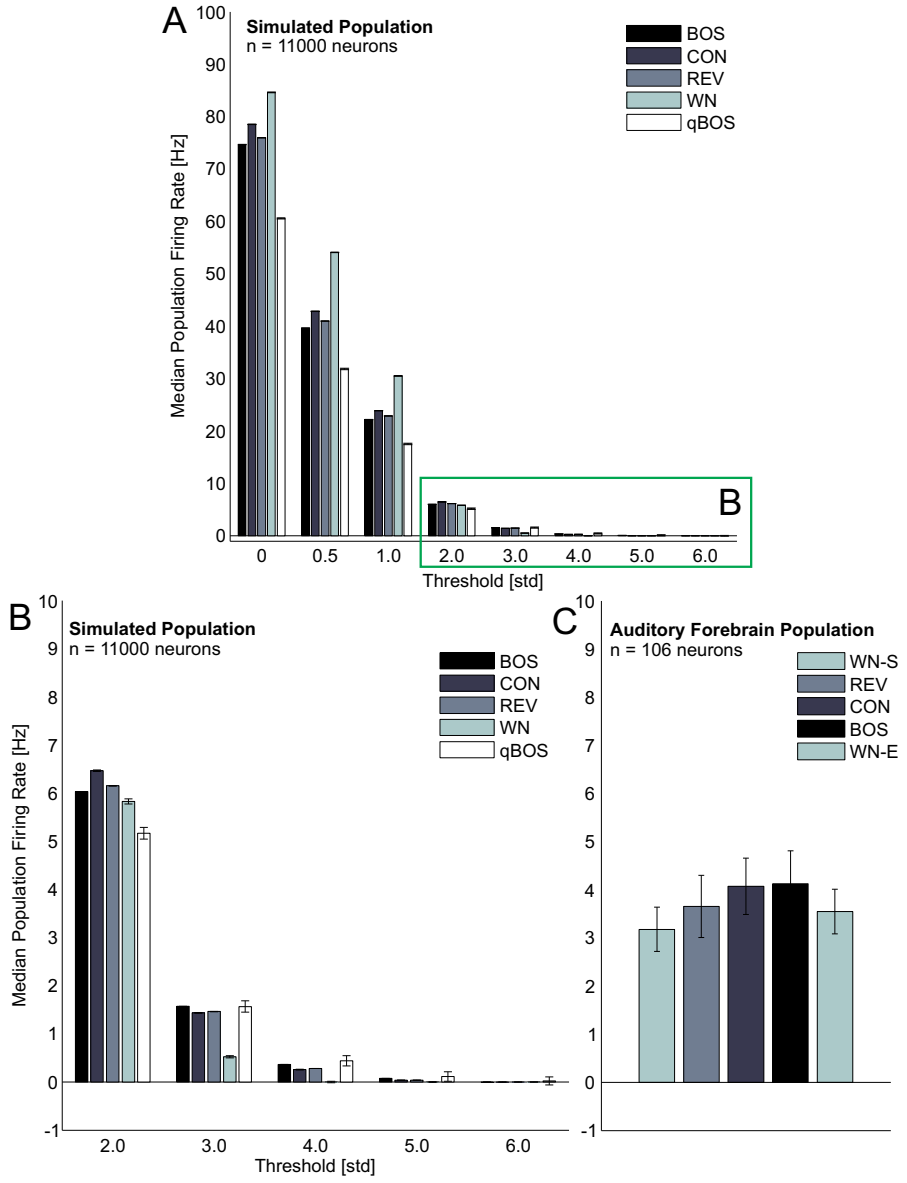

FIGURE 7.6: **Simulated population firing rates as a function of threshold. A)** Median simulated population firing rates for each stimulus class decrease as the firing rate threshold increases. Green box indicates thresholds which are plotted at a larger resolution in **B**. **B)** Median simulated population firing rates for $\Theta \geq 2$. At Θs > 3, the median population firing rate approaches zero. **C)** Median population firing rates for the auditory forebrain population of 106 neurons. Error bars represent the SEM.

increased. Although as for the network firing rates, there was not a simulated firing rate threshold that perfectly matched the both the BOS-CON selectivity and the BOS-REV selectivity, our auditory forebrain population selectivity aligns with network activity at $\Theta = 3$.

In order to compare the results of the simulations with our own results from the songbird auditory forebrain, we pooled the 10,000 mean CCVs across all 11 networks for each stimulus and each threshold, or a total of 110,000 CCVs per stimulus per threshold. We

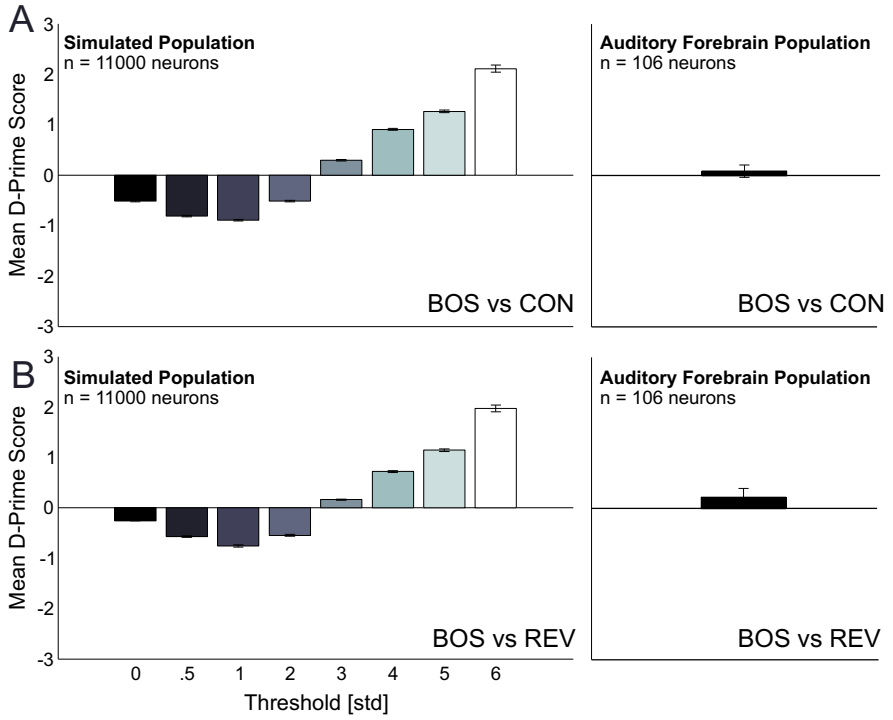

FIGURE 7.7: **Simulated population selectivity as a function of threshold. A)** Mean population d-prime scores for BOS versus CON per threshold for the simulated networks (left) and for the auditory forebrain population (right). **B)** Mean population d-prime scores for BOS versus REV per threshold for the simulated networks (left) and for the auditory forebrain population (right). Error bars represent the SEM.

used the CCV data from short time lags in our subsequent analysis. Specifically, we used the data from -10 ms to +10 ms for each of the CCV functions, which amounted to 3 data points from each pair: -10 ms bin, 0 ms bin, and the +10 ms bin. Based on median network firing rates and mean selectivity scores, we analyzed CCVs for Θs = 2 : 5 (Fig. 7.8).

2.1 Spiking Synchrony is Reduced for BOS for Sparsely Firing Networks

For our population of 106 auditory forebrain neurons (77 neuron pairs), we observed that the neuron pairs were less strongly correlated at short time lags for BOS than for WN, REV, and CON, as evidenced by BOS CCVs from -10 to +10 ms that were significantly smaller than the other stimuli (Fig. 7.8B).

We evaluated the networks at thresholds yielding median firing rates and mean stimulus selectivities that were similar to those we observed in the auditory forebrain. Mean CCVs at short time lags for Θs = 2 : 5 are depicted in Figure 7.8A. Although the CCV values for WN and qBOS are depicted in the figure, we will focus our discussion on the BOS, REV, and CON stimulus classes.

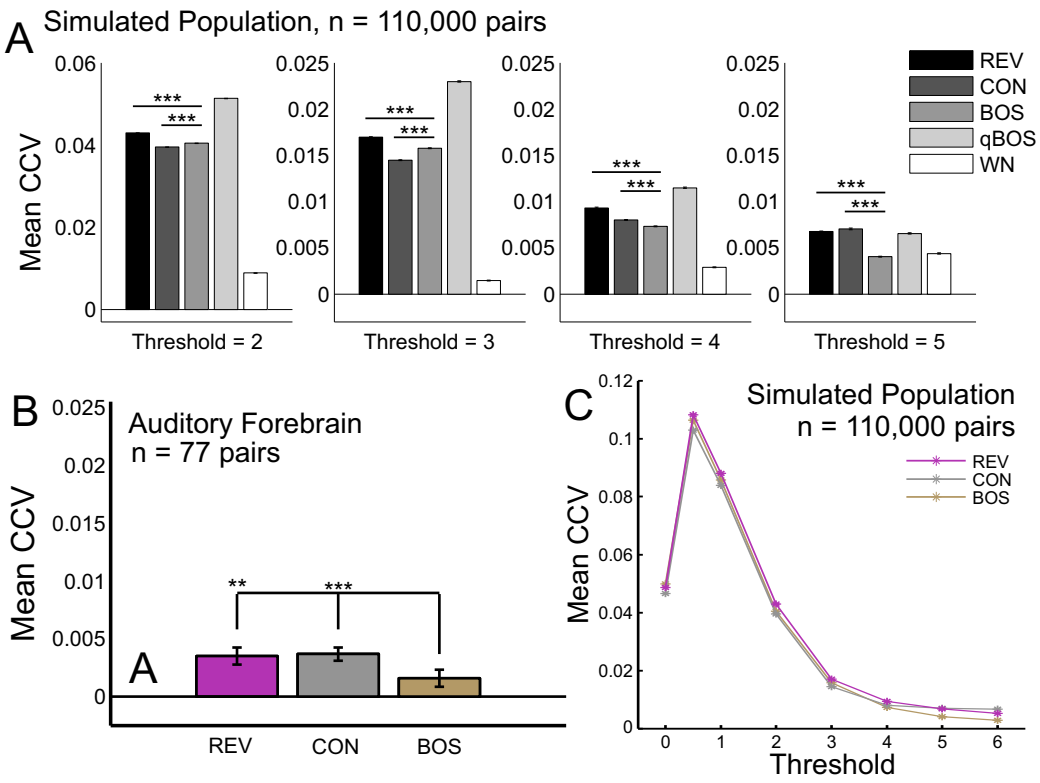

FIGURE 7.8: **BOS CCV values at short time lags are reduced for sparsely-firing, BOS-selective networks. A)** Mean CCVs from -10 to +10 ms lags for the simulated population of 110,000 neuron pairs for Θ = 2 (left) to Θ = 5 (right). At high firing rate thresholds $\Theta \geq 4$, BOS CCV values became significantly smaller than for REV or CON. The CCV values at Θ = 5 most closely resemble the CCV values at short time lags for our auditory forebrain population **(B)**. *** indicates highly significant statistical differences (paired t-test, $p \cong 0$). Error bars represent the SEM. **B)** CCVs from -10 to +10 ms lags for the auditory forebrain population of 106 neurons (77 pairs). BOS CCVs are significantly smaller than those for REV (paired t-test, $p < 0.01$) and CON (paired t-test, $p < 0.001$). **C)** Mean CCVs from -10 to +10 ms lags for REV, CON, and BOS for all thresholds. Although CCV values for BOS are lower than REV for all thresholds, CCV values for BOS become more negative than CON between the thresholds of 3 and 4. Error bars represent the SEM.

When evaluated at Θ = 2, which evoked median firing rate activity that was slightly higher than our auditory forebrain population and network activity that favored CON and REV over BOS, CCV values were an order of magnitude larger than those we obtained in the auditory forebrain. Whereas BOS CCV values at short time lags were significantly different from the other stimulus classes (paired t-test, $p \cong 0$), BOS was only significantly smaller than REV, and CON CCV values were significantly smaller than both REV and BOS.

When the networks were evaluated at Θ = 3, which evoked median firing rate activity that was slightly lower than our auditory forebrain population and network activity that

closely matched the selectivity we observed in our population of auditory forebrain neurons, CCVs were in the range of the values that we observed in the auditory forebrain. Matching the trend observed at $\Theta = 2$, BOS CCV values at short time lags were significantly different from the other stimulus classes (paired t-test, $p \cong 0$), but BOS was only significantly smaller than REV, and CON CCV values were significantly lower than both REV and BOS.

At $\Theta = 4$, the median firing rate activity was lower than what we observed in our population, but the network activity favored the BOS stimuli over CON and REV. At $\Theta \geq 4$, we observed a trend that continued for $\Theta = 5$ and $\Theta = 6$. As the population activity became sparse (low firing rates) and BOS-selective, the CCV values at short time lags for BOS became significantly smaller compared to the other stimuli. This trend started at $\Theta = 4$ and continued to $\Theta = 6$ (Fig. 7.8C).

2.2 Recapitulation

We evaluated spike train cross-covariances in 11 simulated networks of 1000 neurons each to see if our observation of reduced BOS CCVs at short time lags extends to an efficient coding model of the auditory forebrain. To this end, we tested each network with 10 versions each of REV, CON, and BOS, qBOS, and WN, and calculated CCV functions for each neuron in the network paired with 10 other randomly chosen neurons in the network, for a total of 10,000 pairwise comparisons. At low thresholds, CCV functions were characterized by marked periodicities which became less pronounced at high thresholds (Fig. 7.5).

By evaluating median network firing rates (Fig. 7.6) and mean network d-prime scores (Fig. 7.7), we selected a range of firing rate thresholds that most closely resembled the activity that we observed in the auditory forebrain *in vivo* and examined mean CCVs at short time lags (-10 to +10 ms) as a function of threshold. For all birds, BOS CCVs at short time lags were reduced as the firing rate threshold increased. This was also true when we pooled CCVs over all networks (Fig. 7.8). In other words, as the BOS selectivity of the network increased - despite an overall reduction of median network firing rates - the network activity became more decorrelated at short time lags in response to the BOS stimuli.

Overall, these results suggest that stimulus-selectivity and stimulus-specific decorrelation at short time lags could be related phenomena. Although further simulations must be conducted to explore whether networks can be "biased" to show selectivity to other stimuli at high firing rate thresholds (e.g. REV, CON), and to determine whether these biased networks also show stimulus-selective decorrelation, these results provide a first

step in understanding how efficient codes could shape network activity to display both stimulus selectivity and decorrelation.

Part IV

Discussion

Chapter 8

Discussion and Synthesis

In this work, we recorded extracellularly from ensembles of neurons recorded simultaneously in the auditory forebrain of adult male zebra finches. We examined correlated activity between pairs of neurons firing spontaneously during silence and also during auditory stimulation with synthetic and natural stimuli. To examine correlated activity, we used two different correlation metrics: the spike count correlation, which measures correlated trial-to-trial variability of spike counts on large time scales, and the spike-time cross covariance, which measures the correlations in spike timing patterns between neurons at short time scales. The main goal of this project was to examine how different auditory stimuli might affect the correlation dynamics of pairs and of populations of neurons in the auditory forebrain of the zebra finch.

We used a diverse auditory ensemble to drive stimulus-evoked responses in a population of heterogeneously-tuned neurons. This approach was different that often used in other studies of population correlation dynamics. For example, in studies of spiking synchrony (i.e., analyses using variations of the spike-train cross correlogram) a shuffle predictor (similar to our trial-shuffled cross-covariance) is often calculated and directly removed from the cross-correlogram, eliminating stimulus-related correlation effects from the analysis [Eggermont et al., 1983; Huang and Lisberger, 2009; Kimpo et al., 2003; Palm et al., 1988]. Instead of removing the effect of the stimulus, we examined differential effects of the stimulus on correlation dynamics of pairs and populations of neurons.

In the following sections, we will discuss the results of our study in detail. For each section, we will first summarize our major findings and then discuss these findings by placing them in the context of what has been examined in other animal models and sensory modalities.

1 Spike Count Correlations and Temporally Modulated Stimuli

We calculated spike count correlations over long (1.5 s) and short (200 ms) periods of stimulus-evoked activity and spontaneous firing. These durations sample the range of stimulus durations that are typically measured for spike count correlations calculations, which range from a few hundreds of milliseconds to several seconds [Cohen and Kohn, 2011]. Regardless of whether we calculated spike count correlations over short or long durations, we found that there was no mean effect of stimulus class on spike count correlations. During spontaneous firing, the mean spike count correlation in a population of auditory forebrain neurons was low (R_{SC} = 0.051 $\pm$ 0.013; mean $\pm$ SEM) and not significantly different than spike count correlations during auditory stimulation.

Furthermore, for our population of auditory forebrain neurons, stimulus class did not significantly affect the mean spike count correlations calculated over short (200 ms) or long (1.5 s) durations of stimulus-evoked activity (RM Anova, $p > 0.05$), and post-hoc, Bonferroni-corrected, paired t-tests did not reveal significant differences in pair-ordered means between different stimulus-classes. However, when averaged over the population, REV and CON stimulus classes produced mean spike count correlations that were significantly different than zero (t-test, $p < 0.05$), whereas in contrast, the mean spike count correlations for both WN and BOS were not significantly different than zero (t-test, $p > 0.05$).

Although stimulus class did not appear to influence spike count correlations when averaged over the population, individual pairs of neurons did show significant responses that depended on the stimulus class. For example, we often observed that a neuron pair had a significant spike count correlation ($p < 0.05$) in response to one stimulus class, such as BOS, but not for a different stimulus class like REV or CON. We observed a specific example of this when we examined spike count correlations between spatially nearby neurons. In response to BOS and qBOS, we found some nearby neuron pairs showed significant negative spike count correlations, which were not observed for CON, REV, or WN stimuli. In a larger population of auditory forebrain neurons that were not near each other, significant negative spike count correlations were also observed in response to WN. We suspect that by averaging over a population of neurons, and by including all neuron pairs in our analysis regardless of of the Pearson's correlation coefficient p value (see Methods), it is likely that these significant pair-effects were averaged out over the population.

Why would we observe that several pairs had negative spike count correlations in response to BOS and WN stimuli? Statistically speaking, BOS and WN are very different

stimuli. Whereas WN is relatively flat in terms of temporal amplitude modulations, BOS (and other birdsong) has a very rich amplitude modulation envelope. We observed that many neurons in the auditory forebrain were suppressed by WN stimuli (see Fig. 6.3D, F), and negative spike count correlations could arise from the pairing of such a WN-suppressed neuron with another neuron that is driven by WN. However, the prevalence of negatively correlated BOS responses is less easily interpretable, and most likely has to do with the specific tuning properties of the paired neurons.

We examined whether spike count correlations for individual neuron pairs might reflect the tuning similarity between the neurons of the pair. We examined the receptive fields of pairs of neurons that showed significant spike count correlations in response to CON (n = 12 pairs; Fig. 6.11). We did find several pairs of neurons (n = 5) that had significant positive spike count correlations and positive STRF similarity indexes, indicating that there were overlapping regions of the STRFs for the two neurons that might explain increased covariation in firing rate. However, we also found several pairs that had significant positive spike count correlations and anti-correlated receptive field tuning (n = 4), meaning that a pair of neurons still tended to covary their firing rates together, even though the neurons were oppositely tuned. We hypothesized that these results (correlated spike counts and anti-correlated tuning) could be explained by considering the fact that the stimuli we used to drive neural responses were spectro-temporally rich signals that often varied faster than the durations on which we calculated spike count correlations (1.5 s or 200 ms). Therefore, it is likely that spectro-temporal features that drove both, oppositely tuned neurons, could be present within these large calculation windows.

A toy example of this phenomenon would be to examine the spike count correlations of a pair of neurons, one of which is tuned to a stimulus onset, and one of which is tuned to a stimulus offset. Assuming that a song syllable is around 50 ms long, the onset neuron will fire at the onset of the syllable, and the offset neuron will fire at the offset of the syllable around 50 ms later. Both neurons will increase their firing in response to these stimulus features, and when spike counts are compared over time durations greater than 50 ms, it appears that these two neurons covary their firing rates together, despite the difference in their tuning properties.

How can we relate our findings to other studies that have investigated spike count correlations? Although we chose two durations over which to calculate spike count correlations that represent the spectrum of durations that are often used in other studies [Cohen and Newsome, 2008], a major difference between our work and others' relates to the stimuli that we chose to use. Several studies investigating shared trial-to-trial variability explored the primate visual system, and calculated spike count correlations in response

to relatively stable visual stimuli, such as moving sine wave gratings [Kohn and Smith, 2005], moving random dot stimuli [Huang and Lisberger, 2009; Zohary et al., 1994], or visual fixation tasks [Mitchell et al., 2009]. Other studies investigate correlations in the rodent somatosensory cortex using whisker stimulation [Montemurro et al., 2007] or whisker deflection [Middleton et al., 2012]. Such stimuli change relatively slowly over time and lack the highly modulated temporal structure characteristic of birdsong stimuli.

Therefore, we conclude that for fast, time-varying stimuli such as birdsong, spike count correlations calculated over durations larger than the speed at which the spectro-temporal features of the signal change simply represent stimulus-induced covariations in firing rate. We cannot conclude that neurons which covary their firing rates together over long time windows necessarily have similar tuning properties or are performing similar computations. Therefore, if neurons in the auditory forebrain use covariation in firing rate to encode information about auditory stimuli, this must involve integration windows even shorter than the 200 ms that we used in our analysis. We believe a more realistic strategy is that neurons in the songbird auditory forebrain use covariations of firing rates over short integration times, shorter than the speed at which the spectro-temporal features of the signal change (which essentially amounts to spiking synchrony), to convey information about auditory stimuli.

In support of this hypothesis, the following sections will discuss our results pertaining to the spiking synchrony observed in our data.

2 Spiking Synchrony Is Reduced by Auditory Stimulation

We found that spike train cross covariances at short time lags ($\pm$ 10 ms) were significantly and strongly reduced by auditory stimulation compared to spontaneous activity during silence. This finding was true of both a larger population of auditory forebrain neurons (n = 77 neuron pairs) and a smaller population of Field L neurons (n = 45 neuron pairs).

Reduced spiking synchrony in response to sensory stimulation has been previously reported. Shaevitz and Theunissen [2007] recorded from pairs of Field L-HVC neurons and pairs of CLM-HVC neurons during spontaneous and stimulus-evoked states in the zebra finch. Using shuffle-corrected cross-coherency (similar to our noise covariance calculations), the authors found significant coherency peaks during spontaneous firing

that were absent during stimulus-evoked activity (playbacks of BOS, REV, reversed-order BOS, and CON). Specifically, of 32 Field L-HVC paired sites that showed a significant peak during spontaneous activity, only 9 were significantly correlated during stimulus evoked activity. A similar trend was true for CLM-HVC neuron pairs, and the authors observed no cases where a neuron pair was significantly correlated during stimulus-evoked firing and not also during spontaneous firing [Shaevitz and Theunissen, 2007].

Tomita and Eggermont [2005] recorded multi-unit activity in anesthetized cat auditory cortex during periods of auditory stimulation and during silence. They estimated STRFs for the neuron pairs, and based on overlapping regions of the STRFs, separated spikes into "STRF-IN" and "STRF-OUT" categories, where STRF-IN spikes contributed to the overlapping region of the STRFs shared between neurons, and spikes not contributing to the overlapping STRF region were considered STRF-OUT spikes. The authors used cross-correlations to examine spiking synchrony between pairs, and found that the peak correlation coefficient for STRF-OUT spikes was approximately 25% lower than those calculated during periods of silence. In contrast, peak correlation coefficients for STRF-IN spikes were not significantly different from those during silence [Tomita and Eggermont, 2005]. The authors concluded that stimulation reduces the correlation of background activity, and as a result, the signal-to-noise ratio of correlated activity in response to the stimulus is enhanced. They suggested that auditory stimulation disrupts large assemblies of synchronously firing neurons that exists during periods of silence, and transforms the network into many smaller stimulus-activated ensembles.

Middleton et al. [2012] also observed a stimulus-induced reduction of correlations between neuron pairs compared to spontaneous activity and investigated the circuit mechanisms that could be responsible for the observed effect. The authors recorded from regular-spiking (RS; presumably excitatory) and fast-spiking (FS; presumably inhibitory) neurons in layer 2/3 of the the rat barrel cortex during both spontaneous and whisker stimulus-evoked states. The authors observed that for RS-FS pairs, both neurons tended to fire together during the pre-stimulus period. However, during stimulus-evoked activity, the FS neuron was much more silent compared to the RS neuron, suggesting that that FS and RS neurons fire in coordinated fashion during spontaneous activity and their firing becomes more independent with whisker stimulation.

The authors calculated spike count correlations between pairs of simultaneously recorded neurons using 50 ms sliding time windows in 2 ms increments. They found that spike count correlations for FS-RS pairs was significantly reduced during whisker deflection compared to spontaneous firing (paired t-test, $p < 0.001$; [Middleton et al., 2012]). There was no significant difference for RS-RS pairs of neurons (paired t-test, $p = 0.54$).

Through experiments using a model that accurately described the instantaneous firing rates of RS and FS neurons during spontaneous and stimulus-evoked conditions, the authors concluded that the reduced correlation during whisker stimulation could have resulted from an interaction between positively correlated effects resulting from common input to both populations during spontaneous activity and the anti-correlating effects of feedforward inhibitory coupling of the FS population to the RS population during stimulus-evoked states.

Unfortunately, in our study we were not able to determine whether neurons were excitatory or inhibitory based on their extracellular spike shapes. However, studies have shown that high densities of cells in Field L are GABA-positive [Pinaud and Mello, 2007]. Interestingly, distinct populations of GABA-positive neurons were found in Field L, with small-bodied inhibitory neurons located in L1 and L3, whereas large-bodied, GABAergic neurons were predominant in L2a and might serve as projection neurons. Furthermore, 42% of cells that expressed *zenk* in response to song stimulation also expressed *zGAD65* (the zebra finch homologue of the glutamic-acid decarboxylase, the enzyme that synthesizes GABA), and are therefore GABAergic [Pinaud et al., 2004]. Older experiments performed in the auditory forebrain of a non-songbird, the chicken, have shown that inhibitory GABAergic interactions may mediate local inhibition limiting response strength and lateral inhibition responsible for sharpened tuning to stimulus frequencies [Müller and Scheich, 1987, 1988]. Overall, these results provide evidence that GABAergic neurons participate in auditory responses to birdsong, and furthermore, that it is not unreasonable to assume that some of the neuron pairs from which we recorded included inhibitory neurons.

Therefore, a mechanism similar to that proposed by Middleton et al. [2012] may underlie the effect that we observed of auditory stimulation decreasing correlations at short time lags. Specifically, during silence, neuron pairs receive common synaptic input that correlates their activity. During auditory stimulation, both excitatory and inhibitory neurons receive feed-forward, stimulus-related input that drives firing, but excitatory neurons also receive direct feed-forward inhibition from local inhibitory neurons, resulting in an overall decorrelated state during stimulus-evoked conditions. Such a circuit, coupled with a threshold non-linearity, was necessary and sufficient to capture the decorrelation observed for FS-RS neuron pairs during whisker stimulation [Middleton et al., 2012], and we surmise that, given the prevalence of GABAergic neurons in the auditory forebrain of songbirds, a similar role of feedforward inhibition may also explain the effects we observed of decorrelated activity in response to auditory stimulation.

3 Spiking Synchrony is Reduced By BOS but not qBOS

We observed a significant difference between spiking synchrony in response to BOS versus qBOS. CCV and noise-covariance values at short time lags ($\pm$ 10 ms) were significantly larger for qBOS compared to BOS in a population of 63 auditory forebrain pairs. In contrast, spike count correlations calculated over long time periods (1.5 s) were not significantly different than zero, nor was there a significant difference between BOS or qBOS.

Although a study examining the effect of correlations in response to loud and quiet stimulus playbacks has not been reported, Kohn and Smith [2005] performed a study in the primary visual cortex of macaque examining the effect of stimulus contrast on correlation strength. Specifically, they measured the effect of reducing stimulus contrast on correlations, a manipulation that would be expected to alter the strength of the evoked cortical response rather than to alter which neurons are recruited by the stimulus. Using a similar logic, we compared BOS responses to a high-contrast condition, and qBOS responses to a low-contrast condition.

Kohn and Smith [2005] found that spike count correlations (calculated on 2.56 s of data) were significantly larger for intermediate and low contrast stimuli than for high contrast stimuli. Although not significant, our mean spike count correlations values showed an opposite trend: mean spike count correlations were larger for BOS than for qBOS. The larger mean spike count correlation for BOS compared to qBOS may result from the the fact that the population of neurons had a significantly higher z-score in response to BOS compared to qBOS (paired t-test, $p = 0.016$). It has been shown that spike count correlations increase with firing rate [de la Rocha et al., 2007], and because neurons tended to fire more in response to BOS than qBOS, the larger spike count correlation resulting from BOS playback may simply reflect the larger BOS-evoked firing rate.

Kohn and Smith [2005] also examined the effect of spiking synchrony as a function of contrast. Using spike train cross-correlograms corrected for stimulus-induced correlations by subtracting a shift predictor (similar to our noise-covariance functions), the authors found that the peak height was reduced in amplitude for low contrast stimuli, but that there was a compensatory increase in peak width. These results stand in contrast to the significantly larger peak height we observed for noise-covariance functions during the "low contrast" qBOS stimulation. Furthermore, the peak width of BOS, not qBOS, was visibly broader for the mean noise-covariance functions.

The differences in our findings in the auditory forebrain of the songbird compared to the primary visual cortex of primate might have to do with the fact that neurons in Field L have been shown to have different receptive field properties in response to loud and

quiet stimuli [Nagel and Doupe, 2006]. Nagel and Doupe [2008] estimated receptive fields of auditory neurons in Field L using loud (63 dB) and quiet (30 dB) playbacks of a rich synthetic stimulus, and found that STRFs obtained with high stimulus intensities had shorter response latencies, broader and more prominent inhibitory sidebands, and broader excitatory sidebands. In contrast, STRFs obtained using low stimulus intensities had longer response latencies and narrower tuning, with smaller inhibitory and excitatory sidebands [Nagel and Doupe, 2008].

These findings suggest that at high stimulus intensities, neural responses are more broadly tuned, and neurons respond to a larger range of spectro-temporal stimulus combinations. That could mean that during loud stimuli like BOS, broader receptive field tuning could produce many more (noisy) spikes that do not necessarily align in time, leading to the reduced spiking synchrony observed for BOS. In contrast, during quiet stimuli like qBOS, the receptive fields do not have such broad tuning, and therefore the resulting spike times may be more aligned in time, leading to the large spiking synchronization observed for qBOS. This differential processing of loud and quiet birdsong stimuli by auditory forebrain neurons might give songbirds an advantage in discriminating the behavioral relevance of soft and loud songs, i.e., discriminating between the songs of nearby kin and distant neighbors or strangers.

However, we also observed that for spontaneous firing during silence, large and positive spiking synchrony was present at short time lags. It could also be that the correlated responses that we observed in response to qBOS represent an intermediate network state between loud and silent stimulus conditions. In order to determine whether the increased noise-covariance observed in response to qBOS is a result of a specific behavioral meaning implicit in the quiet BOS signal, or rather just an effect of the playback condition, it would be interesting to perform an additional experiment that utilized a range of sound playback levels. If the noise covariance decreased systematically for playbacks ranging from quiet to loud conditions, we could clarify that the observed effect has to do with the network state of the brain.

4 Spiking Synchrony Is Reduced by BOS in a Population of Auditory Forebrain Neurons and in Field L Neurons

We examined the correlated firing of populations of neurons in response to a diverse stimulus ensemble. We found that there was not a significant effect of stimulus class on spike count correlations calculated over 1.5 s of the song bout or over 200 ms of the song motif. However, we observed a significant effect of stimulus class on spiking synchrony

at short time lags ($\pm$ 10 ms). Specifically, in a population of 106 auditory forebrain neurons (77 neuron pairs) CCV in response to BOS was strongly and significantly reduced compared to WN, REV, and CON playback. This trend was also observable in a smaller population consisting of only Field L neurons (62 neurons, 45 neuron pairs). The reduced CCV for BOS resulted from stimulus-evoked anti-correlation between neuron pairs, indicated by the negative values of the trial-shuffled CCV at short time lags in response to BOS.

We observed that such near-zero correlations around zero time lag could arise from broadly tuned neurons that are sensitive to sound onsets, such as the Field L neurons depicted in Fig. 5.19 and Fig. 5.20. These neurons have similar spectral tuning, but differ in their temporal response latencies. The differential response latencies mean that the neurons will rarely fire together in response to song stimuli, but rather that one neuron tends to lead another in firing, shifting the peak of the CCV and trial-shuffled CCV away from zero. Of course, neurons with nearly identical receptive field tuning will still tend to fire together, creating a prominent peak centered at zero time lag, such as that observed in Fig. 5.21.

How can a network of neurons show decorrelated responses to one birdsong stimulus (e.g., BOS), but not to another (e.g., CON)? Furthermore, to what extent does our effect of decorrelated BOS responses relate to the slightly positive stimulus selectivity that our population displayed in response to BOS?

Using an efficient coding algorithm, we found that as the selectivity for a particular stimulus increases in a population of neurons, the spiking synchrony at short time lags becomes significantly decorrelated, matching the effect that we observed in our population of auditory forebrain neurons. Or, stated differently, as the population response to a stimulus becomes sparsely selective, spiking synchrony becomes significantly decorrelated. Neural decorrelation has been linked to stimulus selectivity in Vinje and Gallant [2000]. In this study, the authors found that in primate visual cortex, stimulation of the non-classical receptive field increases the selectivity (*lifetime sparseness*) of individual V1 neurons, increases the sparseness of the population response distribution (*population sparseness*), and strongly decorrelates the responses of neuron pairs. These results mirrored the effects we see as we increase the firing rate threshold of our simulated networks: both the population and lifetime sparseness of the network increase in response to BOS, even as the network becomes more decorrelated in response to BOS at short time lags.

Information theory approaches have shown that whereas positive correlations shared between populations of neurons tend to decrease the information capacity of a network, negative correlations substantially increase the information capacity of the neural

population [Sompolinsky et al., 2001]. We suggest that the decorrelation of spiking synchrony in response to BOS may be the result of an efficient coding strategy that supports the formation of auditory receptive fields that decorrelate spiking responses to prevalently heard BOS but not to other, less frequently heard birdsong stimuli. Overall, these results highlight the role that synchronous and/or asynchronous spiking output may have in the discrimination of behaviorally relevant stimuli.

Bibliography

L. F. Abbott and P. Dayan. The effect of correlated variability on the accuracy of a population code. *Neural computation*, 11(1):91–101, Jan. 1999.

E. D. Adrian and Y. Zotterman. The impulses produced by sensory nerve-endings: Part II. The response of a Single End-Organ. *The Journal of physiology*, 61(2):151–71, Apr. 1926.

E. Akutagawa and M. Konishi. New brain pathways found in the vocal control system of a songbird. *The Journal of Comparative Neurology*, 518(15):3086–100, Aug. 2010.

N. Amin, J. A. Grace, and F. E. Theunissen. Neural response to bird's own song and tutor song in the zebra finch field L and caudal mesopallium. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, 190(6): 469–89, June 2004.

N. Amin, A. J. Doupe, and F. E. Theunissen. Development of selectivity for natural sounds in the songbird auditory forebrain. *Journal of neurophysiology*, 97(5):3517–31, May 2007.

N. Amin, P. Gill, and F. E. Theunissen. Role of the zebra finch auditory thalamus in generating complex representations for natural sounds. *Journal of neurophysiology*, 104(2):784–98, Aug. 2010.

D. Aronov, A. S. Andalman, and M. S. Fee. A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science (New York, N.Y.)*, 320(5876):630–4, May 2008.

S. Atiani, M. Elhilali, S. V. David, J. B. Fritz, and S. A. Shamma. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron*, 61(3):467–80, Feb. 2009.

D. Attwell and S. B. Laughlin. An energy budget for signaling in the grey matter of the brain. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 21(10):1133–45, Oct. 2001.

B. B. Averbeck and D. Lee. Neural noise and movement-related codes in the macaque supplementary motor area. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23(20):7630–41, Aug. 2003.

B. B. Averbeck, P. E. Latham, and A. Pouget. Neural correlations, population coding and computation. *Nature reviews. Neuroscience*, 7(5):358–66, May 2006.

W. Bair, E. Zohary, and W. T. Newsome. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 21(5):1676–97, Mar. 2001.

H. B. Barlow. Summation and inhibition in the frog's retina. *The Journal of physiology*, 119(1):69–88, Jan. 1953.

H. B. Barlow. Single units and sensation: a neuron doctrine for perceptual psychology? *Perception*, 1(4):371–94, Jan. 1972.

E. E. Bauer, M. J. Coleman, T. F. Roberts, A. Roy, J. F. Prather, and R. Mooney. A synaptic basis for auditoryvocal integration in the songbird. *The Journal of Neuroscience*, 28(6):1509, Feb. 2008.

M. J. Berry, D. K. Warland, and M. Meister. The structure and precision of retinal spike trains. *Proceedings of the National Academy of Sciences of the United States of America*, 94(10):5411–6, May 1997.

W. Bialek, F. Rieke, R. R. de Ruyter van Steveninck, and D. Warland. Reading a neural code. *Science*, 252(5014):1854–1857, June 1991.

M. Biederman-Thorson. Auditory evoked responses in the cerebrum (field L) and ovoid nucleus of the ring dove. *Brain research*, 24(2):247–56, Dec. 1970.

B. Bigalke-Kunz, R. Rübsamen, and G. J. Dörrscheidt. Tonotopic organization and functional characterization of the auditory thalamus in a songbird, the European starling. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, 161(2):255–65, Aug. 1987.

F. Blättler and R. H. R. Hahnloser. An efficient coding hypothesis links sparsity and selectivity of neural responses. *PloS one*, 6(10):e25506, Jan. 2011.

J. J. Bolhuis, G. G. Zijlstra, A. M. den Boer-Visser, and E. A. Van Der Zee. Localized neuronal activation in the zebra finch brain is related to the strength of song learning. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (5):2282–5, Feb. 2000.

T. Boumans, S. M. H. Gobes, C. Poirier, F. E. Theunissen, L. Vandersmissen, W. Pintjens, M. Verhoye, J. J. Bolhuis, and A. Van der Linden. Functional MRI of auditory responses in the zebra finch forebrain reveals a hierarchical organisation based on signal strength but not selectivity. *PloS one*, 3(9):e3184, Jan. 2008.

M. S. Brainard and A. J. Doupe. What songbirds teach us about learning. *Nature*, 417 (6886):351–8, May 2002.

M. Brecht and B. Sakmann. Dynamic representation of whisker deflection by synaptic potentials in spiny stellate and pyramidal cells in the barrels and septa of layer 4 rat somatosensory cortex. *The Journal of physiology*, 543(Pt 1):49–70, Aug. 2002.

E. A. Brenowitz. Altered perception of species-specific song by female birds after lesions of a forebrain nucleus. *Science (New York, N.Y.)*, 251(4991):303–5, Jan. 1991.

E. A. Brenowitz. Birdsong: integrating physics, physiology, and behavior. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 188(11):827–828, Dec. 2002.

E. A. Brenowitz and A. P. Arnold. Interspecific comparisons of the size of neural song control regions and song complexity in duetting birds: evolutionary implications. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 6(10): 2875–9, Oct. 1986.

E. A. Brenowitz and M. D. Beecher. Song learning in birds: diversity and plasticity, opportunities and challenges. *Trends in neurosciences*, 28(3):127–32, Mar. 2005.

C. D. Brody. Correlations without synchrony. *Neural computation*, 11(7):1537–51, Oct. 1999.

S. Cardoso de Oliveira, A. Thiele, and K. P. Hoffmann. Synchronization of neuronal activity during stimulus expectation in a direction discrimination task. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 17(23):9248–60, Dec. 1997.

C. E. Carr and M. Konishi. A circuit for detection of interaural time differences in the brain stem of the barn owl. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 10(10):3227–46, Oct. 1990.

C. K. Catchpole and P. J. B. Slater. *Bird Song: Biological Themes and Variations*. Cambridge University Press, Cambridge, 2 edition, 2008.

S. J. Chew, C. V. Mello, F. N. Nottebohm, E. D. Jarvis, and D. S. Vicario. Decrements in auditory responses to a repeated conspecific song are long-lasting and require two

periods of protein synthesis in the songbird forebrain. *Proceedings of the National Academy of Sciences of the United States of America*, 92(8):3406–10, Apr. 1995.

S. J. Chew, D. S. Vicario, and F. N. Nottebohm. A large-capacity memory system that recognizes the calls and songs of individual birds. *Proceedings of the National Academy of Sciences of the United States of America*, 93(5):1950–5, Mar. 1996.

G. B. Christianson, M. Sahani, and J. F. Linden. Depth-Dependent Temporal Response Properties in Core Auditory Cortex. *Journal of Neuroscience*, 31(36):12837–12848, Sept. 2011.

N. S. Clayton. The effects of cross-fostering on selective song learning in estrildid finches. *Behaviour*, 109(3):163–175, 1989.

M. R. Cohen and A. Kohn. Measuring and interpreting neuronal correlations. *Nature neuroscience*, 14(7):811–9, July 2011.

M. R. Cohen and W. T. Newsome. Context-dependent changes in functional circuitry in visual area MT. *Neuron*, 60(1):162–73, Oct. 2008.

C. Constantinidis and P. S. Goldman-Rakic. Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. *Journal of neurophysiology*, 88(6):3487–97, Dec. 2002.

S. J. Cruikshank, T. J. Lewis, and B. W. Connors. Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex. *Nature neuroscience*, 10(4):462–8, Apr. 2007.

Y. Dan, J. M. Alonso, W. M. Usrey, and R. C. Reid. Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nature neuroscience*, 1 (6):501–7, Oct. 1998.

P. Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press, Cambridge, 2001.

J. de la Rocha, B. Doiron, E. Shea-Brown, K. Josić, and A. D. Reyes. Correlation between neural spike trains increases with firing rate. *Nature*, 448(7155):802–6, Aug. 2007.

R. R. de Ruyter van Steveninck, G. D. Lewen, S. P. Strong, R. Koberle, and W. Bialek. Reproducibility and variability in neural spike trains. *Science (New York, N.Y.)*, 275 (5307):1805–8, Mar. 1997.

A. F. Dean. The variability of discharge of simple cells in the cat striate cortex. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 44(4):437–40, Jan. 1981.

G. C. DeAngelis, G. M. Ghose, I. Ohzawa, and R. D. Freeman. Functional micro-organization of primary visual cortex: receptive field analysis of nearby neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 19(10): 4046–64, May 1999.

S. Deneve, P. E. Latham, and A. Pouget. Reading population codes: a neural implementation of ideal observers. *Nature neuroscience*, 2(8):740–5, Aug. 1999.

R. Desimone and S. J. Schein. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of neurophysiology*, 57(3):835–68, Mar. 1987.

R. Desimone, T. D. Albright, C. G. Gross, and C. Bruce. Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 4(8):2051–62, Aug. 1984.

K. Diba, H. A. Lester, and C. Koch. Intrinsic noise in cultured hippocampal neurons: experiment and modeling. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(43):9723–33, Oct. 2004.

A. J. Doupe. Songand orderselective neurons develop in the songbird anterior forebrain during vocal learning. *Journal of neurobiology*, 33(5):694–709, Nov. 1997.

a. J. Doupe and P. K. Kuhl. Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience*, 22:567–631, Jan. 1999.

J. Dugas-Ford, J. J. Rowell, and C. W. Ragsdale. Cell-type homologies and the origins of the neocortex. *Proceedings of the National Academy of Sciences of the United States of America*, 109(42):16974–9, Oct. 2012.

S. E. Durand, J. M. Tepper, and M. F. Cheng. The shell region of the nucleus ovoidalis: a subdivision of the avian auditory thalamus. *The Journal of comparative neurology*, 323(4):495–518, Sept. 1992.

P. Dutar, H. M. Vu, and D. J. Perkel. Multiple cell types distinguished by physiological, pharmacological, and anatomic properties in nucleus HVc of the adult zebra finch. *Journal of neurophysiology*, 80(4):1828–38, Oct. 1998.

A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, N. K. Logothetis, and A. S. Tolias. Decorrelated neuronal firing in cortical microcircuits. *Science (New York, N.Y.)*, 327 (5965):584–7, Jan. 2010.

A. S. Ecker, P. Berens, A. S. Tolias, and M. Bethge. The Effect of Noise Correlations in Populations of Diversely Tuned Neurons. *Journal of Neuroscience*, 31(40):14272–14283, Oct. 2011.

J. J. Eggermont. Properties of correlated neural activity clusters in cat auditory cortex resemble those of neural assemblies. *Journal of neurophysiology*, 96(2):746–64, Aug. 2006.

J. J. Eggermont, W. J. Epping, and A. M. Aertsen. Stimulus dependent neural correlations in the auditory midbrain of the grassfrog (Rana temporaria L.). *Biological cybernetics*, 47(2):103–17, Jan. 1983.

G. B. Ermentrout, R. F. Galán, and N. N. Urban. Reliability, synchrony and noise. *Trends in neurosciences*, 31(8):428–34, Aug. 2008.

M. A. Escabi and C. E. Schreiner. Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(10):4114–31, May 2002.

A. A. Faisal, L. P. J. Selen, D. M. Wolpert, and A. Aldo Faisal. Noise in the nervous system. *Nature reviews. Neuroscience*, 9(4):292–303, Apr. 2008.

P. Fatt and B. Katz. Some observations on biological noise. *Nature*, 166(4223):597–8, Oct. 1950.

E. S. Fortune and D. Margoliash. Cytoarchitectonic organization and morphology of cells of the field L complex in male zebra finches (Taenopygia guttata). *The Journal of comparative neurology*, 325(3):388–404, Nov. 1992.

E. S. Fortune and D. Margoliash. Parallel pathways and convergence onto HVc and adjacent neostriatum of adult zebra finches (Taeniopygia guttata). *The Journal of comparative neurology*, 360(3):413–41, Sept. 1995.

E. F. Foster and S. W. Bottjer. Axonal connections of the high vocal center and surrounding cortical regions in juvenile and adult male zebra finches. *The Journal of comparative neurology*, 397(1):118–38, July 1998.

C. A. Fox and J. W. Barnard. A quantitative study of the Purkinje cell dendritic branchlets and their relationship to afferent fibres. *Journal of anatomy*, 91(3):299–313, July 1957.

R. W. Friedrich, C. J. Habermann, and G. Laurent. Multiplexing using synchrony in the zebrafish olfactory bulb. *Nature neuroscience*, 7(8):862–71, Aug. 2004.

P. Fries, J. H. Reynolds, A. E. Rorie, and R. Desimone. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science (New York, N.Y.)*, 291 (5508):1560–3, Mar. 2001.

L. Gabernet, S. P. Jadhav, D. E. Feldman, M. Carandini, and M. Scanziani. Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. *Neuron*, 48(2):315–27, Oct. 2005.

J. Gautrais and S. Thorpe. Rate coding versus temporal order coding: a theoretical approach. *Bio Systems*, 48(1-3):57–65, 1998.

T. Q. Gentner. Neural systems for individual song recognition in adult birds. *Annals of the New York Academy of Sciences*, 1016:282–302, June 2004.

T. Q. Gentner and D. Margoliash. Neuronal populations and single cells representing learned auditory objects. *Nature*, 424(6949):669–74, Aug. 2003.

N. Gordon, T. M. Shackleton, A. R. Palmer, and I. Nelken. Responses of neurons in the inferior colliculus to binaural disparities: insights from the use of Fisher information and mutual information. *Journal of neuroscience methods*, 169(2):391–404, Apr. 2008.

G. D. Graña, C. P. Billimoria, and K. Sen. Analyzing variability in neural responses to complex natural sounds in the awake songbird. *Journal of neurophysiology*, 101(6): 3147–57, June 2009.

J. A. Grace, N. Amin, N. C. Singh, and F. E. Theunissen. Selectivity for conspecific song in the zebra finch auditory forebrain. *Journal of neurophysiology*, 89(1):472–87, Jan. 2003.

D. S. Greenberg, A. R. Houweling, and J. N. D. Kerr. Population imaging of ongoing neuronal activity in the visual cortex of awake rats. *Nature neuroscience*, 11(7):749–51, July 2008.

Y. Gu, S. Liu, C. R. Fetsch, Y. Yang, S. Fok, A. Sunkara, G. C. DeAngelis, and D. E. Angelaki. Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron*, 71(4):750–61, Aug. 2011.

D. A. Gutnisky and V. Dragoi. Adaptive coding of visual information in neural populations. *Nature*, 452(7184):220–4, Mar. 2008.

R. H. R. Hahnloser and A. Kotowicz. Auditory representations and memory in birdsong learning. *Current opinion in neurobiology*, 20(3):332–339, Mar. 2010.

R. H. R. Hahnloser, A. A. Kozhevnikov, and M. S. Fee. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*, 419(6902):65–70, Sept. 2002.

N. G. Hatsopoulos, C. L. Ojakangas, L. Paninski, and J. P. Donoghue. Information about movement direction obtained from synchronous activity of motor cortical neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 95 (26):15706–11, Dec. 1998.

A. L. Hodgkin and A. F. Huxley. Action potentials recorded from inside a nerve fibre. *Nature*, 144(3651):710–711, 1939.

A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117 (4):500–44, Aug. 1952.

A. R. Houweling and M. Brecht. Behavioural report of single neuron stimulation in somatosensory cortex. *Nature*, 451(7174):65–8, Jan. 2008.

T. Hromádka, M. R. Deweese, and A. M. Zador. Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS biology*, 6(1):e16, Jan. 2008.

A. Hsu, S. M. N. Woolley, T. E. Fremouw, and F. E. Theunissen. Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(41):9201–11, Oct. 2004.

X. Huang and S. G. Lisberger. Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. *Journal of neurophysiology*, 101(6):3012–30, June 2009.

D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148:574–91, Oct. 1959.

D. Huber, L. Petreanu, N. Ghitani, S. Ranade, T. Hromádka, Z. F. Mainen, and K. Svoboda. Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice. *Nature*, 451(7174):61–4, Jan. 2008.

K. Immelmann. Song development in the zebra finch and other estrildid finches. In R. A. Hinde, editor, *Bird vocalizations*. Cambridge University Press, 1969.

M. Ito and H. Komatsu. Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(13):3313–24, Mar. 2004.

E. D. Jarvis. Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016:749–77, June 2004.

E. D. Jarvis. Neural systems for vocal learning in birds and humans: a synopsis. *Journal of ornithology / DO-G*, 148(1):35–44, Dec. 2007.

E. D. Jarvis and F. N. Nottebohm. Motor-driven gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8):4097–102, Apr. 1997.

J. M. Jeanne, J. V. Thompson, T. O. Sharpee, and T. Q. Gentner. Emergence of learned categorical representations within an auditory forebrain circuit. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(7):2595–606, Feb. 2011.

J. K. Jeong, T. A. Terleph, K. Burrows, L. A. Tremere, and R. Pinaud. Expression and rapid experience-dependent regulation of type-A GABAergic receptors in the songbird auditory forebrain. *Developmental neurobiology*, 71(10):803–17, Oct. 2011.

H. J. Karten. The ascending auditory pathway in the pigeon (Columba livia). II. Telencephalic projections of the nucleus ovoidalis thalami. *Brain research*, 11(1):134–53, Oct. 1968.

H. J. Karten. Homology and evolutionary origins of the 'neocortex'. *Brain, behavior and evolution*, 38(4-5):264–72, Jan. 1991.

L. C. Katz and M. E. Gurney. Auditory responses in the zebra finch's motor system for song. *Brain research*, 221(1):192–7, Sept. 1981.

G. B. Keller and R. H. R. Hahnloser. Neural processing of auditory feedback during vocal practice in a songbird. *Nature*, 457(7226):187–90, Jan. 2009.

D. B. Kelley and F. N. Nottebohm. Projections of a telencephalic auditory nucleus-field L-in the canary. *The Journal of comparative neurology*, 183(3):455–69, Feb. 1979.

G. Kim and A. J. Doupe. Organized representation of spectrotemporal features in songbird auditory forebrain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(47):16977–90, Dec. 2011.

R. Kimpo, F. E. Theunissen, and A. J. Doupe. Propagation of correlated activity through multiple stages of a neural circuit. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23(13):5750–61, July 2003.

D. Knudsen and T. Q. Gentner. Mechanisms of song perception in oscine birds. *Brain and language*, 115(1):59–68, May 2010.

E. I. Knudsen. Mechanisms of experience-dependent plasticity in the auditory localization pathway of the barn owl. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, 185(4):305–21, Oct. 1999.

A. Kohn and M. A. Smith. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(14):3661–73, Apr. 2005.

P. König, A. K. Engel, and W. Singer. Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends in neurosciences*, 19(4):130–7, Apr. 1996.

M. Konishi. The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift für Tierpsychologie*, 22(7):770–83, Dec. 1965.

M. Konishi. Comparative neurophysiological studies of hearing and vocalizations in songbirds. *Zeitschrift für Vergleichende Physiologie*, 66(3):257–272, 1970.

M. Konishi. Birdsong: from behavior to neuron. *Annual review of neuroscience*, 8: 125–70, Jan. 1985.

A. A. Kozhevnikov and M. S. Fee. Singing-related activity of identified HVC neurons in the zebra finch. *Journal of neurophysiology*, 97(6):4271–83, June 2007.

N. O. E. Krützfeldt, P. Logerot, M. F. Kubke, and J. M. Wild. Connections of the auditory brainstem in a songbird, Taeniopygia guttata. II. Projections of nucleus angularis and nucleus laminaris to the superior olive and lateral lemniscal nuclei. *The Journal of comparative neurology*, 518(11):2135–48, June 2010.

V. A. F. Lamme, H. Supèr, and H. Spekreijse. Feedforward, horizontal, and feedback processing in the visual cortex. *Current opinion in neurobiology*, 8(4):529–35, Aug. 1998.

I. Lampl, I. Reichova, and D. Ferster. Synchronous membrane potential fluctuations in neurons of the cat visual cortex. *Neuron*, 22(2):361–74, Feb. 1999.

D. Lee, N. L. Port, W. Kruse, and a. P. Georgopoulos. Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 18(3): 1161–70, Feb. 1998.

H. Lei and R. Mooney. Manipulation of a central auditory representation shapes learned vocal output. *Neuron*, 65(1):122–34, Jan. 2010.

H. J. Leppelsack. Funktionelle Eigenschaften der Hörbahn im Feld L des Neostriatum caudale des Staren (Sturnus vulgaris L., Aves). *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, 88(3):271–320, 1974.

H. J. Leppelsack and M. Vogt. Responses of auditory neurons in the forebrain of a songbird to stimulation with species-specific sounds. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, 107(3):263–274, 1976.

M. S. Lewicki. Intracellular characterization of song-specific neurons in the zebra finch auditory forebrain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(18):5855–63, Sept. 1996.

M. S. Lewicki. Efficient coding of natural sounds. *Nature neuroscience*, 5(4):356–63, Apr. 2002.

M. S. Lewicki and B. J. Arthur. Hierarchical organization of auditory temporal context sensitivity. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(21):6987–98, Nov. 1996.

M. S. Lewicki and M. Konishi. Mechanisms underlying the sensitivity of songbird forebrain neurons to temporal order. *Proceedings of the National Academy of Sciences of the United States of America*, 92(12):5582–6, June 1995.

J. F. Linden, R. C. Liu, M. Sahani, C. E. Schreiner, and M. M. Merzenich. Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *Journal of neurophysiology*, 90(4):2660–75, Oct. 2003.

P. Logerot, N. O. E. Krützfeldt, J. M. Wild, and M. F. Kubke. Subdivisions of the Auditory Midbrain (N. Mesencephalicus Lateralis, pars dorsalis) in Zebra Finches Using Calcium-Binding Protein Immunocytochemistry. *PLoS ONE*, 6(6):e20686, June 2011.

M. London, A. Roth, L. Beeren, M. Häusser, and P. E. Latham. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature*, 466 (7302):123–7, July 2010.

M. A. Long, D. Jin, and M. S. Fee. Support for a synaptic chain model of neuronal sequence generation. *Nature*, 468(7322):394–9, Nov. 2010.

T. Lu, L. Liang, and X. Wang. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature neuroscience*, 4(11):1131–8, Nov. 2001.

D. M. MacKay and W. S. McCulloch. The limiting information capacity of a neuronal link. *The Bulletin of Mathematical Biophysics*, 14(2):127–135, June 1952.

Z. F. Mainen and T. J. Sejnowski. Reliability of spike timing in neocortical neurons. *Science (New York, N.Y.)*, 268(5216):1503–6, June 1995.

D. Maney and R. Pinaud. Estradiol-dependent modulation of auditory processing and selectivity in songbirds. *Frontiers in neuroendocrinology*, 32(3):287–302, Dec. 2010.

D. Margoliash. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 3(5):1039–57, May 1983.

D. Margoliash. Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 6(6):1643–61, June 1986.

P. Marler. A comparative approach to vocal learning: Song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology*, 71(2, Pt.2):1–25, 1970a.

P. Marler. Birdsong and speech development: could there be parallels? *American scientist*, 58(6):669–73, 1970b.

P. Marler. Innateness and the instinct to learn. *Anais da Academia Brasileira de Ciências*, 76(2):189–200, June 2004.

P. Marler and S. Peters. Structural Changes in Song Ontogeny in the Swamp Sparrow Melospiza georgiana. *The Auk*, 99(3):446–458, 1982.

N. Masuda and K. Aihara. Dual coding hypotheses for neural information representation. *Mathematical biosciences*, 207(2):312–21, June 2007.

J. McCasland and M. Konishi. Interaction between auditory and motor activities in an avian song control nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 78(12):7815–9, Dec. 1981. p

C. D. Meliza, Z. Chi, and D. Margoliash. Representations of conspecific song by starling secondary forebrain auditory neurons: toward a hierarchical framework. *Journal of neurophysiology*, 103(3):1195–208, Mar. 2010.

C. V. Mello, D. S. Vicario, and D. F. Clayton. Song presentation induces gene expression in the songbird forebrain. *Proceedings of the National Academy of Sciences of the United States of America*, 89(15):6818–22, Aug. 1992.

J. W. Middleton, C. Omar, B. Doiron, and D. J. Simons. Neural correlation is stimulus modulated by feedforward inhibitory circuitry. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 32(2):506–18, Jan. 2012.

J. A. Mitchell and G. Hall. Paleostriatal lesions and instrumental learning in the pigeon. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology*, 36(2):93–117, May 1984.

J. F. Mitchell, K. A. Sundberg, and J. H. Reynolds. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron*, 63(6):879–88, Sept. 2009.

M. A. Montemurro, S. Panzeri, M. Maravall, A. Alenda, M. R. Bale, M. Brambilla, and R. S. Petersen. Role of precise spike timing in coding of dynamic vibrissa stimuli in somatosensory thalamus. *Journal of neurophysiology*, 98(4):1871–82, Oct. 2007.

R. Mooney. Different subthreshold mechanisms underlie song selectivity in identified HVc neurons of the zebra finch. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 20(14):5420–36, July 2000.

R. Mooney, M. J. Rosen, and C. B. Sturdy. A bird's eye view: top down intracellular analyses of auditory selectivity for learned vocalizations. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, 188(11-12):879–95, Dec. 2002.

G. P. Moore, J. P. Segundo, D. H. Perkel, and H. Levitan. Statistical signs of synaptic interaction in neurons. *Biophysical journal*, 10(9):876–900, Sept. 1970.

F. Mormann, S. Kornblith, R. Q. Quiroga, A. Kraskov, M. Cerf, I. Fried, and C. Koch. Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(36):8865–72, Sept. 2008.

D. Moshitch, L. Las, N. Ulanovsky, O. Bar-Yosef, and I. Nelken. Responses of neurons in primary auditory cortex (A1) to pure tones in the halothane-anesthetized cat. *Journal of neurophysiology*, 95(6):3756–69, June 2006.

V. B. Mountcastle. Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J Neurophysiol*, 20(4):408–434, 1957.

C. M. Müller and H. J. Leppelsack. Feature extraction and tonotopic organization in the avian auditory forebrain. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 59(3):587–99, Jan. 1985.

C. M. Müller and H. Scheich. GABAergic inhibition increases the neuronal selectivity to natural sounds in the avian auditory forebrain. *Brain research*, 414(2):376–80, June 1987.

C. M. Müller and H. Scheich. Contribution of GABAergic inhibition to the response characteristics of auditory units in the avian forebrain. *Journal of neurophysiology*, 59(6):1673–89, June 1988.

K. I. Nagel and A. J. Doupe. Temporal processing and adaptation in the songbird auditory forebrain. *Neuron*, 51(6):845–59, Sept. 2006.

K. I. Nagel and A. J. Doupe. Organizing principles of spectro-temporal encoding in the avian primary auditory area field L. *Neuron*, 58(6):938–55, June 2008.

K. I. Nagel, G. Kim, H. McLendon, and A. J. Doupe. A bird brain's view of auditory processing and perception. *Hearing research*, 273(1-2):123–33, Mar. 2011.

B. E. Nixdorf-Bergweiler and H. J. Bischof. *A Stereotaxic Atlas of the Brain of the Zebra Finch, Taeniopygia Guttata: With Special Emphasis on Telencephalic Visual and Song System Nuclei in Transverse and Sagittal Sections*. University of Kiel, 2007.

F. N. Nottebohm. Ontogeny of bird song. *Science (New York, N.Y.)*, 167(3920):950–6, Feb. 1970.

F. N. Nottebohm and A. P. Arnold. Sexual dimorphism in vocal control areas of the songbird brain. *Science (New York, N.Y.)*, 194(4261):211–3, Oct. 1976.

F. N. Nottebohm, T. M. Stokes, and C. M. Leonard. Central control of song in the canary, Serinus canarius. *The Journal of comparative neurology*, 165(4):457–86, Feb. 1976.

F. N. Nottebohm, D. B. Kelley, and J. A. Paton. Connections of vocal control nuclei in the canary telencephalon. *The Journal of comparative neurology*, 207(4):344–57, June 1982.

I. E. Ohiorhenuan, F. Mechler, K. P. Purpura, A. M. Schmid, Q. Hu, and J. D. Victor. Sparse coding and high-order correlations in fine-scale cortical networks. *Nature*, 466(7306):617–21, July 2010.

B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–9, June 1996.

K. Padmanabhan and N. N. Urban. Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. *Nature neuroscience*, 13(10):1276–82, Oct. 2010.

G. Palm, A. M. Aertsen, and G. L. Gerstein. On the significance of correlations among neuronal spike trains. *Biological cybernetics*, 59(1):1–11, Jan. 1988.

S. Panzeri, R. S. Petersen, S. R. Schultz, M. Lebedev, and M. E. Diamond. The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron*, 29(3):769–77, Mar. 2001.

J. Perez-Orive, O. Mazor, G. C. Turner, S. Cassenaer, R. I. Wilson, and G. Laurent. Oscillations and sparsening of odor representations in the mushroom body. *Science (New York, N.Y.)*, 297(5580):359–65, July 2002.

J. Perez-Orive, M. Bazhenov, and G. Laurent. Intrinsic and circuit properties favor coincidence detection for decoding oscillatory input. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(26):6037–47, June 2004.

D. H. Perkel, G. L. Gerstein, and G. P. Moore. Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophysical journal*, 7(4):419–40, July 1967a.

D. H. Perkel, G. L. Gerstein, and G. P. Moore. Neuronal spike trains and stochastic point processes. I. The single spike train. *Biophysical journal*, 7(4):391–418, July 1967b.

R. Pinaud and C. V. Mello. GABA immunoreactivity in auditory and song control brain areas of zebra finches. *Journal of chemical neuroanatomy*, 34(1-2):1–21, Sept. 2007.

R. Pinaud, T. A. F. Velho, J. K. Jeong, L. A. Tremere, R. M. Leão, H. von Gersdorff, and C. V. Mello. GABAergic neurons participate in the brain's response to birdsong auditory stimulation. *The European journal of neuroscience*, 20(5):1318–30, Sept. 2004.

C. Poo and J. S. Isaacson. Odor representations in olfactory cortex: "sparse" coding, global inhibition, and oscillations. *Neuron*, 62(6):850–61, June 2009.

A. Pouget, P. Dayan, and R. Zemel. Information processing with population codes. *Nature reviews. Neuroscience*, 1(2):125–32, Dec. 2000.

J. F. Prather, S. Peters, S. Nowicki, and R. Mooney. Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature*, 451(7176):305–10, Jan. 2008.

J. F. Prather, S. Nowicki, R. G. Andrzejak, S. Peters, and R. Mooney. Neural correlates of categorical perception in learned vocal communication. *Nature neuroscience*, 12 (2):221–8, Feb. 2009.

R. Q. Quiroga and S. Panzeri. Extracting information from neuronal populations: information theory and decoding approaches. *Nature reviews. Neuroscience*, 10(3): 173–85, Mar. 2009.

R. Q. Quiroga, L. Reddy, G. Kreiman, C. Koch, and I. Fried. Invariant visual representation by single neurons in the human brain. *Nature*, 435(7045):1102–7, June 2005.

A. Reiner, D. J. Perkel, L. L. Bruce, A. B. Butler, A. Csillag, W. Kuenzel, L. Medina, G. Paxinos, T. Shimizu, G. Striedter, M. Wild, G. F. Ball, S. Durand, O. Güntürkün, D. W. Lee, C. V. Mello, A. Powers, S. A. White, G. Hough, L. Kubikova, T. V.

Smulders, K. Wada, J. Dugas-Ford, S. Husband, K. Yamamoto, J. Yu, C. Siang, E. D. Jarvis, and O. Gütürkün. Revised nomenclature for avian telencephalon and some related brainstem nuclei. *The Journal of comparative neurology*, 473(3):377–414, May 2004.

A. D. Reyes. Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. *Nature neuroscience*, 6(6):593–9, June 2003.

A. Riehle, S. Grün, M. Diesmann, and A. Aertsen. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science (New York, N.Y.)*, 278(5345):1950–3, Dec. 1997.

F. Rieke, D. Warland, R. R. de Ruyter van Steveninck, and W. Bialek. *Spikes: Exploring the Neural Code*. MIT Press, Cambridge, Mass., 1997.

M. Riesenhuber. Getting a handle on how the brain generates complexity. *Network (Bristol, England)*, 23(3):123–7, Jan. 2012.

M. Riesenhuber and T. Poggio. Hierarchical models of object recognition in cortex. *Nature neuroscience*, 2(11):1019–25, Nov. 1999.

P. R. Roelfsema, V. A. F. Lamme, and H. Spekreijse. Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nature neuroscience*, 7 (9):982–91, Sept. 2004.

M. Rose. Über die cytoarchitektonische Gliederung des Vorderhirns der Vögel. *Journal für Psychologie und Neurologie*, 21(November):278–352, 1914.

G. Rothschild, I. Nelken, and A. Mizrahi. Functional organization and population dynamics in the mouse primary auditory cortex. *Nature neuroscience*, 13(3):353–60, Mar. 2010.

P. A. Salin and J. Bullier. Corticocortical connections in the visual system: structure and function. *Physiological reviews*, 75(1):107–54, Jan. 1995.

E. Salinas and T. J. Sejnowski. Correlated neuronal activity and the flow of neural information. *Nature reviews. Neuroscience*, 2(8):539–50, Aug. 2001.

H. Scheich, G. Langner, and D. Bonke. Responsiveness of units in the auditory neostriatum of the guinea fowl (Numida meleagris) to species-specific calls and synthetic stimuli. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, 132(3):243–255, 1979.

P. H. Schiller, B. L. Finlay, and S. F. Volman. Short-term response variability of monkey striate neurons. *Brain research*, 105(2):347–9, Mar. 1976.

J. W. Schumacher, D. M. Schneider, and S. M. N. Woolley. Anesthetic state modulates excitability but not spectral tuning or neural discrimination in single auditory midbrain neurons. *Journal of neurophysiology*, 106(2):500–14, Aug. 2011.

J. P. Segundo, G. P. Moore, L. J. Stensaas, and T. H. Bullock. Sensitivity of Neurones in Aplysia To Temporal Pattern of Arriving Impulses. *The Journal of experimental biology*, 40:643–67, Dec. 1963.

K. Sen, F. E. Theunissen, and A. J. Doupe. Feature analysis of natural sounds in the songbird auditory forebrain. *Journal of neurophysiology*, 86(3):1445–58, Sept. 2001.

H. S. Seung and H. Sompolinsky. Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences of the United States of America*, 90 (22):10749–53, Dec. 1993.

M. N. Shadlen and W. T. Newsome. Noise, neural codes and cortical organization. *Current opinion in neurobiology*, 4(4):569–79, Aug. 1994.

M. N. Shadlen and W. T. Newsome. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 18(10):3870–96, May 1998.

S. S. Shaevitz and F. E. Theunissen. Functional connectivity between auditory areas field L and CLM and song system nucleus HVC in anesthetized zebra finches. *Journal of neurophysiology*, 98(5):2747–64, Nov. 2007.

M. Shamir and H. Sompolinsky. Implications of neuronal diversity on population coding. *Neural computation*, 18(8):1951–86, Aug. 2006.

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

P. W. Sherman, H. K. Reeve, and D. W. Pfennig. Recognition Systems. In J. R. Krebs and N. B. Davies, editors, *Behavioral Ecology*, chapter 4, pages 69–96. Blackwell Publishing, Oxford, 4th editio edition, 1997.

M. A. Smith and A. Kohn. Spatial and temporal scales of neuronal correlation in primary visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(48):12591–603, Nov. 2008.

W. R. Softky. Simple codes versus efficient codes. *Current opinion in neurobiology*, 5 (2):239–47, Apr. 1995.

H. Sompolinsky, H. Yoon, K. Kang, and M. Shamir. Population coding in neuronal systems with correlated noise. *Physical Review E*, 64(5):1–11, Oct. 2001.

P. N. Steinmetz, A. Roy, P. J. Fitzgerald, S. S. Hsiao, K. O. Johnson, and E. Niebur. Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature*, 404(6774):187–90, Mar. 2000.

J. A. Swets. Detection theory and psychophysics: A review. *Psychometrika*, 26(1): 49–63, Mar. 1961.

T. Tchumatchenko, T. Geisel, M. Volgushev, and F. Wolf. Spike correlations - what can they tell about synchrony? *Frontiers in neuroscience*, 5(May):68, Jan. 2011.

T. A. Terleph, C. V. Mello, and D. S. Vicario. Auditory topography and temporal response dynamics of canary caudal telencephalon. *Journal of neurobiology*, 66(3): 281–92, Feb. 2006.

T. A. Terleph, C. V. Mello, and D. S. Vicario. Species differences in auditory processing dynamics in songbird auditory telencephalon. *Developmental neurobiology*, 67(11): 1498–510, Sept. 2007.

F. E. Theunissen. From synchrony to sparseness. *Trends in Neurosciences*, 26(2):61–64, Feb. 2003.

F. E. Theunissen and S. S. Shaevitz. Auditory processing of vocal sounds in birds. *Current opinion in neurobiology*, 16(4):400–7, Aug. 2006.

F. E. Theunissen, K. Sen, and A. J. Doupe. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 20(6):2315–31, Mar. 2000.

F. E. Theunissen, S. V. David, N. C. Singh, A. Hsu, W. E. Vinje, and J. L. Gallant. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network (Bristol, England)*, 12(3):289–316, Aug. 2001.

F. E. Theunissen, N. Amin, S. S. Shaevitz, S. M. N. Woolley, T. Fremouw, and M. E. Hauber. Song selectivity in the song system and in the auditory forebrain. *Annals of the New York Academy of Sciences*, 1016:222–45, June 2004a.

F. E. Theunissen, S. M. N. Woolley, A. Hsu, and T. Fremouw. Methods for the analysis of auditory processing in the brain. *Annals of the New York Academy of Sciences*, 1016:187–207, July 2004b.

J. A. Thompson and F. Johnson. HVC microlesions do not destabilize the vocal patterns of adult male zebra finches with prior ablation of LMAN. *Developmental neurobiology*, 67(2):205–18, Feb. 2007.

J. V. Thompson and T. Q. Gentner. Song recognition learning and stimulus-specific weakening of neural responses in the avian auditory forebrain. *Journal of neurophysiology*, 103(4):1785–97, Apr. 2010.

S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381(6582):520–2, June 1996.

P. H. E. Tiesinga, J.-M. Fellous, E. Salinas, J. V. José, and T. J. Sejnowski. Inhibitory synchrony as a mechanism for attentional gain modulation. *Journal of physiology, Paris*, 98(4-6):296–314, 2005.

D. J. Tolhurst, J. A. Movshon, and A. F. Dean. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision research*, 23(8):775–85, Jan. 1983.

M. Tomita and J. J. Eggermont. Cross-correlation and joint spectro-temporal receptive field properties in auditory cortex. *Journal of neurophysiology*, 93(1):378–92, Jan. 2005.

M. J. Tovée, E. T. Rolls, A. Treves, and R. P. Bellis. Information encoding and the responses of single neurons in the primate temporal visual cortex. *Journal of neurophysiology*, 70(2):640–54, Aug. 1993.

L. A. Tremere, J. K. Jeong, and R. Pinaud. Estradiol shapes auditory processing in the adult brain by regulating inhibitory transmission and plasticity-associated gene expression. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(18):5949–63, May 2009.

L. G. Ungerleider and J. V. Haxby. 'What' and 'where' in the human brain. *Current opinion in neurobiology*, 4(2):157–65, Apr. 1994.

W. M. Usrey. The role of spike timing for thalamocortical processing. *Current opinion in neurobiology*, 12(4):411–7, Aug. 2002.

G. E. Vates, B. M. Broome, C. V. Mello, and F. N. Nottebohm. Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. *The Journal of comparative neurology*, 366(4):613–42, Mar. 1996. doi: 10.1002/(SICI)1096-9861(19960318)366:4\<613::AID-CNE5\>3.0.CO;2-7.

W. E. Vinje and J. L. Gallant. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science (New York, N.Y.)*, 287(5456):1273–6, Feb. 2000.

R. von der Heydt, E. Peterhans, and G. Baumgartner. Illusory contours and cortical neuron responses. *Science (New York, N.Y.)*, 224(4654):1260–2, June 1984.

J. Wade and A. P. Arnold. Sexual differentiation of the zebra finch song system. *Annals of the New York Academy of Sciences*, 1016:540–59, June 2004.

H.-P. Wang, D. Spencer, J.-M. Fellous, and T. J. Sejnowski. Synchrony of thalamocortical inputs maximizes cortical reliability. *Science (New York, N.Y.)*, 328(5974):106–9, Apr. 2010a.

Y. Wang, A. Brzozowska-Prechtl, and H. J. Karten. Laminar and columnar auditory cortex in avian brain. *Proceedings of the National Academy of Sciences of the United States of America*, 107(28):12676–81, July 2010b.

S. Waydo, A. Kraskov, R. Q. Quiroga, I. Fried, and C. Koch. Sparse representation in the human medial temporal lobe. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(40):10232–4, Oct. 2006.

M. A. Whittington, R. D. Traub, and J. G. Jefferys. Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature*, 373(6515):612–5, Feb. 1995.

L. Wilbrecht and F. N. Nottebohm. Vocal learning in birds and humans. *Mental retardation and developmental disabilities research reviews*, 9(3):135–48, Jan. 2003.

J. M. Wild, N. O. E. Krützfeldt, and M. F. Kubke. Connections of the auditory brainstem in a songbird, Taeniopygia guttata. III. Projections of the superior olive and lateral lemniscal nuclei. *The Journal of comparative neurology*, 518(11):2149–67, June 2010.

R. W. Williams and K. Herrup. The control of neuron number. *Annual review of neuroscience*, 11:423–53, Jan. 1988.

B. Willmore and D. J. Tolhurst. Characterizing the sparseness of neural codes. *Network (Bristol, England)*, 12(3):255–70, Aug. 2001.

B. D. B. Willmore and A. J. King. Auditory cortex: representation through sparsification? *Current biology : CB*, 19(24):R1123–5, Dec. 2009.

F. Wolf, A. R. Houweling, and M. Brecht. Sparse and powerful cortical spikes. *Current opinion in neurobiology*, 20(3):306–312, Apr. 2010.

T. Womelsdorf, B. Lima, M. Vinck, R. Oostenveld, W. Singer, S. Neuenschwander, and P. Fries. Orientation selectivity and noise correlation in awake monkey area V1 are modulated by the γ cycle. *Proceedings of the National Academy of Sciences of the United States of America*, 109(11):4302–7, Mar. 2012.

S. M. N. Woolley and J. H. Casseday. Response properties of single neurons in the zebra finch auditory midbrain: response patterns, frequency coding, intensity coding, and spike latencies. *Journal of neurophysiology*, 91(1):136–51, Jan. 2004.

S. M. N. Woolley and J. H. Casseday. Processing of modulated sounds in the zebra finch auditory midbrain: responses to noise, frequency sweeps, and sinusoidal amplitude modulations. *Journal of neurophysiology*, 94(2):1143–57, Aug. 2005.

S. M. N. Woolley, P. R. Gill, and F. E. Theunissen. Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26 (9):2499–512, Mar. 2006.

S. M. N. Woolley, P. R. Gill, T. Fremouw, and F. E. Theunissen. Functional groups in the avian auditory system. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(9):2780–93, Mar. 2009.

J. Yu and D. Ferster. Membrane potential synchrony in primary visual cortex during sensory stimulation. *Neuron*, 68(6):1187–201, Dec. 2010.

R. Zann. Ontogeny of the Zebra Finch Distance Call: I. Effects of Cross-fostering to Bengalese Finches. *Zeitschrift für Tierpsychologie*, 68(1):1–23, Apr. 1985.

M. D. Zaretsky and M. Konishi. Tonotopic organization in the avian telencephalon. *Brain research*, 111(1):167–71, July 1976.

S. Zeng, X. Zhang, W. Peng, and M. Zuo. Immunohistochemistry and neural connectivity of the Ov shell in the songbird and their evolutionary implications. *The Journal of comparative neurology*, 470(2):192–209, Mar. 2004.

E. Zohary, M. N. Shadlen, and W. T. Newsome. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370(6485):140–3, July 1994.