Modulation and Manipulation of Sound Representation in the Auditory Cortex

By

Jessica Liberty Sackville Hamilton

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Shaowen Bao, Chair
Professor Frederic Theunissen
Professor Daniel Feldman
Professor Keith Johnson

Fall 2013

Modulation and Manipulation of Sound Representation in the Auditory Cortex

Abstract

Modulation and Manipulation of Sound Representation in the Auditory Cortex
by
Jessica Liberty Sackville Hamilton
Doctor of Philosophy in Neuroscience
University of California, Berkeley
Professor Shaowen Bao, Chair

The brain contains neurons of many different types interacting in complex functional circuits. To process sensory information these cells work in concert to form representations of the external world. In the auditory cortex, this involves integrating information from different cell types across an orderly anatomical structure of layers and columns. Representations can be observed at the level of single cells, cortical microcircuits, and large-scale sensory maps. The relationship between single cell properties and circuits within the auditory cortex, however, is still poorly understood. Furthermore, the structure-function relationships uncovered by neuroscientific study may crucially depend on the stimuli used to probe the system. This thesis brings together work from each of these different levels to describe how sounds are represented in the cortex, how this representation changes with experience, and how different cells contribute to cortical representation.

First, I describe how the statistics of sound stimuli influence response properties in the mouse primary auditory cortex by comparing responses to pure tones and natural sounds (ultrasonic vocalizations). I also compare these responses to a temporally reversed vocalization to determine whether a sound with similar spectrotemporal content but no ethological relevance is represented similarly. When comparing pure tones and vocalizations, I find that the temporal response properties are similar, but that spectral response properties (e.g. frequency selectivity) often differ substantially. In particular, there are multiple sites that responded to vocalizations with frequency content outside their classical tone-derived receptive field, suggesting some specificity for behaviorally relevant sounds. When comparing forward and backward vocalizations, temporal responses are similar, but frequency bandwidth and characteristic frequency differs significantly across the population. Thus, the behaviorally relevant sound appears to be represented differently from non-behaviorally relevant synthetic and naturalistic sounds.

The response properties of auditory neurons are not fixed, but rather depend on experience. In the next study, I examine how exposure to pulsed noise during different sensitive windows of the auditory critical period affects single site properties as well as circuit-level dynamics. On the single site level, I find that early exposure to pulsed noise increases receptive field thresholds and decreases frequency selectivity, while late noise exposure increases frequency bandwidths as well as spontaneous and evoked firing rates. To describe changes in functional microcircuits, I use the Ising model, which describes pairwise interactions between simultaneously recorded sites in the auditory cortex as well as interactions between sites and the stimuli that modulate them. I find that early noise exposure decreases stimulus drive, whereas late noise exposure does not change the strength of sound inputs but rather decreases the spread of functional connections from the deep to the superficial layers across sites with different frequency selectivity.

Finally, I use a combination of optogenetic tools and computational methods to describe how the activity of a specific class of inhibitory neurons affects network connectivity in the auditory cortex. I examine the contribution of parvalbumin-positive (PV+) inhibitory interneurons, which make up around half of the inhibitory neurons in the cortex. These neurons are known to be involved in the generation of gamma oscillations, and their maturation corresponds with the end of the auditory critical period for plasticity. Using Ising models in tandem with linear-nonlinear vector autoregressive models, I show that stimulating PV+ neurons increases feedforward information flow through cortical circuits without changing lateral interactions within the same layers.

This dissertation is dedicated to my family and friends.

Contents

List of Figures

Acknowledgments

The work involved in producing this thesis would not have been possible without a lot of generous and talented people. I am so happy to have had the chance to work at a place like UC Berkeley where people are so giving with their time, advice, and expertise.

First of all, thanks to my adviser, Shaowen Bao, for his imparting his knowledge, creativity, and unwavering optimism. Thank you for pushing me when I needed it, but also providing me with autonomy. Thanks to my lab mates, both past (Heesoo Kim, Hania Köver, Michele Insanally, Sungchil Yang, Benjamin Weiner, Stella Zhang) and present (Asako Miyakawa and Robert Gibboni). Thanks especially to Heesoo for teaching me everything I needed to know about electrophysiology and spending many patient hours with me in the basement. Thanks to Asako and Robert for being a great support throughout grad school, and for being fun and entertaining while providing great scientific feedback.

Thank you to my thesis committee, Frédéric Theunissen, Dan Feldman, and Keith Johnson. Special thanks to Frédéric for taking me on as an adopted lab member so that I could work with and get feedback from two talented research groups. Also thanks to Dan for helpful discussions about circuits and plasticity, and to Keith for help with current and future endeavors in auditory neuroscience.

Thanks to my undergraduate research mentees, Jeffrey Gu and Trevor Flynn, for their hard work and good attitudes.

Thanks to Jascha Sohl-Dickstein, for agreeing to collaborate with me and eventually fostering the development of two (!) great projects. Thanks for being so easy to work with and fun to be around!

Thanks to Vanessa Carels, for being an amazing co-author and friend. Thanks for being so cheerful about sitting with me for hours with a paint brush and tiny brain slices, and for being an amazing teacher for all things histology-related.

To members of the Dan lab: Seung-Hee Lee, Siyu Zhang, and Lucas Pinto for helping me get our optogenetics setup going in the first place. Thanks especially to Seung-Hee for letting me shadow her experiments and providing me with much-needed advice on protocols. Thanks also to Feldman lab members Ray Shao, Justin Elstrott, Brian Isett, Kelly Clancy, Toshio Miyashita, and Sam Harding-Forrester for generously allowing me to use their equipment!

Thanks to Kati Markowicz, Tony Leonard, and everyone at the Helen Wills Neuroscience Institute for being amazingly helpful.

To Wendy de Heer and Natalia Bilenko, thanks for your incredible friendship and helping maintain my mental health. Our weekly meetings were instrumental in navigating my future and I hope we have many more.

I would also like to thank my research mentor Katherine Narr at the University of California, Los Angeles, who guided me through an amazing pre-graduate school journey that persuaded me to go into science. Thank you for imparting your knowledge on good science, good writing, and how to navigate the academic world.

Thank you to my parents, Andrew and Catherine, and to my sister, Wildrose, for believing in me and always being there for me. Thank you to Barbro and Gerry for giving me a home away from home and for reminding me that "unconventional science" is sometimes the best science.

And last but not least, to Alex Huth, who has been an amazing partner in science and in life. We came to graduate school together, and two PhDs later I can't imagine having gone through it without you.

CHAPTER 1

Introduction: Sensory representations, circuits, and plasticity

Our ability to perceive a beautiful symphony, converse with a friend, or be awakened by our alarm clock in the morning relies on the coordination of vast numbers of neurons in the auditory centers of the brain. After sound waves enter the ears and vibrate the basilar membrane, an electrical signal with information about each sound's properties is transduced from the level of the auditory nerve all the way up to the highest processing center in the auditory pathway, the auditory cortex (Kandel et al. 2000). Whether we find certain chords in the symphony dissonant, or whether we understand the words that our friend is saying to us, depend on early experience. How exactly this perception occurs—which cells are involved, how they work together, how they are organized, and how they can change—is a fundamental problem in systems neuroscience.

1.1 Auditory processing: hierarchical levels of analysis

Studying sound representation in the brain involves analysis at multiple levels—from single cells to microcircuits to sensory maps across an auditory area (Fig. 1.1). At the single cell level or at the level of small populations of neurons in multi-unit recordings, researchers have found selectivity for many different acoustic features. These are generally measured using simple, synthesized sound stimuli, such as pure tone pips, dynamic chord tones, noise bursts, frequency-modulated (FM) sweeps, temporally orthogonal ripple combinations (TORCs), or modulation-limited noise (Ahrens et al. 2008; Blake & Merzenich 2002; Depireux et al. 2001; Escabi & Schreiner 2002; Fritz et al. 2003; Han et al. 2007; Kilgard & Merzenich 1999; Kim & Bao 2009; Linden & Schreiner 2003; Machens et al. 2003). However, the use of naturalistic or ethologically relevant stimuli with spectrotemporally complex statistical structure is becoming more common (Liu et al. 2007; Woolley et al. 2006; Theunissen et al. 2001; Syka et al. 2005; Šuta et al. 2007; Portfors et al. 2009; Wang et al. 1995; Gehr et al. 2000). By investigating properties at single sites, researchers have shown that cells in the auditory cortex show selectivity for specific sound frequencies (Polley et al. 2007), FM sweep direction (Insanally et al. 2009), sound intensities (Polley et al. 2006; Wu et al. 2006; Watkins & Bar-

bour 2008, 2011), and spectral contrast (Barbour & Wang 2003). Single sites also exhibit different temporal response properties, including the latency between stimulus onset and spiking activity. For example, cells in the primary auditory cortex (AI) tend to show faster response latencies than cells in secondary areas such as secondary auditory cortex (AII). Cells in different layers of the cortex also exhibit different temporal and spectral response characteristics (Sugimoto et al. 1997; Wallace & Palmer 2008; Sakata & Harris 2009; Christianson et al. 2011), with the shortest response latencies observed in the infragranular layers IIIb and IV, corresponding to regions of direct thalamic input (Linden & Schreiner 2003).

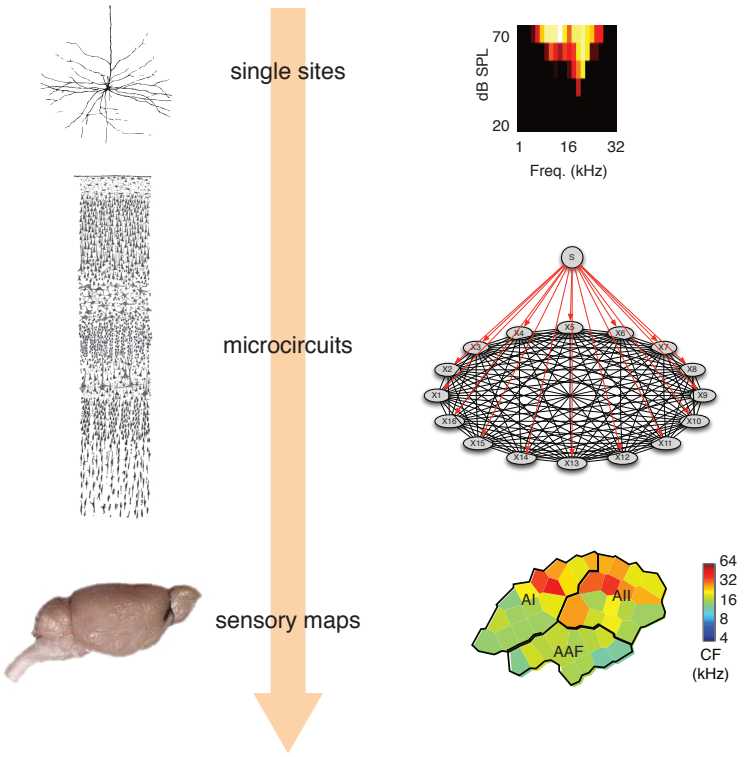

Figure 1.1: Different levels of analysis of auditory processing in the brain. Low level analysis may be performed on single cells or small populations of multi-units, where properties can be described by classical receptive fields (top right) or spectrotemporal receptive fields (see Chapter 2). At the microcircuit level, it is necessary to account for interactions between cells, which can be performed using graphical models as pictured at center right and described in detail in Chapters 3 and 4. At a more global level, multiple single site recordings can be used to construct a sensory map of selectivity for different sound features, as shown at the bottom right (and Chapter 3). Drawing of cortical cells by Santiago de Ramon y Cajal and mouse brain image are available on Wikimedia Commons.[1]

Though research on representation of external stimuli in single sites can tell us which stimulus properties might be important for perception and discrimination, neurons do not act in isolation. In 1957, Mountcastle (Mountcastle 1957) described the existence of the cortical column, thought to be the fundamental functional unit of the cortex. In his work, he showed that neurons in the primate somatosensory cortex were arranged vertically into columns, and that neurons in different columns had non-overlapping receptive fields. Later, the idea of the "canonical cortical circuit" was proposed by Douglas and Martin (Douglas et al. 1989; Douglas & Martin 1991) from experiments in the visual cortex. They showed that input from the thalamus arrives at the infragranular input layer of the cortex, then propagates to more superficial layers which then send feedback to deeper layers. Still, the idea that this microcircuit is truly "canonical" in the sense that it is uniform throughout

[1] `http://en.wikipedia.org/wiki/File:Cajal_cortex_drawings.png`, `https://en.wikipedia.org/wiki/File:Mouse_brain.jpg`

cortex has proven contentious (Douglas & Martin 2007). In this thesis, I hope to shed some light on how these circuits might be represented in the auditory cortex and how different types of cells contribute to information flow through the layers and columns of the cortex.

At the highest level described here, the level of sensory maps, researchers have divided the auditory cortex into a number of subfields based on their response properties and their spatial organization. The mouse auditory cortex consists of primary auditory cortex (AI) which receives direct input from the ventral medial geniculate body (MGBv) of the thalamus, anterior auditory field (AAF), which also receives direct input, secondary auditory cortex (AII), the dorso-posterior field (DP), and the ultrasonic field (UF) (Stiebler et al. 1997). Many studies have focused on representation and plasticity in AI (Han et al. 2007; Insanally et al. 2009, 2010; Zhang et al. 2002; Schreiner & Winer 2007; Kilgard et al. 2001; Kim & Bao 2009; Kim et al. 2013; Köver et al. 2013), while considerably fewer report changes to other auditory areas (Polley et al. 2006; Diamond & Weinberger 1986; Noreña & Eggermont 2002; Eggermont & Kenmochi 1998). This thesis describes sound representation and circuit manipulation in AI in Chapters 2, 3, and 4, and in AAF and AII in Chapter 3.

1.2 Cortical representation of natural and non-natural sounds

Like other sensory cortices, the auditory cortex appears to exhibit some fundamental organizing principles regarding sound representation (Kaas 1997). Perhaps the most salient of these principles is the topographic organization of sound by frequency selectivity, also called tonotopy (Clopton et al. 1974). Tonotopy exists at the level of the cochlea, where the resonant properties of the basilar membrane give it ordered selectivity to different sound frequencies (Müller 1991), and persists up to the level of the inferior colliculus (Stiebler & Ehret 1985), medial geniculate nucleus of the thalamus (Hackett et al. 2011), and the auditory cortex (Kaas 1997; Stiebler et al. 1997; Han et al. 2007). It is thought that this orderly representation of external features may underlie efficient coding and permit effective sensory discrimination (Kaas 1997; Schreiner & Winer 2007; Kohonen 1990). However, tonotopic representations are not sufficient to predict the responses of auditory neurons to more complex sounds. In mice, for example, populations of neurons in the inferior colliculus (Portfors et al. 2009; Holmstrom et al. 2010) respond to conspecific vocalizations without responding to pure tones with the same frequency content, suggesting some specificity for ethologically relevant sounds. Thus, the spectrotemporal content of the signal alone does not predict the brain's response to a sound.

In Chapter 2, I contrast the representation of mouse conspecific vocalizations with non-natural sounds—the pure tones used in many studies of auditory perception and plasticity. We find multiple instances of sites in the auditory cortex that respond to conspecific vocalizations but that will not respond to tone pips with similar frequency content, consistent with findings by Holmstrom et al. (2010). This suggests that tonotopy and responses to pure tones alone do not predict responses to more complex sounds, and that behavioral relevance may also play a role in complex sound representations. Still, we find that there are some similarities in response properties to natural and non-natural sounds. Temporal latencies, mean firing rate, and temporal bandwidths are generally

correlated across both classes of stimuli, while spectral properties such as frequency selectivity are not.

1.3 Plasticity in single sites and in cortical circuits

Previous studies have shown that properties of single sites, microcircuits, and maps are dynamic and can change with experience. These changes can be particularly drastic if exposure to sounds occurs during a "critical period" for plasticity, which in rodents is during the first month of life (Zhang et al. 2001; de Villers-Sidani et al. 2007), and which in human children may last until age 7 (Johnson & Newport 1989). During this critical period, brain circuits are particularly plastic and change their representations in response to the statistics of the external environment. Exposing rat pups to repeated 7 kHz pure tones, for example, results in an increase in the area of the tonotopic map sensitive to 7 kHz and narrows tuning bandwidths in the exposure frequency range (Han et al. 2007). Exposure to different types of sound stimuli during the critical period has varying effects on cortical representation—exposure to pulsed white noise reduces frequency selectivity, decreases receptive field quality, increases receptive field thresholds, and increases receptive field bandwidths (Zhang et al. 2002; Zhou & Merzenich 2008; Insanally et al. 2010). Exposure to continuous white noise, on the other hand, can delay the closure of the critical period, leaving circuits open to passive experience-dependent modification in adulthood (Chang & Merzenich 2003).

Though the critical period in rodents lasts for several weeks, it can be further subdivided into different sensitive windows, where exposure in each window affects different response properties of auditory neurons. For example, changes to the characteristic frequency of neurons appear to require exposure to sounds during early epochs (Insanally et al. 2009, 2010), whereas changes to response bandwidth and threshold occur with later exposure (Insanally et al. 2009, 2010). It is still unclear, however, how these changes in response properties are related to circuit changes. Zhang et al. (2002) showed that pulsed noise exposure decreases temporal correlations between pairs of cells, but more complex interactions between cells across layers were not described.

In Chapter 3, I describe how single sites in different functional fields of the mouse auditory cortex change with early experience. I then compare these findings to changes at the circuit level, incorporating interactions between pairs of simultaneously recorded sites across different layers and columns of the auditory cortex. To analyze changes at the circuit level I use the Ising model (Ising 1925), a graphical model that describes pairwise interactions (couplings) between nodes on a graph. The original Ising model was invented by the German physicist Wilhelm Lenz as a problem for his student, Ernst Ising (Lenz 1920), and was formulated to describe the configuration of atomic spins in a lattice, which allowed for the identification of phase transitions between ferromagnetic and paramagnetic states of matter (Ising 1925). However, this model is not limited to studies of atomic spins, and has been used to describe interactions in neural systems (Ganmor et al. 2011b,a; Köster et al. 2013; Marre et al. 2009; Ohiorhenuan et al. 2010; Roudi et al. 2009b; Schaub & Schultz 2012; Schneidman et al. 2006; Shlens et al. 2006, 2009; Tang et al. 2008) as well as protein folding (Bryngelson & Wolynes 1987; Bakk & Høye 2003) and even social networks (Klemm et al. 2003). In this thesis, the nodes in the Ising model are multi-unit sites recorded from a silicon polytrode in the

auditory cortex. A positive coupling between two sites means that the two sites tend to spike (or not spike) at the same time. A negative coupling means that when one site is active, the other site is not, and vice versa. However, this coupling is not the same as a direct correlation, since direct connections can be differentiated from false-positive connections that arise as a result of common inputs (Schneidman et al. 2006; Hamilton et al. 2013). For example, given three sites A, B, and C with some known connectivity (A is connected to B, and A is connected to C, but B and C are not connected), the Ising model will correctly uncover the true connectivity, whereas pairwise correlation would suggest a direct connection between B and C (Fig. 1.2).

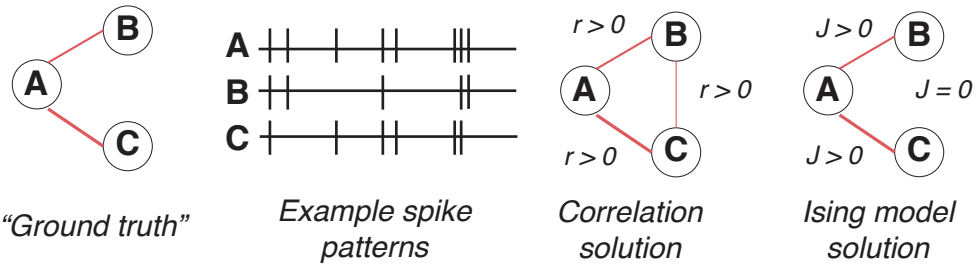

Figure 1.2: Differences between Ising model couplings and pairwise correlation. For a toy model with three sites with known connectivity (left), B and C appear to be correlated (and therefore connected) because they tend to spike at the same time. However, this connection can be explained away by their mutual connection to A.

By modifying the traditional Ising model to incorporate couplings to external stimuli as well as couplings between recording sites, I show how the model can be used to dissociate changes in different auditory cortex. Couplings between sound stimuli and deep layer inputs, for example, can be used as a proxy for thalamocortical connections. Couplings from deep to superficial layers can be examined to look at corticocortical changes. What we find is that exposure to pulsed noise during different sensitive windows of the auditory critical period differentially affects different auditory circuits. When we relate this to single site properties, we find that reduced frequency selectivity and increased thresholds may result from reduced input coupling, and that increased bandwidths are not related to functional corticocortical spread, but are more likely driven by increases in intrinsic firing rate.

1.4 Circuit manipulation

To fully understand the nature of circuit computations in the brain during auditory processing, it is necessary to both record from and manipulate cells of different types. Genetically-encoded optogenetic tools have recently made it possible to control specific subsets of neurons with millisecond precision using light (Boyden et al. 2005; Zhang et al. 2010; Fenno et al. 2011). With this technique, genes encoding light-sensitive opsins derived from microbial organisms are inserted into cells of interest. This can be performed using adeno-associated viruses (AAVs) and lentiviruses with or without specific genetic targeting via Cre-lox (Sauer 1998; Zhang et al. 2010), or for layer-specific

expression the opsins can be inserted via *in utero* electroporation (Zhang et al. 2010).

In this work, I use a Cre-inducible virus to express the intrinsically blue-light-sensitive ion channel channelrhodopsin-2 (Nagel et al. 2003) in parvalbumin-positive (PV+) inhibitory interneurons in the mouse auditory cortex. This channel is naturally found in the single-cell green alga *Chlamydomonas reinhardtii* (Fig. 1.3A), which uses it for cell motility (Nagel et al. 2003). When expressed in cortical neurons, shining blue light directly on the cortex will induce a conformational change in the ChR2 channel, causing it to open and depolarize the PV+ cells in which it is expressed (see Fig. 1.3B).

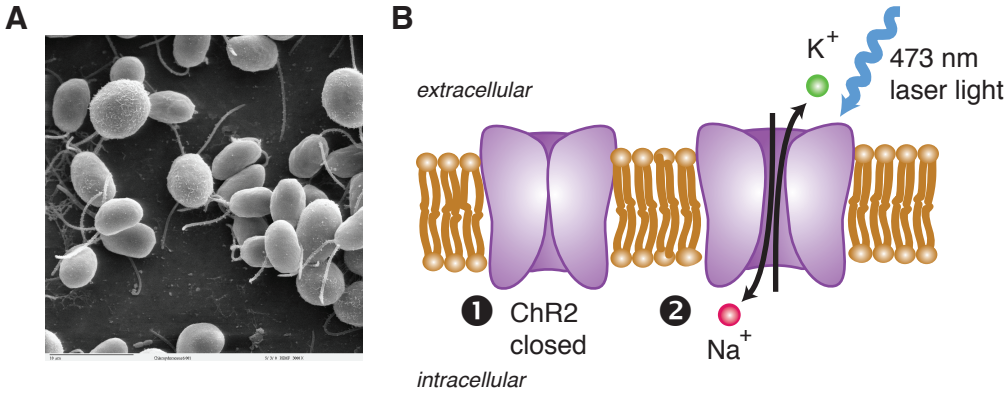

Figure 1.3: (A) Scanning electron microscope image of *Chlamydomonas reinhardtii*, the single celled green alga from which ChR2 is derived. Image courtesy of Dartmouth Electron Microscope Facility, Dartmouth College and available through Wikimedia Commons.[2] **(B)** Schematic of the light-sensitive ion channel channelrhodopsin-2 (ChR2). ChR2 is a non-specific cation channel that is activated by blue light ($\approx$ 473 nm wavelength for our laser). ChR2 contains the light-isomerizable complex all-*trans*-retinal which is converted to 13-*cis* retinal when it absorbs a photon, inducing a conformational change in the channel, opening of the pore, and an influx of positive ions into the cell.

The virus (pAAV–EF1a–DIO–hChR2(H134R)–EYFP–WPRE–pA) used here contains a non-specific but strong promoter (elongation factor 1a [EF1a], found in microtubules) and a DIO (double-floxed inverted) copy of ChR2. The addition of an extra pair of mutated lox sites allows for strong, irreversible expression of the channel once lox sites are cleaved in the presence of Cre (see Fig. 1.4).

Using optogenetics in vertebrate animals is relatively new and is still in development by multiple labs (Fenno et al. 2011; Bernstein & Boyden 2011; Boyden 2011). Recent work has shown that optogenetically stimulating PV+ neurons in visual and somatosensory cortices enhances feature selectivity, and that it may improve the signal-to-noise ratio (SNR) in cortical circuits (Atallah et al. 2012; Lee et al. 2012; Sohal et al. 2009; Wilson et al. 2012). In the auditory system, researchers have used optogenetics to probe the function of PV+ neurons by labeling and recording from them (Moore & Wehr 2013). Moore & Wehr (2013) found that PV+ neurons are fairly well-tuned for frequency and have short response latencies, suggesting that they play a minor role in frequency tuning, but may rather be involved in dynamic gain control of responses. Li et al. (2013) corroborate this, showing that strong optogenetic stimulation of PV+ neurons to silence putative corticocortical connections

[2] `http://commons.wikimedia.org/wiki/File:Chlamydomonas6-1.jpg`

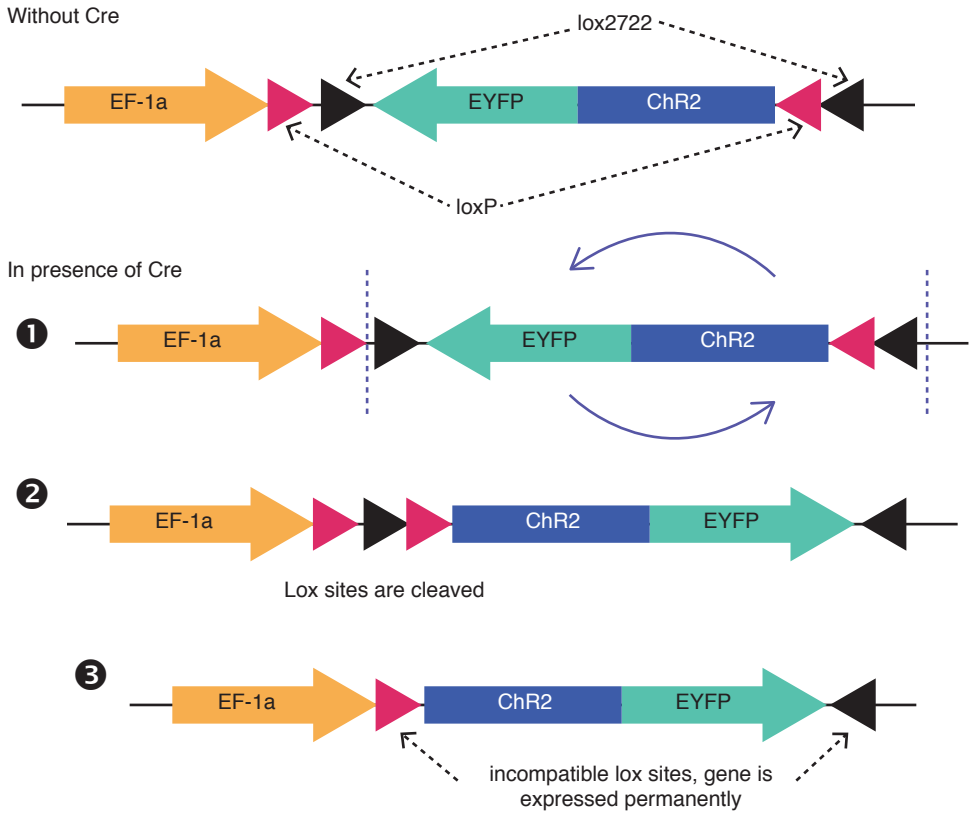

Figure 1.4: The virus used in this work encodes a reversed copy of ChR2 fused to the fluorescent marker EYFP, flanked by loxP and lox2722 sites on each side. The promoter EF1a allows the virus to express in all neurons, regardless of type. However, when the Cre protein is not present, ChR2-EYFP remains in the nonfunctional reversed configuration and is not expressed. In the presence of Cre, such as in a PV-Cre mouse, where Cre is expressed in PV+ inhibitory interneurons, the Cre protein will recognize the lox sites and catalyze DNA recombination between pairs of compatible lox sites (Sauer 1998). Depending on the direction of the lox non-palindromic core region, the DNA will either be inverted or excised. In this virus, Cre can recognize either pair of lox sites, and the DNA in between them will invert, since the pairs of lox sites face one another. After this occurs, one side of the ChR2-EYFP will have three lox sites in the same orientation. Thus, two of these lox sites will be excised, leaving incompatible lox sites flanking the ChR2-EYFP gene, which is now expressed in its functional direction and can no longer undergo recombination.

in the auditory cortex does not change auditory tuning, but instead linearly scales thalamic input. These studies suggest that PV+ neurons have an important role in sensory processing. Still, these studies were performed at the level of single sites or cells, and so do not describe changes to the network as a whole. In Chapter 4, I use optogenetics and computational methods (including the Ising model described above) to show how stimulating PV+ neurons affects functional connections in auditory cortical circuits. I also describe how PV+ neuron stimulation affects responses at the single site level. These results have important implications for how PV+ neurons contribute to sound perception and sensory gating.

1.5 Summary and motivation

Studies on function and representation in auditory cortex have generally focused on processing at one level of the hierarchy—either single cell/multi-unit response properties, microcircuit interactions through pairwise correlations, or map representation. This thesis brings together multiple approaches to provide a unified perspective on neural encoding in the auditory cortex.

CHAPTER 2

Representation of natural and synthetic sounds in mouse auditory cortex

2.1 Abstract

Rodents are highly vocal animals, with complex conspecific calls that facilitate social interactions. Most studies of rodent auditory processing have used synthetic sounds, which often have vastly different statistical structure compared to conspecific vocalizations. Neuronal selectivity derived with synthetic sounds often does not predict responses to conspecific vocalizations. Thus, to understand neural representations of conspecific calls in rodents, we compared auditory cortical responses in the adult female mouse to pure tones, male mouse courtship vocalizations, and the same mouse vocalizations reversed in time. Spectrotemporal receptive fields (STRFs) derived with vocalizations varied widely, ranging from spectrally and temporally narrow to broad. Despite similar secondary stimulus statistics for forward and backward vocalizations as calculated by the modulation power spectrum, STRFs derived with the two ensembles were sometimes quite different for a given recording site. Specifically, the characteristic frequency and frequency bandwidth for STRFs from forward and backward vocalizations tended to be different, while response latencies, STRF temporal bandwidth, and firing rate in response to the two stimuli were similar. Interestingly, STRFs sometimes showed excitatory bands in frequencies well outside the classical receptive field reconstructed from responses to pure tones. These results suggest that summing neural responses to simple frequency components may be inadequate to describe responses to natural sound, and that distinct mechanisms may be involved.

2.2 Introduction

Processing communication sounds is a main function of the auditory system. Some auditory neurons preferentially respond to conspecific vocalizations over time-reversed or heterospecific vocal-

izations (Hauber et al. 2007; Šuta et al. 2007; Wang & Kadia 2001), and may even respond robustly to sounds outside of the frequency range defined by their classical receptive field (Holmstrom et al. 2010; Portfors et al. 2009; Rauschecker et al. 1995). Such selectivity cannot be fully captured by simple frequency-intensity receptive fields of the neurons, as it involves complex spectrotemporal interactions and nonlinear integration of stimulus features into a neural response. These processes may require the complex statistical structure of natural stimuli that is absent in synthetic stimuli typically used in electrophysiological experiments (Attias & Schreiner 1997; Singh & Theunissen 2003). To elucidate neuronal selectivity for such complex acoustic features, researchers have attempted to model the relationship between the stimulus and response using linear and nonlinear spectrotemporal receptive field models (STRFs), which can describe neuronal selectivity for complex features (Aertsen & Johannesma 1981; deCharms et al. 1998; Eggermont et al. 1983; Theunissen et al. 2001; Theunissen & Shaevitz 2006; Wu et al. 2006). Evidence indicates that STRFs are highly dependent on stimulus context — they can be quite different when derived from tone pips, ripples or natural vocalizations (Blake & Merzenich 2002; Gill et al. 2008; Hauber et al. 2007; Linden & Schreiner 2003; Machens et al. 2004; Woolley et al. 2006). Among these different stimulus ensembles, natural vocalizations are advantageous because of their ethological relevance and their ability to drive auditory neurons. Examining STRFs derived from conspecific vocalizations may improve our understanding of neuronal processing of these sounds.

Rodents are good model animals to study neuronal processing of species-specific vocalizations. They produce a wide variety of ultrasonic vocalizations, including pup isolation calls and adult courtship or encounter calls (Hahn & Lavooy 2005; Holy & Guo 2005; Kim & Bao 2009; Liu et al. 2003). The calls are important for nursing and courtship behaviors, and may be categorically perceived like human speech sounds (Ehret 1992; Ehret & Haack 1981; Liu et al. 2003). The syllables are highly structured with characteristic frequency modulations and harmonic stacks produced in highly stereotyped sequences with precise temporal structure (Holy & Guo 2005). Preferential representations of these spectral and temporal features may contribute to selective responses to conspecific communication sounds. In addition, rodents have a short developmental time window and robust cortical plasticity, allowing systematic investigation of the role of experience in representations of species-specific vocalizations (Bao et al. 2004; Henry & McGinn 1992; Kim & Bao 2009). Despite these advantages of the model system, surprisingly little is known about how rodent auditory neurons respond to complex features in conspecific vocalizations. Most previous studies have either examined STRFs in rodents with synthetic sounds (Ahrens et al. 2008; Linden & Schreiner 2003), or looked at other measures of neural selectivity for conspecific calls (Holmstrom et al. 2010; Kao et al. 1997; Liu et al. 2003; Portfors et al. 2009). Some electrophysiological studies have also calculated STRFs using neural responses to jungle sounds (Asari & Zador 2009). Still, it is largely unknown what spectrotemporal features in rodent conspecific vocalizations optimally activate cortical neurons, and how stimulus context or ethological relevance relates to encoding of these sounds.

In this study, we characterized neural responses to both synthetic and natural sounds in the mouse auditory cortex and compared classical receptive fields obtained with pure tones to STRFs acquired with forward and reversed courtship vocalizations. We found that classical receptive field structure is similar to STRF structure in the temporal, but not spectral, domain, and that STRFs are stimulus dependent even when calculated with sounds containing almost identical secondary

stimulus statistics.

2.3 Experimental procedures

2.3.1 Subjects and experimental design

The University of California-Berkeley Animal Care and Use Committee approved all procedures. Six adult (≈P50) female B6;129P2-Pvalb$^{tm1(cre)Arbr/J}$ (The Jackson Laboratory) were used in this study. The animals were housed with litter mates of the same gender and provided with food and water *ad libitum* before surgical procedures were performed. We recorded neural responses to pure tones, forward vocalizations, and reversed vocalizations in the auditory cortex and analyzed their properties.

2.3.2 Electrophysiology and stimuli

The right auditory cortex was mapped for each mouse under anesthesia using a cocktail of ketamine (100 mg/kg) and xylazine (10 mg/kg) and procedures described previously (Han et al. 2007). Following deflection of the temporal muscle, exposure of the auditory cortex, and removal of the dura mater, multi-unit activity was recorded using tungsten microelectrodes (1–2 MΩ at 1kHz, FHC, Inc.) lowered orthogonally to the cortex at depth of 300-400 μm below the pial surface (approximately layer 4). The extracellular signal was obtained using a TDT amplifier connected to TDT RX5 hardware (Tucker Davis Technologies, Inc.). Spike times were calculated by thresholding the extracellular signal at 2 times the standard deviation of the signal-to-noise ratio, and were logged using custom software running on a Windows XP computer. Recordings were made in response to three types of stimuli: (1) pure tones, (2) male mouse ultrasonic song, and (3) reversed song identical to (2) but backward in time. Sites that did not respond to either sound stimulus were excluded from further analysis. A total of 153 sites were sampled, with 124 of these included in further analysis.

To reconstruct the classical frequency-intensity receptive field (RF) of each site, we presented 25-ms pure tone pips (5-ms squared cosine ramp) to the contralateral ear using a cannulated speaker (EC-1, Tucker Davis Technologies) calibrated to ensure a flat output over the entire frequency range of sounds played. We used tone pips of 22 frequencies (4–73.5 kHz, 0.2 octave spacing) and 7 sound intensities (20–80 dB SPL, 10 dB spacing), repeating each frequency-intensity pair 3 times. Pure tones were played at a rate of 3 pips per second.

Natural sound stimuli for forward and backward vocalizations were obtained from recordings of male mouse courtship vocalizations (Holy & Guo 2005). 26.8 seconds of selected mouse vocalizations (sampling rate 156 kHz, bandpass filtered between 15 and 75 kHz, played at 60 dB), repeated 20 times, were presented while recording from the same site as in (1). The vocalization stimulus included a wide variety of calls and song sequences (described in detail in Holy & Guo (2005)), from frequency-modulated sweeps to complex syllables containing harmonic stacks, upsweeps, downsweeps, and short "grace notes". Due to hardware concerns, the original vocalization waveforms were resampled to 156 kHz and shifted 0.5 octaves down in pitch while maintaining their original temporal rate. Responses were then recorded to 20 repetitions of a backward vocalization stimulus,

which was identical to the previous 26.8-second vocalization but was reversed in time. Since differences in the secondary stimulus statistics may influence the structure of the STRFs calculated from responses to them (Woolley et al. 2006), we calculated the modulation power spectra for both vocalizations (Fig. 2.1B) and backward vocalizations (Fig. 2.1D) to determine whether asymmetries in the presence of upsweeps versus downsweeps could bias our analysis. The modulation power spectrum used here is the amplitude component of the 2D Fourier transform of the stimulus spectrogram, and describes the joint statistics of the spectrotemporal modulations in a sound across its entire duration (Singh & Theunissen 2003). Power in the left quadrant represents power due to upsweeps in the stimulus, whereas power in the right quadrant represents power due to downsweeps. To quantify the asymmetries in the sound stimulus, we calculated the relative power in the first and second quadrants:

$$\alpha_{\text{asym}} = \frac{P_{\text{down}} - P_{\text{up}}}{P_{\text{down}} + P_{\text{up}}} \tag{2.1}$$

where P_{down} is the total power in the right quadrant and P_{up} is the total power in the left quadrant. Values of α_{asym} range between -1 and 1, where a value of 0 indicates that the number of upsweeps and downsweeps in the stimulus is equal. If α_{asym} is positive, there are more downsweeps than upsweeps in the sound, if α_{asym} is negative, there are more upsweeps.

2.3.3 Data analysis

Pure tone data analysis: classical receptive field calculation

To reconstruct the classical frequency-intensity receptive field, we took responses from the pure tone pips described above and counted the number of spikes occurring in response to a given tone within a window defined by the peak of the post-stimulus-time histogram (PSTH). The characteristic frequency for pure tones (RF CF) was identified as the tip of the frequency tuning curve, or the frequency at which the neuron fired responses reliably at threshold intensity level, and the frequency bandwidth was calculated at 30 dB above threshold. The mean firing rate in response to tones was calculated as the average number of spikes fired in the same time window used to calculate the receptive field. The kurtosis of the tone-evoked PSTH was used as a measure for temporal bandwidth.

Vocalization analysis: signal reliability calculation

We used a coherence-based signal-to-noise metric described in Borst & Theunissen (1999) to obtain an upper and lower bound on information carried in the spike train in response to vocalizations for each neuron. Briefly, an estimate of the true signal $s(t)$ was obtained by averaging all trial spike histograms, $R(t)$. The noise estimate $n(t)$ for each trial was then calculated as the residual error after subtracting the signal $s(t)$ from $R(t)$. The signal power $S(f)$ and noise power $N(f)$ as a function of a frequency f were calculated using Welch's method for power spectral density. The

coherence between S and R, which describes the degree of linear relationship between S and R at each frequency f (Hsu et al. 2004), was calculated as follows:

$$\gamma^2_{s,s+n} = \frac{\langle S^2(f)\rangle}{\langle S^2(f)\rangle + \langle N^2(f)\rangle} \tag{2.2}$$

Where $R = S + N$, as stated before. This measure can be used to calculate the information lower bound, measured in bits per second:

$$I = -\int_{f=0}^{F_b} \log_2(1 - \gamma^2_{s,s+n})df \tag{2.3}$$

The integration was performed up to a frequency bound of $F_b = 100$ Hz, after which the relationship between signal and noise power remained flat. To calculate the information upper bound, the same procedure was followed, but the estimate of the noise for each trial was jack-knifed by leaving out data from the trial of interest during the calculation of the mean signal (Hsu et al. 2004). The information metric reported in our results is the mean of the upper and lower information bound.

STRF calculation

To determine the relationship between the vocalization stimuli and the neural response, we calculated the spectrotemporal receptive field (STRF) for each recording site. The STRF is the time-varying stimulus-response function for a neuron, and describes which combinations of acoustical features will excite or inhibit a neuron over time (Aertsen & Johannesma 1981; Eggermont et al. 1983; Theunissen et al. 2001). The relationship between the STRF, H, stimulus, S, and estimated response, $\hat{r}(t)$, are given by the following equation:

$$\hat{r}(t) = r_0 + \sum_{i=0}^{M-1}\sum_{\tau=0}^{N-1} H(\tau, f)\cdot S(t-\tau, f) \tag{2.4}$$

where N is the number of delays of length τ after which the STRF will be estimated (reflecting memory for the stimulus), and M is the number of frequency bands in the spectrogram. The baseline spontaneous firing rate is given by r_0. Positive values of H indicate components of the stimulus that are correlated with increased neuronal firing, and negative values of H indicate components correlated with decreased neuronal firing.

The inputs to the model are the spectrogram of the sound stimulus and the spike times, represented as the peri-stimulus time histogram (PSTH). The spectrogram is chosen as a reasonable representation of the stimulus since auditory neurons are not generally phase-locked to oscillations in the sound waveform, but do respond to changes in sound frequency (Kowalski et al. 1996). The log-magnitude spectrograms for the vocalization and backward vocalizations were calculated using a short-time Fourier Transform with sampling steps of 5 ms in time and 2 kHz in frequency,

and frequency minimum and maximum bounds of 15 and 75 kHz, respectively. Spectrogram plots are shown with a logarithmic y-axis to facilitate interpretation between the STRF and the classical receptive field results. The PSTH was calculated for each recording site by counting the number of spikes in each 5-ms time bin during the trial and averaging across all 20 presentations of the stimulus.

STRF estimation by ridge regression

To solve for the STRF parameters $H(\tau, f)$ directly one may use multi-dimensional linear regression to minimize the mean-square-error between the actual response, $r(t)$, and the response estimated by the STRF, $\hat{r}(t)$, normalizing by the autocorrelation in the stimulus. This is referred to as normalized reversed correlation and is a common method used to estimate STRFs from natural stimuli (Theunissen et al. 2001). However, such an estimate can yield poor predictions as often the stimulus does not adequately sample the full stimulus space, and overfitting of noisy responses is likely to occur. Thus, we chose to use ridge regression, which minimizes the mean-square-error between the actual and estimated response while imposing a constraint on the sum of the regression coefficients with a ridge parameter $\lambda \geq 0$, where increasing values of λ correspond to greater shrinkage of the regression coefficients (Asari & Zador 2009; Machens et al. 2003; Wu et al. 2006). This is also referred to as L2 norm regularization or Tikhonov regularization. Ridge regression minimizes the following error function, where $S(t, f)H(t, f)$ is equal to the estimated response $\hat{r}$ from the STRF:

$$E(H(t, f)) = ||r(t) - S(t, f)H(t, f)||^2 + \lambda||H(t, f)||^2 \tag{2.5}$$

A consequence of regularization via the ridge parameter is a smooth STRF structure, since stimulus correlations are assumed to be smaller. In our implementation, we estimated STRFs using 80% of the data as our training set and 20% as the validation set. We repeated this process for each possible unique 20% validation set. From the training data, we found the optimal ridge parameter for each STRF pair. To do this, we chose 5 random chunks of 500 ms taken from the training data to serve as a held out set. STRFs were estimated for each possible ridge parameter using the remaining training data. Model performance was then evaluated on the held out set. This was performed 25 times to obtain a bootstrapped estimate for the ridge parameter with the best average correlation between the actual and predicted response from the STRF (Efron 1979). For each neuron, the STRF for the forward and backward vocalization were fit using the same ridge parameter, chosen as the value of λ that gave the best mean correlation across the two STRF types (forward and backward) on the held out sets. We sampled 45 ridge parameters, logarithmically spaced from 100 to 105. Constraining the ridge parameters to be identical for forward and backward vocalizations enabled us to more directly compare their STRFs, since any differences in the STRF structure could not be attributed to differences in regularization of the regression coefficients.

STRF validation

STRF performance was calculated for each training and validation set as the raw correlation between the actual response to the validation stimulus and the response predicted by the STRF. The

final STRF was computed as the average of all the STRFs from this jackknifing procedure, and its performance was calculated as the mean of the correlation coefficients for each of the jackknifing iterations. Predicted response traces shown in results figures are all calculated using the intermediate STRFs, where the stimulus chunk for predictions was left out of the STRF calculation.

2.3.4 STRF structure and characteristics

STRF types and clustering

STRFs were automatically assigned to clusters using classical multidimensional-scaling (MDS). The pairwise distance between all STRFs was calculated as the square root of the sum of the squared differences between each pixel in the STRF. This distance matrix was then used as input for the MDS algorithm, which calculates a configuration space in which each STRF, represented by a point, is placed in the space so that closer distances between it and other STRFs indicate that the STRFs are more similar. The first 3 dimensions of the MDS chosen for analysis represent roughly 46% of the variance in the data, with diminishing returns as more dimensions are added to the analysis. To get an idea of the types of STRF structures, we then performed k-means clustering on the points in MDS space to partition the STRFs into clusters, using 5 cluster centroid locations. Since it is clear from our results that the STRFs do not separate evenly into discrete clusters, we do not claim that there are exactly 5 STRF types in our dataset—rather, this method was used as a tool to explore the variety of data structures and find how they were related.

Spectrotemporal separability

The spectrotemporal separability of STRFs is a feature with important implications for neural coding. A STRF is considered fully separable if it can be decomposed into the cross product of one purely spectral and one purely temporal impulse response (Ahrens et al. 2008; Depireux et al. 2001). A neuron with a fully separable STRF would respond equally to upsweeps and downsweeps, with no direction selectivity, whereas direction selectivity would by definition require an inseparable STRF. To describe the separability of our STRFs, we performed singular value decomposition (SVD) on the STRF matrices and calculated the ratio of the first singular value relative to all other singular values, and used these to calculate the degree of inseparability α_{SVD} (Depireux et al. 2001).

$$\alpha_{SVD} = 1 - \frac{\lambda_1^2}{\sum_i \lambda_i^2} \tag{2.6}$$

α_{SVD} ranges in value from 0 to 1, with 0 being fully separable, and values close to 1 indicating higher degrees of inseparability.

STRF characteristic frequency, temporal latency, and spectral and temporal bandwidth, and E-I ratio

To estimate the STRF characteristic frequency (CF), temporal latency, and spectral and temporal bandwidth, we thresholded STRF matrices at 20% of the maximum value to obtain a "cleaned" STRF. For the STRF CF, we calculated the mean by collapsing across all delay bins and found the frequency at which the STRF showed its maximum excitatory response. STRF temporal latency was calculated as the time at which the STRF showed its maximum excitatory response, collapsing across all frequencies. STRF frequency bandwidth was calculated as the width around the peak of this frequency response. STRF temporal bandwidth was calculated similarly by collapsing across all frequency bins and calculating the width of the temporal response. The excitatory-inhibitory (E/I) ratio was calculated as the absolute value of the maximum peak amplitude of the STRF divided by the minimum STRF amplitude, and was performed on the raw STRF.

STRF similarity

To compare STRFs calculated from forward and backward vocalizations, we used a similarity index to describe the pixel by pixel correlation of each pair of STRFs (DeAngelis et al. 1999; Escabi & Schreiner 2002). The SI is calculated as the vector inner product of the two STRFs divided by the product of their vector norms, and is identical to the Pearson correlation coefficient:

$$\text{SI} = \frac{\langle \text{STRF}_A, \text{STRF}_B \rangle}{||\text{STRF}_A|| \cdot ||\text{STRF}_B||} \tag{2.7}$$

An SI of 1 indicates that two STRFs are identical, whereas an SI of 0 indicates that there is no correlation between the two STRFs.

2.4 Results

In this study, we characterized the structure of STRFs calculated in response to natural vocalizations and their time-reversed counterparts. We recorded responses to pure tones, forward vocalizations, and backward vocalizations from 144 sites in 6 mice. Stimulus spectrograms and modulation power spectra, as well as raster plots and PSTHs from one recording site in response to a segment of the forward and backward vocalization are shown in Figure 2.1. The secondary stimulus statistics were similar for forward and backward vocalizations (Fig. 2.1B and 2.1D). The forward vocalization had slightly more power in the left quadrant, corresponding to upsweeps, whereas the backward vocalization showed the opposite. The asymmetry index for forward vocalizations (α_{asym} = -0.0001) and backward vocalization (α_{asym} = 0.0001) was very close to 0, indicating that the stimulus is essentially symmetric.

2.4.1 STRF fitting characteristics

Out of 124 STRFs, 76 STRFs calculated from forward vocalizations and 62 STRFs calculated from backward vocalizations yielded predictions above $r = 0.2$, with 60 sites yielding STRFs with pre-

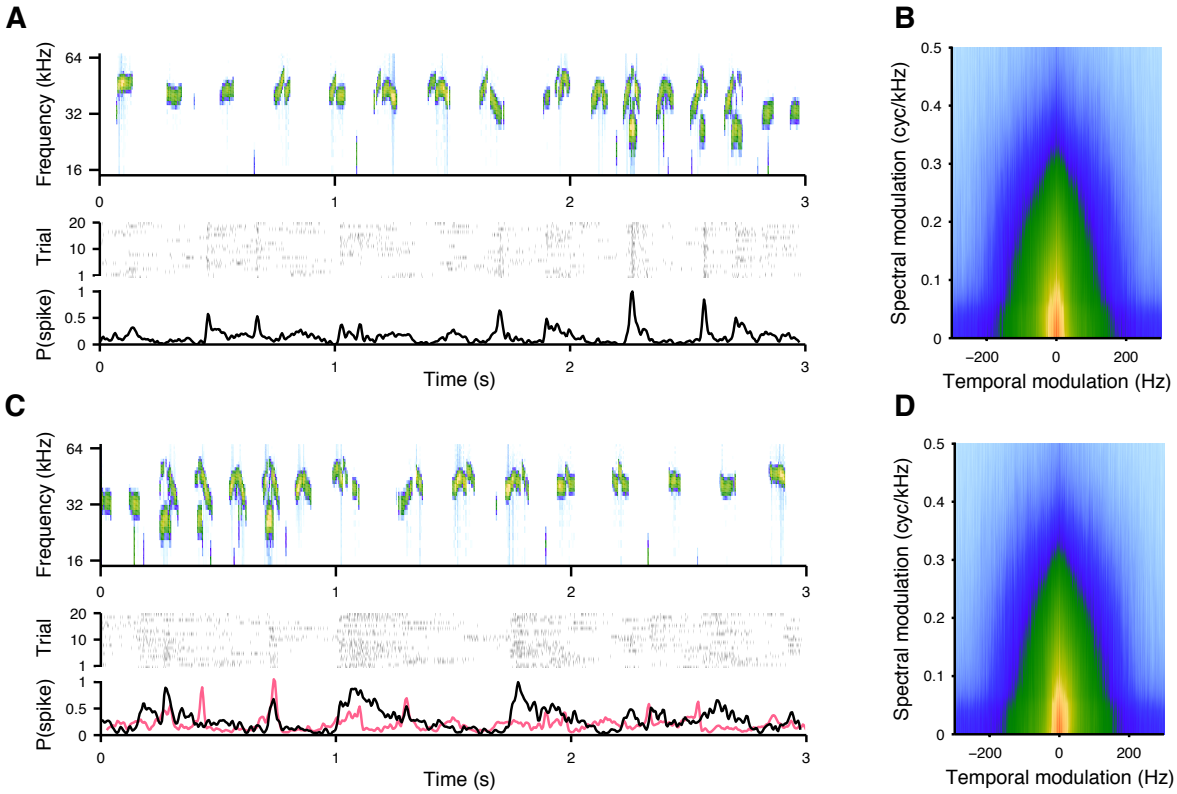

Figure 2.1: Stimulus characteristics and example responses. (A) Top: spectrogram of 3-second segment of forward vocalization, bottom: corresponding spike raster and peri-stimulus time histogram (PSTH). (B) Modulation power spectrum for forward vocalizations. (C) and (D) same as (A) and (B) but depicting stimulus and response properties for backward vocalizations. 3-second segment shown is identical to that shown in (A), but reversed in time. Magenta PSTH overlay shows the same PSTH as in (A) but reversed in time to show how selectivity to individual syllables changes in the forward and backward case.

dictions over this threshold criterion for both forward and backward vocalizations. When comparing STRF properties between backward and forward vocalizations or between STRFs and receptive fields, only STRFs with $r > 0.2$ for both conditions were included in our analysis, except where indicated. Figure 2.2 shows a summary of the STRF fitting characteristics for forward vocalizations only, with a histogram of correlation coefficients for the actual response and the response predicted by the STRF, an example of a predicted and actual response to a segment of the forward vocalization, and relationships between STRF predictability and response characteristics. STRF prediction performance was significantly correlated with the average firing rate to the vocalizations (Fig. 2.2C, $R^2 = 0.22, p = 1.1 \times 10^{-6}$), with higher mean firing rates yielding better STRFs. We hypothesized that neurons with a lower pure tone characteristic frequency (RF CF) might show lower predictions since the vocalization frequencies would be outside of the neurons' receptive fields. This did indeed appear to be the case, since STRF performance was positively correlated with the CF in response to pure tones (Fig. 2.2D, $R^2 = 0.32, p = 8.1 \times 10^{-10}$). Still, some neurons with low RF CFs showed strong prediction performance, on the other hand, there were mid- and high-frequency RF CF neu-

rons that yielded poor STRFs. Trial-by-trial signal repeatability was positively correlated with STRF prediction performance (Fig. 2.2E, $R^2 = 0.52, p = 4.1 \times 10^{-24}$), indicating that the more robust and reliable the neural response, the better the STRF.

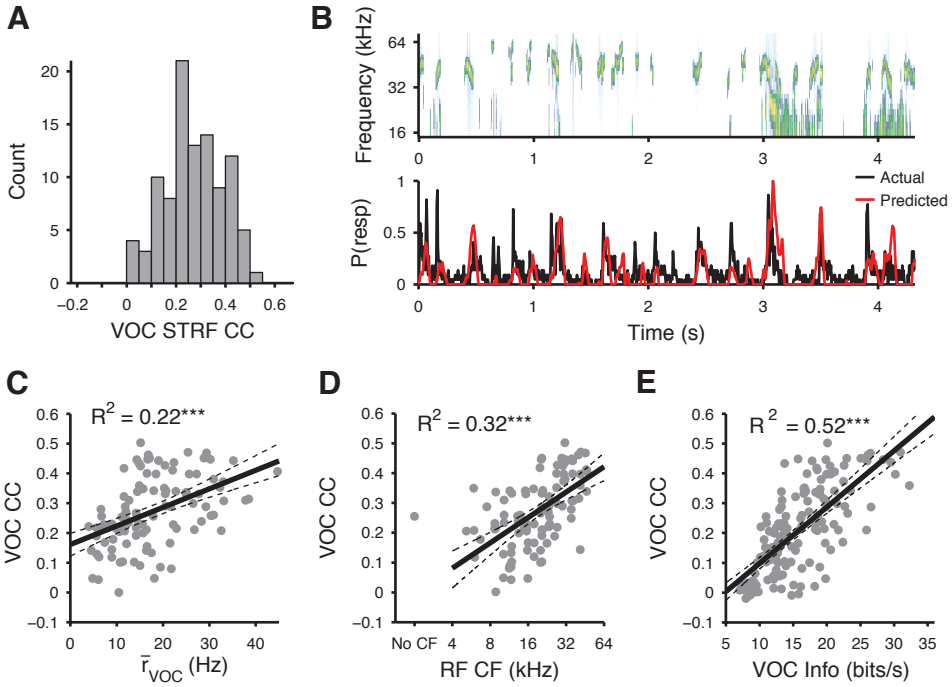

Figure 2.2: STRF fitting characteristics. (A) Histogram of STRF prediction accuracy, as measured by correlation (CC) between actual response and response predicted by the STRF. (B) Top panel: spectrogram of ≈ 4 s segment of the vocalization stimulus. Bottom: example predicted (red) and actual neural response (black) to the stimulus above. (C) VOC STRF performance (correlation coefficient) is positively correlated with mean spike rate in response to vocalizations ($R^2 = 0.22, p = 1.1 \times 10^{-6}$). (D) STRF correlation coefficient versus characteristic frequency (CF) of classical receptive field. Sites with higher CFs tended to produce STRFs with better prediction performance ($R^2 = 0.32, p = 8.1 \times 10^{-10}$). (E) STRF correlation coefficient versus cross-trial signal repeatability of the neural response (expressed as information in bits/s). High response repeatability was strongly positively correlated with STRF performance ($R^2 = 0.52, p = 4.1 \times 10^{-24}$). $^\star\, p < 0.05$, $^{\star\star}\, p < 0.01$, $^{\star\star\star}\, p < 0.001$.

2.4.2 STRF types in the auditory cortex

We estimated the STRF from responses to trains of vocalizations with complex spectrotemporal interactions. These vocalizations, described in detail in Holy and Guo (2005) include a number of syllables presented as natural trains. Syllable types included frequency-modulated sweeps, harmonic stacks, and frequency jumps. The resulting STRFs showed that neurons in the auditory cortex respond to specific elements of the vocalizations and that their responses may be limited temporally, spectrally, or both. Furthermore, these responses do not necessarily overlap with the response frequency range predicted by responses to pure tones (Fig. 2.3A). To classify STRF types into groups

automatically, we used multi-dimensional scaling coupled with k-means clustering to assign each STRF to a class (see Methods). In Fig. 2.3B, each point in the scatterplot represents one STRF, and the distance between points is related to how similar a STRF is to others in the dataset. While the orientation of the MDS axes is arbitrary, there is a clear progression of temporal bandwidth from low to high and from low to high receptive field characteristic frequency within the MDS space, as indicated by the arrows. Though the STRFs themselves do not obviously separate into distinct clusters, this analysis allows us to investigate the multidimensional space describing STRF structure in a principled way. For example, in the five classes of STRF shown in Fig. 2.3, sites that responded to middle frequencies (between 16 and 32 kHz, panel i) in the vocalization over a wide temporal range did not respond reliably to pure tones at all. Other sites that responded robustly to low frequency pure tones between 4 and 16 kHz showed responses to much higher frequencies in the vocalization STRF (Fig. 2.3A, panel ii), though it is possible that the lack of overlap in the low frequencies is due to the inability to sample those frequencies with the vocalization sounds, which were band-pass filtered between 15 and 75 kHz. Some STRFs showed excitatory responses both in and out of the classical receptive field for the neuron (Fig. 2.3A, panel iii), whereas others responded almost precisely at the CF or within the classical receptive field for the neuron (Fig. 2.3A, panels iv and v).

An important property of the STRF that also describes how a neuron responds to joint spectral and temporal modulations is its separability. If a STRF can be described as the cross product of one purely spectral and one purely temporal component, then it is considered to be fully separable (Depireux et al. 2001). More complex STRFs, including those that are direction selective (to either up-sweeps or down-sweeps) cannot be fully separable by definition. Thus, the separability of a neuron's STRF has implications for how that neuron encodes stimulus features. Here we calculated the degree of inseparability α_{SVD} of the STRFs by performing singular value decomposition (SVD) on each STRF and calculating the ratio of the squared weight of the first singular value to the sum of the squared weights for all the singular values (see Methods). The larger the first singular value is relative to the others, the more separable the STRF. The index α_{SVD} ranges from 0 to 1, with 0 being fully separable into a frequency and temporal component, and values approaching 1 indicating increasing inseparability. By convention, values higher than 0.2 are considered more "inseparable", though there is no clear boundary between separable and inseparable STRFs. The STRFs we calculated from responses in auditory cortex ranged from highly separable (Fig. 2.4A) to moderately inseparable (Fig. 2.4C). This can be observed intuitively in Figs. 2.4B and 2.4D, where the spectral and temporal weights (calculated as the mean STRF weights collapsed across the opposite dimension, shown in purple and green in Fig. 2.4A and 2.4C) are combined with the cross-product to create a fully separable STRF. In Fig. 2.4B, the resulting fully separable STRF is highly correlated with the actual STRF (Fig. 2.3F, $R^2 = 0.67$), whereas in Fig. 2.4C, the fully separable STRF looks very little like the original STRF in 4C ($R^2 = 0.005$), which shows sensitivity to a harmonic stack that includes an upsweep in its mid-frequencies. While the fully separable STRF in 4D captures the flat frequency response in the high frequency range (just below 64 kHz), it completely obscures the response to sweeps. The overall separability of the entire population of STRFs is shown in Fig. 2.4E, where many STRFs were comparably separable to the example shown in Fig. 2.4A, but others were more inseparable. The separability of the STRFs was not significantly correlated with STRF performance ($R^2 = 0.02$, $p = 0.31$, data not shown).

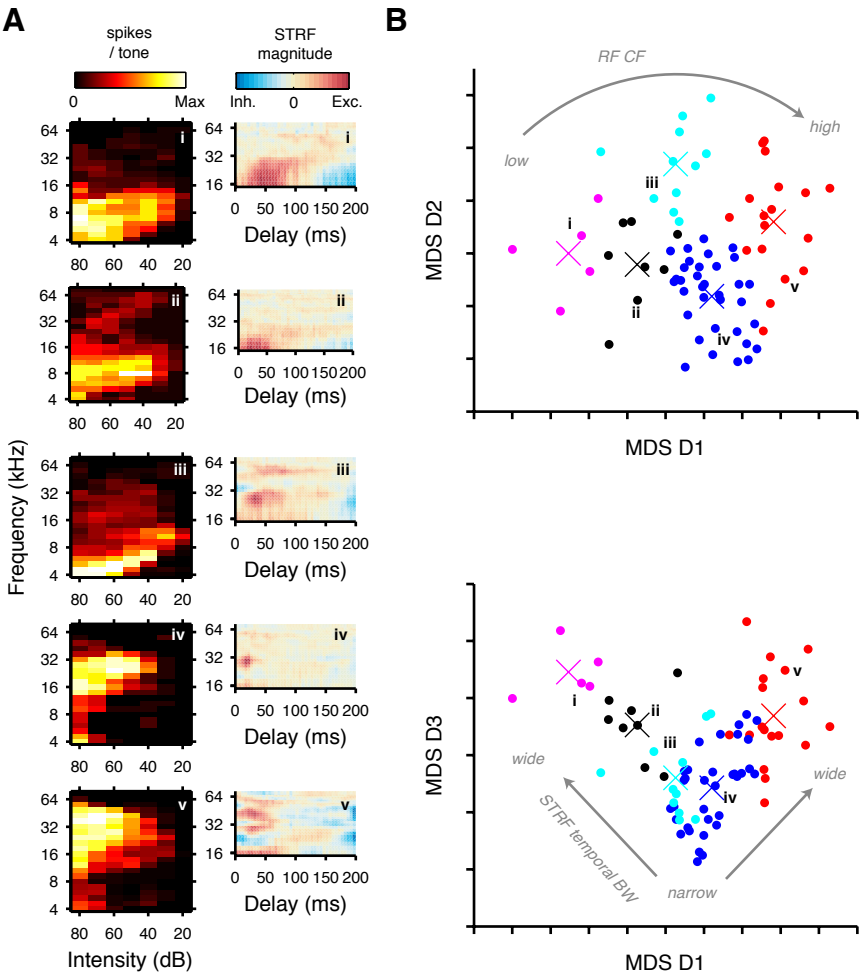

Figure 2.3: Example classes of STRFs calculated from responses to vocalizations. Stimulus parameters resulting in increased firing (excitation) are shown in red, parameters resulting in decreased firing (inhibition) are shown in blue. (A) Five examples of different types of STRF (left) and their corresponding receptive fields (right): (i) temporally broad responses to lower frequency features, (ii) more temporally and spectrally narrow responses to lower frequency features (iii) temporally restricted responses to two distinct frequency components, (iv) similar to (iii), but more restricted in frequency and temporal response, (v) temporally restricted response to harmonic stacks in the vocalization. Prediction performance for each STRF is as follows: (i) CC = 0.31, (ii) CC = 0.26, (iii), CC = 0.33, (iv) CC = 0.45, (v) CC = 0.47. (B) Multidimensional scaling (MDS) analysis of STRFs separates STRFs into clusters shown in (A). First three dimensions of MDS analysis are shown.

2.4.3 Comparison of classical receptive field properties with STRF properties

Our data suggest that neurons may be specialized for natural sounds, as some neurons did not respond to pure tones but responded robustly to vocalization stimuli, and some neurons responded to vocalization stimuli outside their classical pure tone-derived receptive fields. We compared the responses to pure tones and vocalization stimuli, and found that the temporal response parameters

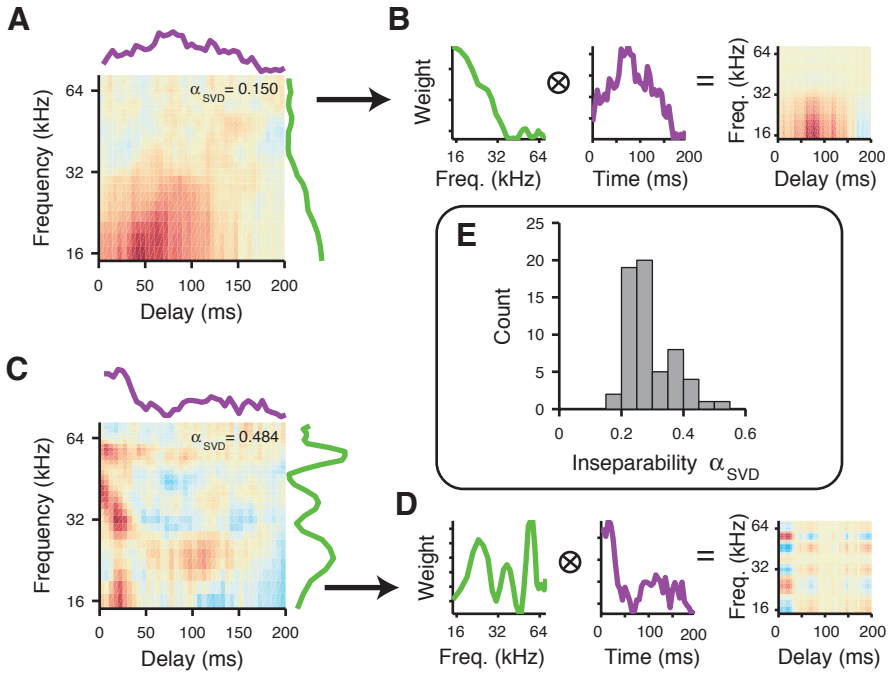

Figure 2.4: STRFs calculated in response to vocalizations vary in linear inseparability. (A) Example of highly separable STRF ($\alpha_{SVD} = 0.15$), along with spectral and temporal components. (B) The resulting linearly separable STRF calculated by taking the cross product of the spectral and temporal mean response shown in (A). This fully separable STRF is well correlated with the original STRF ($R^2 = 0.67$). (C) Example of inseparable STRF ($\alpha_{SVD} = 0.48$), with spectral and temporal components. (D) Same as (B) for STRF in (C)—this second example, calculated from the inseparable STRF, is a poor match for the original STRF ($R^2 = 0.005$). (E) Histogram of inseparability index, α_{SVD} shows that STRFs in auditory cortex range from highly separable (α_{SVD} close to 0) to moderately inseparable ($\alpha_{SVD} > 0.2$ is commonly considered more inseparable).

tended to be better correlated between the two stimulus types than frequency response parameters. Mean firing rates in response to pure tones were significantly positively correlated ($R^2 = 0.43$, $p = 1.6 \times 10^{-10}$) with mean firing rates in response to the vocalization stimulus (Fig. 2.5A). The STRF characteristic frequency (CF, Fig. 2.5B) was only weakly correlated with CF for the pure-tone receptive field (RF) ($R^2 = 0.06$, $p = 0.04$). STRF temporal latency (Fig. 2.5C) was positively correlated with the latency in response to pure tones ($R^2 = 0.32$, $p = 9.0 \times 10^{-8}$) and the values did not differ significantly ($p = 0.17$, Wilcoxon signed-rank test). STRF frequency bandwidth (Fig. 2.5D) was not significantly correlated with frequency bandwidth for the classical receptive field ($R^2 \approx 0$, $p = 1$). These observations are corroborated by visual inspection of the STRFs in Fig. 2.3, which shows that some STRF responses overlap significantly with the frequency range of the classical RF, but others may not overlap at all. This suggests that a different encoding strategy may be used to encode natural versus synthetic sounds. STRF temporal bandwidth was significantly negatively correlated with the kurtosis of the PSTH in response to tones (Fig. 2.5E, $R^2 = 0.17$, $p = 2.6 \times 10^{-4}$), meaning neurons with more temporally limited (higher kurtosis) responses for

pure tones also exhibited more temporally limited excitation in their response to the vocalization.

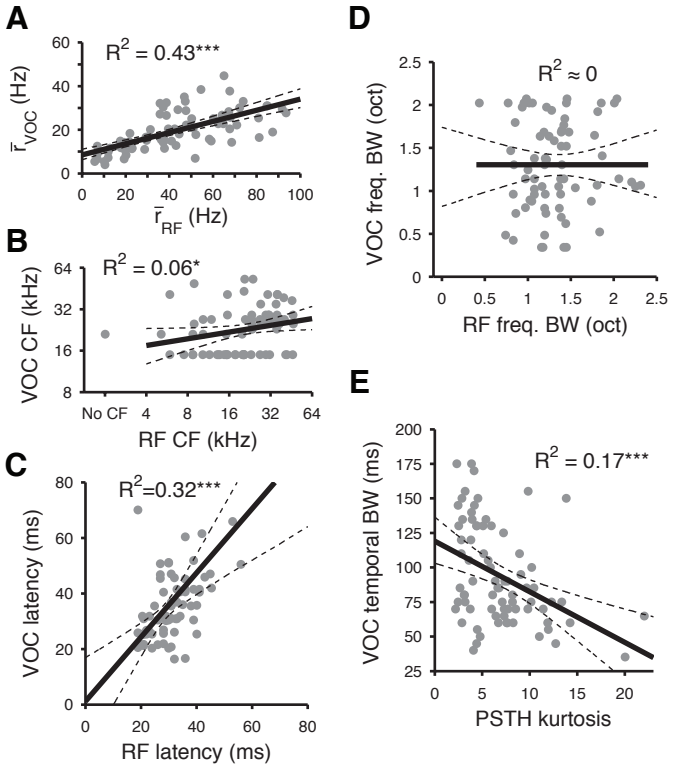

Figure 2.5: Comparison of classical receptive field and pure tone response properties with STRFs calculated from vocalizations. Solid black line indicates least-squares regression line, dashed curves indicate boundaries of bootstrapped 95% confidence band for the regression line. (A) Mean firing rate of neurons in response to vocalizations is significantly positively correlated ($R^2 = 0.43$, $p = 1.6 \times 10^{-10}$) with firing rate of neurons in response to pure tones. (B) STRF characteristic frequency (CF) is weakly correlated with receptive field (RF) characteristic frequency ($R^2 = 0.06$, $p = 0.04$). (C) STRF temporal latency is positively correlated with latency in response to pure tones ($R^2 = 0.32$, $P = 9.0 \times 10^{-8}$). (D) STRF frequency bandwidth (BW) is not correlated with receptive field frequency bandwidth ($R2 \approx 0$, $p = 1$). (E) STRF temporal bandwidth is significantly negatively correlated ($R^2 = 0.17, p = 2.6 \times 10^{-4}$) with the peakiness (as measured by kurtosis) of the post-stimulus time histogram (PSTH) calculated for pure tones. Higher values of PSTH kurtosis correspond to smaller temporal bandwidths for tones, thus, temporal bandwidths of responses to vocalizations are positively correlated with temporal bandwidths of pure tone responses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2.4.4 Comparison of forward and backward vocalization STRFs

Though it is clear that vocalizations may elicit strong responses even in neurons that would not otherwise respond to individual frequency components of the stimulus, we wanted to investigate whether a sound with similar secondary statistical structure would be encoded similarly by cortical neurons. Thus, we recorded responses to the same ultrasonic vocalization trains, but reversed

in time, which we call backward vocalizations (VBK). This alters the temporal structure of the vocalization while preserving its spectral characteristics. Still, as shown in Fig. 2.1B and 2.1D, the secondary stimulus statistics are similar for the forward and backward vocalizations, with a very slight increase in the power in the left quadrant, corresponding to upsweeps, for the forward vocalization (α_{asym} = -0.0001), and vice versa in the backward vocalization (α_{asym} = 0.0001). To test whether this small asymmetry would affect the structure of the STRFs, we calculated VBK STRFs using the stimulus subspace from the forward vocalization, as is done for stimuli with very different secondary stimulus statistics (Woolley et al. 2006), and there was no difference in STRF structure (data not shown). The STRFs reported here were calculated using their own stimulus subspace (that is, VOC for VOC, and VBK for VBK). Despite the almost identical secondary stimulus statistics for forward and backward vocalizations, the STRFs calculated for the two stimuli were often different, suggesting stimulus specificity in the neural response or nonlinear interactions between the stimulus and response. Example STRFs for forward and backward vocalizations at four different sites in the auditory cortex are shown in Figure 2.6, along with their corresponding classical receptive fields.

For many STRFs, the features to which the neuron responded for the forward vocalization appeared different from those to which it responded for the backward vocalization (Fig. 2.6, rows iii and iv), and for other STRFs, response patterns seemed similar (Fig. 2.6, rows i and ii). For example, in (ii), the direction selectivity of the neuron changes slightly for forward and backward vocalizations—in the forward case, the neuron responds robustly by increasing its spike rate for a vocalization with an up and then a downsweep centered around 32 kHz, and in the backward case, the neuron increases its firing rate in response to a downsweep (recall that these STRFs are displayed with the delay as the x-axis, thus the stimulus that evokes a response would be the reflection of the STRF about the y-axis). Still, the frequency and temporal bandwidths for the two STRFs are similar. In the third panel of Fig. 2.6, which corresponds to the same site whose response trace and raster plots are shown in Fig. 2.1, the STRF structure overlaps somewhat, but is more temporally smeared for the backward vocalization. Furthermore, the inhibitory areas between 16 and 32 kHz in the VOC STRF disappear in the VBK STRF. In one of the more extreme examples, the temporal bandwidth is narrower in the VBK STRF compared to the VOC STRF, and the structure of the harmonic stacks within the VOC STRF is markedly different from that in the VBK STRF. The distribution of similarity indices for the VOC and VBK STRFs can be seen in Fig. 2.7B. STRFs ranged in similarity from moderately similar (SI = 0.4) to highly similar (SI = 0.8).

Despite the differences in STRF structure, the prediction performance of VOC and VBK STRFs was highly correlated (Fig. 2.7A, $R^2 = 0.79$, $p = 7.7 \times 10^{-50}$), and there was no significant difference between the performance of the VOC and VBK STRFs across cortical sites ($p = 0.71$, Wilcoxon signed-rank test). This suggests that differences seen in the structure of the STRFs were not due to poor fitting of the response data. In addition, using VBK STRFs to predict VOC responses and vice versa showed that responses from one stimulus were not adequate to predict the other (Fig. 2.7C and 2.7D). That is, the prediction performance of VBK STRFs when predicting forward vocalization responses was significantly worse than predicting forward vocalization responses from VOC STRFs (Fig. 2.7C, $p = 0.002$, Wilcoxon signed-rank test). Likewise, VOC STRFs predicted responses to the backward stimulus significantly more poorly than when using the VBK STRFs (Fig.

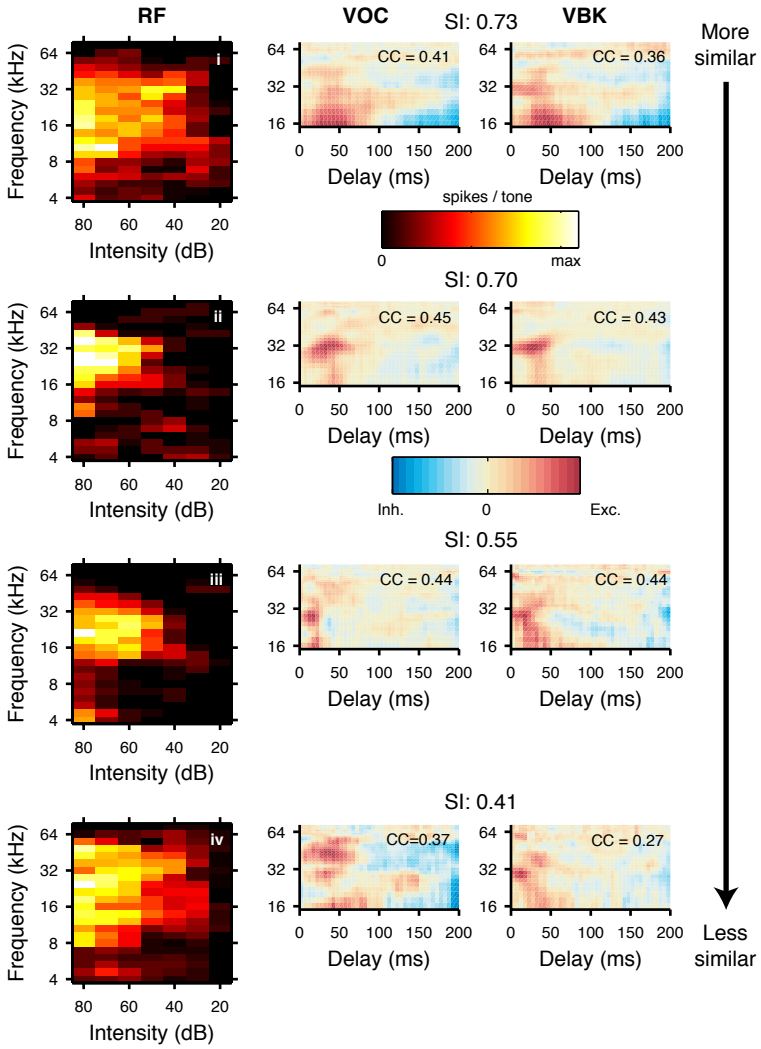

Figure 2.6: Example STRFs for forward vocalizations (VOC) and backward vocalizations (VBK) and corresponding classical receptive field from responses to pure tones (RF). Four rows show STRFs and RFs from four different sites in the auditory cortex. The STRFs shown in row (iii) were taken from the same cell whose responses are shown in Figure 2.1. STRFs are ordered according to their similarity indices, with more similar STRFs at the top of the plot and less similar STRFs at the bottom.

2.7D, $p = 1.4 \times 10^{-9}$, Wilcoxon signed-rank test).

To quantify the changes observed in the gross structure of the STRFs, we compared the same response properties from the pure tone analysis in Fig. 2.5 between VOC and VBK STRFs. In general, response parameters including CF, temporal bandwidth, and excitatory-inhibitory (E/I) ratio were positively correlated between the two STRF types (Fig. 2.8), but the values themselves differed markedly across individual STRFs. As was the case in the pure tone analysis, temporal response parameters for the VOC and VBK STRFs appeared to be more similar than spectral response pa-

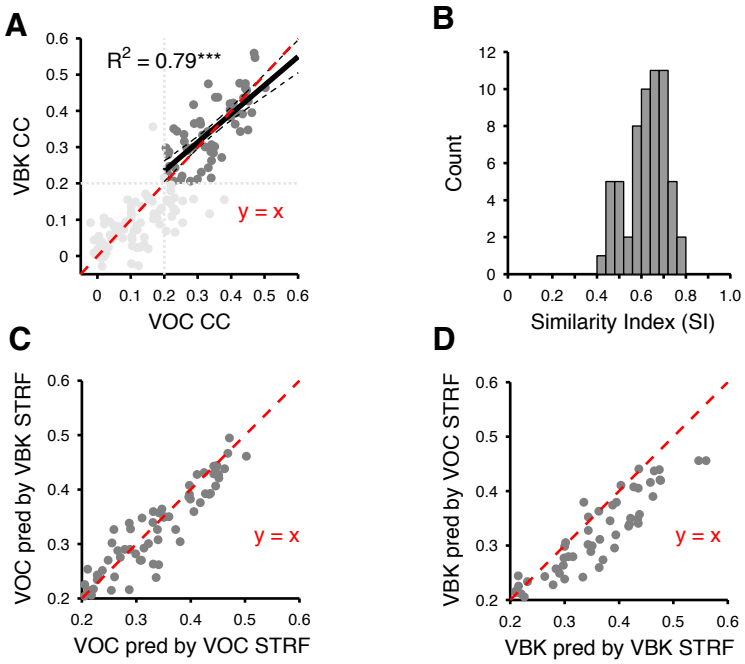

Figure 2.7: (A) Relationship between STRF prediction accuracy (correlation between predicted and actual response, CC) for VOC and VBK. STRF prediction performance was positively correlated between VOC and VBK ($R^2 = 0.79, p = 7.7 \times 10^{-50}$) and not significantly different between the two conditions ($p = 0.71$, Wilcoxon signed-rank test). STRFs with good performance in both VOC and VBK conditions were used in subsequent analyses and are shown in dark gray. (B) Distribution of similarity indices (SI) of STRFs for forward and backward vocalizations. SI ranges from 0 to 1, with 0 being most dissimilar and 1 indicating that the two STRFs are identical. SIs were calculated for STRF pairs where both had performance correlations > 0.2. (D and E) Cross-stimulus prediction performance indicates that forward vocalizations are not well predicted from backward vocalization responses and vice versa. (C) Response prediction performance was significantly greater for forward vocalizations when using the VOC STRF than when using the VBK STRF ($p = 0.002$, Wilcoxon signed-rank test). (D) Response prediction performance was significantly greater for backward vocalizations when using VBK STRFs over VOC STRFs ($p = 1.4 \times 10^{-9}$, Wilcoxon signed-rank test). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

rameters. STRF CF (Fig. 2.8A) was significantly different for VOC and VBK STRFs ($p = 0.007$, Wilcoxon signed-rank test), but linearly correlated ($R^2 = 0.16, p = 0.001$) where VOC CFs were on average lower than VBK CFs. There was no consistent relationship between frequency bandwidths for the VOC and VBK STRFs (Fig. 2.8B, $R^2 = 0, p = 0.97$) though the VOC STRFs exhibited a wider range of possible bandwidths compared to VBK, which were more spectrally limited. STRF latencies were strongly correlated for VOC and VBK (Fig. 2.8C, $R^2 = 0.30, p = 5.5 \times 10^{-6}$) and did not differ significantly ($p = 0.91$, Wilcoxon signed-rank test). The temporal bandwidths were also similar for the two STRF types, and were both highly correlated (Fig. 2.8D, $R^2 = 0.34$, $p = 1.2 \times 10^{-6}$) and similar in value ($p = 0.10$, Wilcoxon signed-rank test). The ratios of the excitatory to inhibitory peak of the VOC and VBK STRFs were also positively correlated (Fig. 2.8E, $R^2 = 0.13, p = 0.005$), and were mostly greater than 1, indicating more pronounced excitation

compared to inhibition in the STRFs overall. In general, we saw very few STRFs with strong inhibitory bands, though those that did exhibit this characteristic tended to show spectrally-limited inhibition in the middle of a harmonic stack (Fig 2.3A, panel iii and v). Finally, VOC and VBK STRFs showed similar degrees of inseparability (Fig. 2.8F, $p = 0.28$, Wilcoxon signed-rank test) that were linearly correlated ($R^2 = 0.35$, $p = 6.8 \times 10^{-7}$), indicating that the relative complexity of the STRFs was similar for the two stimuli.

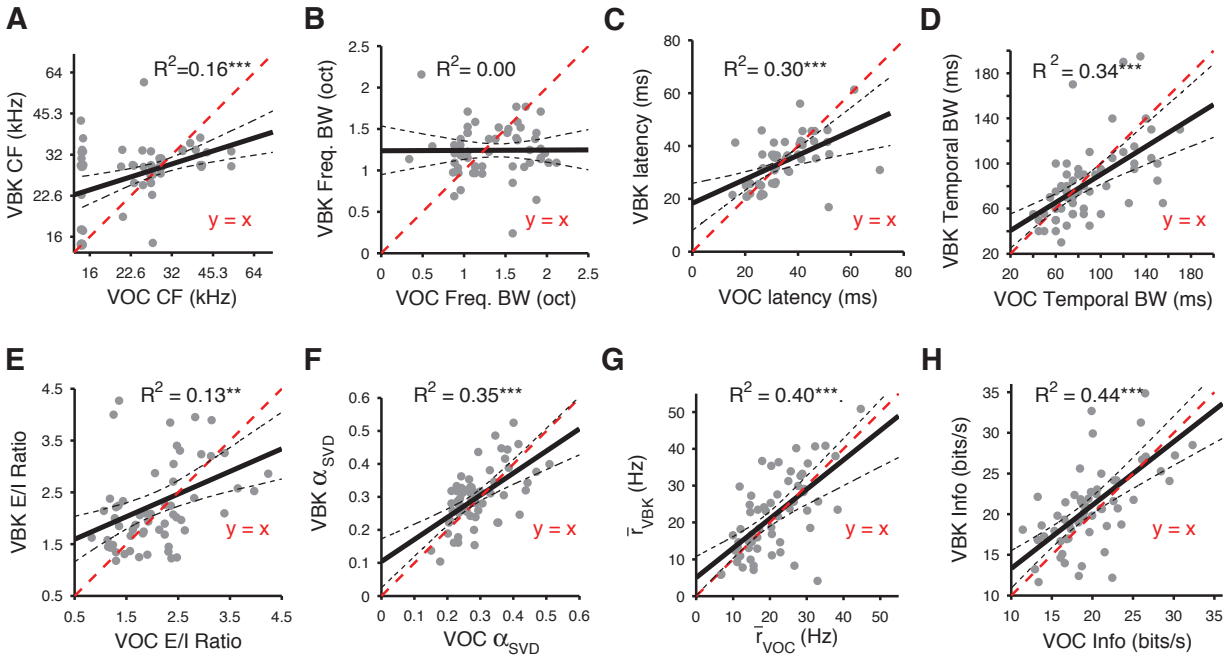

Figure 2.8: Comparison of STRF characteristics for forward (VOC) and backward (VBK) vocalizations. Despite similar mean firing rate and response reliability, the structure of the STRFs calculated from forward and backward vocalizations differed substantially. Solid black lines indicate least squares regression line; dashed black lines indicate the boundaries of the bootstrapped 95% confidence band for the regression. **(A)** STRF characteristic frequency (CF) for VOC and VBK STRFs differ significantly ($p = 0.007$, Wilcoxon signed-rank test) but are positively correlated ($R^2 = 0.16$, $p = 0.001$); **(B)** STRF frequency bandwidths (BW) are not significantly correlated ($R^2 = 0$, $p = 0.97$) for forward and backward vocalizations. **(C)** STRF temporal latencies were strongly correlated for VOC and VBK STRFs ($R^2 = 0.30$, $p = 5.5 \times 10^{-6}$) and were not significantly different ($p = 0.91$, Wilcoxon signed-rank test); **(D)** STRF temporal bandwidths are strongly positively correlated ($R^2 = 0.34$, $p = 1.2 \times 10^{-6}$) but are not statistically different ($p = 0.10$, Wilcoxon signed-rank test); **(E)** ratio of the excitatory and inhibitory peak amplitudes of the STRF (E/I ratio) are positively correlated ($R^2 = 0.13$, $p = 0.005$); **(F)** Degree of inseparability was positively correlated ($R^2 = 0.35$, $p = 6.8 \times 10^{-7}$) and similar for VOC and VBK STRFs ($p = 0.28$, Wilcoxon signed-rank test); **(G)** Mean firing rate for vocalizations and backward vocalizations was similar ($p = 0.31$, Wilcoxon signed-rank test) and highly correlated ($R^2 = 0.40$, $p = 5.6 \times 10^{-8}$); **(H)** Mean information (as measured by trial-by-trial coherence) conveyed by the spike train for VOC and VBK is slightly higher for VBK compared to VOC ($p = 0.046$, Wilcoxon test) and strongly positively correlated ($R^2 = 0.44$, $p = 7.5 \times 10^{-9}$). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The STRF gives insight into the fine details of how a neuron responds to a stimulus in a way that

might not otherwise be seen when looking at mean firing rate or response reliability. Indeed, in our neurons there was no significant difference in mean firing rate for forward or backward vocalizations (Fig. 2.8G, $p = 0.31$, Wilcoxon signed-rank test) and the firing rates were well correlated ($R^2 = 0.40$, $p = 5.6 \times 10^{-8}$), which demonstrates that neurons may preferentially encode different features of these stimuli without an overall increase or decrease in spike rate. Signal repeatability (Fig. 2.8H), as measured by information in the spike train (see Methods) was highly correlated for VOC and VBK ($R^2 = 0.44$, $p = 7.5 \times 10^{-9}$), with a slight increase in information in the spike trains for backward vocalizations ($p = 0.046$, Wilcoxon signed-rank test). Furthermore, there was no significant difference in the sparseness of the neural response to forward and backward vocalizations as measured by PSTH kurtosis ($p = 0.1$, Wilcoxon signed-rank test).

2.5 Discussion

In natural environments, animals encounter a rich variety of sounds that are highly structured and non-random. Those sounds vary in both frequency and time, and the ability to represent and process their joint spectrotemporal modulations is crucial for communication and survival. Though it is reasonable to assume that the auditory system evolved to represent the statistical structure of the natural acoustic environment in an efficient manner (Lewicki 2002; Smith & Lewicki 2006), historically it has been difficult to parameterize natural sounds in a systematic way, and thus many researchers have focused their efforts on describing neural responses to synthetic, easily manipulated sounds including pure tones, white noise, chirps, dynamic ripples, and temporally orthogonal ripple combinations (TORCs) (Ahrens et al. 2008; Blake & Merzenich 2002; Depireux et al. 2001; Escabi & Schreiner 2002; Fritz et al. 2003; Han et al. 2007; Kilgard & Merzenich 1999; Kim & Bao 2009; Linden & Schreiner 2003; Machens et al. 2003). That work has been crucial to our understanding of sound encoding in general (Polley et al. 2007; Schreiner & Sutter 1992). Nevertheless, many studies have shown that neural responses to synthetic sounds may be inadequate to describe how neurons encode natural sounds (Kao et al. 1997; Woolley et al. 2006). This study extends those findings by comparing cortical feature selectivity for natural versus reversed vocalizations, and natural versus synthetic sounds.

2.5.1 Responses to natural vocalizations outside of the classic receptive field

By recording responses to pure tones, ultrasonic courtship vocalizations, and their time-reversed counterparts in the mouse auditory cortex, we were able to directly compare how the brain processes synthetic versus natural and altered natural sounds in the joint spectrotemporal domain. We found that vocalizations strongly drove spiking in auditory cortical neurons. Neurons responded preferentially to the vocalizations even if the frequency range of the vocalization was well outside the neuron's classical receptive field (Fig. 2.3). We found that spectral parameters were not well correlated with STRF RF frequency parameters, but that temporal response properties seemed to be better matched. In particular, CF and frequency bandwidth differed greatly for pure tone and vocalization responses, but temporal latency, temporal bandwidth, and mean firing rate were relatively consistent. Similar phenomena have been reported in the mouse inferior colliculus (IC)

(Holmstrom et al. 2010; Portfors et al. 2009) and monkey auditory cortex (Rauschecker et al. 1995), where robust responses to vocalizations have been observed in neurons that did not respond to pure tones in the same frequency range. It remains to be determined whether the mismatch observed in the cortex is a feed-forward manifestation of similar phenomena in subcortical auditory nuclei.

2.5.2 Different feature selectivity for forward versus backward vocalizations

Although the forward and backward vocalizations in our study were well matched in terms of the presence of upsweeps and downsweeps, they elicited different responses in the auditory cortex (Fig. 2.1). To compare spectrotemporal feature selectivity, we derived STRFs separately with responses to forward and backward vocalizations using cross-validated ridge regression with matched constraints, and quantified STRF properties including CF, latency, frequency and temporal bandwidth, and E-I ratios. As shown in Fig. 2.8, the CFs of the VOC STRFs and VBK STRFs were very different in some neurons. Although there was no consistent relationship between VOC and VBK STRF frequency bandwidth (that is, frequency bandwidth was not systematically wider or narrower in either case), the range of frequency bandwidths was larger in the VOC STRFs, indicative of a more heterogeneous set of STRFs across the neural population for better stimulus encoding (Holmstrom et al. 2010). These results indicate that neurons can respond more strongly to different spectral elements of a sound if they occur in different temporal contexts, and that the frequency range over which they respond may also depend on the precise relationship between the spectral and temporal components of the sound. This finding is particularly striking because only the temporal order of frequency presentation changed across the two stimulus conditions, and we did not manipulate the spectral content of the vocalization in any way before reversing it. It should be noted that interpreting changes in CF may be complicated by the presence of many STRFs with complex spectral structure that is not adequately described by a CF and frequency bandwidth alone (Fig. 2.3A, panels iii and v, and Fig. 2.6, panels i, iii, and iv). For example, a change in CF could manifest from a VOC and VBK STRF with harmonic stacked structure (as in Fig. 2.3A, panel v) where there is a shift in power from one harmonic to another. Thus, these changes should be taken together in context with changes in bandwidth, similarity index, and other measures. Still, the responses differed enough that VBK STRFs were significantly worse at predicting responses to forward vocalizations and vice versa (Fig. 2.7C and 2.7D).

Temporal response properties for VOC and VBK STRFs were more similar than frequency response properties, with a strong positive correlation between temporal latency (Fig. 2.7C) and temporal bandwidth for both VBK and VOC STRFs (Fig. 2.7D). This shows that reversing the vocalization can shift neural responses spectrally while maintaining selectivity for similar temporal modulations. Temporal response properties are closely related to the intrinsic biophysical properties of the neurons (Atencio & Schreiner 2008), whereas spectral responses may be more affected by the strength and summation of a neuron's synaptic inputs. Furthermore, spectral responses may be more likely modified according to temporal properties including order, repetition rate, or coincidence of sound presentation (Chang et al. 2005; Schneider & Woolley 2011).

The stability of spectral and temporal characteristics of STRFs may depend on the stimulus and the context/task. (Fritz et al. 2005) reported changes in both the spectral and temporal parame-

ters of the STRF during frequency discrimination tasks in the awake ferret primary auditory cortex. Though those authors did not compute CF, frequency bandwidth, and temporal bandwidth as in the current study, a qualitative assessment of their STRFs seems to indicate that shifts in spectral tuning are more striking than changes in temporal bandwidth for their task. In the frequency discrimination task, they showed facilitation at the target frequency and suppression at a reference tone frequency. In a later study, the same laboratory showed that temporal plasticity in A1 can be induced by training ferrets on a gap detection task—in this case, shifts in latency and temporal bandwidth were most pronounced (Fritz et al. 2007). In the absence of a behavioral task, these researchers showed that most A1 neurons in the ferret exhibit stable STRFs (calculated in response to TORCs) over time scales of 30-120 minutes, although others have shown spontaneous variability in the spectral and temporal domains (Elhilali et al. 2007).

Similar overall firing rates to forward and backward vocalizations. Despite the clear differences in STRF structure, the mean firing rate in response to forward and backward vocalizations was similar in our study. Previous reports have shown weaker (Šuta et al. 2003), stronger (Šuta et al. 2003, 2007; Wang & Kadia 2001), or similar responses (Gehr et al. 2000; Glass & Wollberg 1983; Šuta et al. 2003) to forward over reversed vocalizations. Our results indicate that a lack of gross-scale changes in firing rate or response magnitude does not preclude the existence of changes in fine timing or shifts in power in the spectral response.

2.5.3 Temporal asymmetry of the vocalizations

For auditory neurons to respond differently to forward versus backward vocalizations, the vocalizations must be asymmetrical. One type of asymmetry could arise from a difference in the presence of upsweeps and downsweeps, another could emerge from differences in the relative ratios of attack and decay of each syllable. Vocalization asymmetry may vary greatly across different species. For example, while the mouse courtship vocalizations used in the present study are largely symmetrical in terms of the presence of upsweeps and downsweeps (Fig. 2.1), some primate calls and bird songs are highly asymmetrical (Schnupp et al. 2006; Singh & Theunissen 2003; Wang et al. 2005). In addition, some vocalizations, such as marmoset calls, have salient attacks at their onset followed by slow decays (Wang et al. 1995). Auditory neurons generally respond strongly to acoustic transients such as these asymmetric features, which could likely contribute to selectivity for forward vocalizations compared to time-reversed ones. Although rodents' cortical neurons respond equally strongly to forward and reversed conspecific calls, they are nonetheless sensitive to stimulus asymmetry. For example, rats can discriminate human speech of different languages when it is played forward but not in reverse (Toro et al. 2005, 2003). Thus, the properties of the vocalization stimulus itself, such asymmetry in upsweeps and downsweeps and the sound attack and decay, may play a role in generating different responses to the reversed vocalizations.

2.5.4 Species-specificity versus experience

Wang & Kadia (2001) reported higher firing rate in response to conspecific vocalizations in marmoset primary auditory cortex compared to the same vocalization played in reverse, but observed no difference in firing rate when they played the same vocalization forward and backward to another

species—the cat. Such species-specificity may be innate or dependent on prior experience. It cannot be explained solely by acoustic asymmetry of the vocalizations, and likely requires behavioral significance being associated with the vocal sounds (Seyfarth & Cheney 2010; Syka et al. 2005).

Animals in our experiments were not exposed to courtship vocalizations used as experimental stimuli. Although they had experienced pup calls their siblings made during early development, those calls are characteristically different from adult calls Kim & Bao (2009); Liu et al. (2003). Thus the results reported here are likely innate and independent of experience. Others have shown that experience to early experience to heterospecific vocalizations may lead to better more efficient encoding of these sounds without necessarily conferring specificity for natural over reversed vocalizations (Schnupp et al. 2006; Woolley et al. 2010). Still, it remains to be determined how experience could change the cortical representation of the conspecific vocalizations played here.

2.5.5 Anesthesia effects

Because brain states and behavioral/emotional relevance of the stimuli may play a role in modulating brain activity, anesthesia, which removes/alters those influences, will affect how brain process vocalizations (Elhilali et al. 2002; Schumacher et al. 2011; Syka et al. 2005). For example, ketamine/xylazine anesthesia has been shown to change the trial-to-trial variability of neuronal responses to sound in rat auditory cortex, where response variability was lowest under light anesthesia and highest under a medium level of anesthesia (Kisley & Gerstein 1999). In songbirds, Schumacher et al. (2011) found that urethane anesthesia did not change spectral tuning or single neuron discriminability, but that it did significantly depress spike rates to conspecific song as well as spontaneous firing rates. In guinea pigs, researchers have shown depressed and sometimes temporally broader or more poorly phase-locked responses to conspecific vocalizations under ketamine/xylazine anesthesia compared to awake animals (Syka et al. 2005). Comparing STRFs calculated from TORCs in anesthetized and awake ferrets, Elhilali et al. (2002) report similar varieties of STRFs in the two cases, but increased spike rates, less separable STRFs, more complex spectral processing, and increased direction selectivity in the awake case. The lack of difference between responses to forward and backward courtship calls in the present study may be due to the anesthesia. However, similar cortical responses to forward and backward conspecific vocalizations were also observed in awake squirrel and macaque monkeys (Glass & Wollberg 1983; Recanzone 2008).

2.5.6 Summary

Conspecific vocalizations belong to a class of complex and highly structured acoustic stimuli. Efficient representation of this type of sound is important for the survival of the organism and species. Elucidating how conspecific sounds are represented in auditory cortex will likely advance our understanding on sensory coding mechanisms. Results of the present study, together with those in previous reports, indicate that auditory cortex processes natural vocalizations in ways that are not predicted from responses to simple synthetic sounds. Further studies are needed to reveal how the unique sensory processing of conspecific vocalizations develops through genetically determined or experience-dependent processes.

Acknowledgments

This work was performed in collaboration with Matar Haller, Claire Oldfield, and Frédéric Theunissen. We would also like to thank Dr. Timothy Holy and Zhongsheng Guo for providing the stimuli used in this project. The authors also would like to thank Alexander Huth for helpful discussion and expertise on STRF estimation techniques.

CHAPTER 3

Early and late critical period sound exposure differentially affect functional connectivity in auditory cortical circuits

3.1 Abstract

The neural representation of sound is shaped by acoustic experience during sensitive periods of sensory development. The cortical circuits mediating functional plasticity in these sensitive periods are not well characterized. We acquired simultaneous multi-unit recordings in putative layers 2–4 of the mouse primary auditory cortex (AI), anterior auditory field (AAF), and secondary auditory cortex (AII) after pulsed noise exposure during different sensitive windows of the auditory critical period. We then used an Ising model to describe functional connectivity within auditory cortical circuits. Noise exposure in early windows reduced stimulus-related functional connections in all auditory fields, with slight differences between AI, AAF, and AII. Late noise exposure, in contrast, was associated with increased intrinsic firing rate and reduced spread of functional connectivity between deep and superficial layer sites in AI. We compared these functional connectivity results with receptive field response properties to determine how single-site feature selectivity is related to circuit properties. Our results suggest that changes to distinct auditory circuits underlie plastic changes observed at the single site level.

3.2 Introduction

The auditory cortex undergoes rapid maturation during early postnatal development, as manifested by the emergence and refinement of cortical sound representations (Zhang et al. 2002; Chang et al. 2005; de Villers-Sidani et al. 2007; Insanally et al. 2009). This process is shaped by acoustic experience in a "critical period" of heightened plasticity (Hensch 2005; de Villers-Sidani et al. 2007). Within the critical period, sensory experience in a series of time windows changes the representation of different acoustic features, including the characteristic frequency (CF), tuning bandwidth

(BW), frequency modulation (FM) and binaural selectivity (Insanally et al. 2009, 2010; Popescu & Polley 2010). For example, early critical period experience shapes the cortical frequency map (de Villers-Sidani et al. 2007; Insanally et al. 2009; Liu et al. 2011), whereas later critical period experience shapes tuning bandwidth and frequency modulation selectivity (Insanally et al. 2009, 2010).

The existence of multiple sensitive periods for different acoustic features suggests that the CF, BW, FM, and binaural selectivity of cortical neurons are determined by different circuits that undergo experience-dependent modifications in distinct time windows. However, the relative contributions of thalamocortical and corticocortical circuits in mediating changes in these sensitive periods are still poorly understood. For example, some studies suggest that CF is determined by thalamocortical connections and bandwidth is modulated by corticocortical connections (Kaur et al. 2004; Happel et al. 2010; Barkat et al. 2011; Li et al. 2013), but others have posited that corticocortical connections play a role in CF determination (Liu et al. 2007). In general, it is thought that thalamocortical connections in sensory systems are more likely altered during early development and that adult plasticity occurs primarily in corticocortical circuits (Feldman et al. 1999; Fox 2002), though recent studies have shown that thalamocortical plasticity is possible in adulthood (Oberlaender et al. 2012; Yu et al. 2012).

To investigate differential changes in functional connectivity in auditory circuits, we use the Ising model, a tool common in physics (Ising 1925) and machine learning (Ackley et al. 1985) that describes a network in terms of pairwise couplings between nodes. When applied to a network of neurons, the Ising model assigns couplings to pairs of neurons based on observed population firing patterns. Although the Ising model consists of only pairwise interactions or couplings, these couplings are determined while taking into account all other activity in the network, and are thus distinct from pairwise correlations. Specifically, it is possible to disambiguate the relationships between pairs of neurons that appear correlated as a result of common input or serial connections (Schneidman et al. 2006; Hamilton et al. 2013).

This method has been used to model interactions between ganglion cells in the retina (Schneidman et al. 2006; Shlens et al. 2006, 2009; Ganmor et al. 2011b,a), cultured cortical cells (Schneidman et al. 2006), simulated cortical neurons (Roudi et al. 2009a; Schaub & Schultz 2012), and cortical neurons *in vitro* (Tang et al. 2008) and *in vivo* (Marre et al. 2009; Ohiorhenuan et al. 2010; Hamilton et al. 2013). In our previous work, we showed how the Ising model can uncover dynamic changes in neuronal selectivity as well as functional connectivity between neurons (Hamilton et al. 2013). Here, we use it to describe network-level changes in thalamocortical and corticocortical circuits as a result of plasticity.

In the present study, we recorded neuronal activity at multiple sites across layers and columns of the auditory cortical network. We found that early exposure to pulsed noise altered functional connectivity in thalamocortical circuits, whereas late exposure preferentially altered corticocortical circuitry. These findings indicate that different cortical circuits undergo experience-dependent reorganization in distinct sensitive periods.

3.3 Experimental procedures

3.3.1 Data acquisition

Subjects

The University of California-Berkeley Animal Care and Use Committee approved all procedures. Litters of mouse pups (strain B6;129P2-Pvalb$^{tm1(cre)Arbr/J}$, the Jackson Laboratory) and their mothers were placed in a sound-attenuation chamber (MedAssociates, Inc.), given food and water ad libitum, and kept on a 12 hour light/dark cycle. Groups were housed with 5-10 animals per cage pre-weaning (at P21), and split into groups of 6 or fewer post-weaning. Animals were exposed to trains of pulsed noise for one of three time windows (window 1, P8–15 ($n = 6$, 5F/1M); window 2, P16–23 ($n = 5$, 4F/1M); and window 3, P24–P31 ($n = 4$, 3F/1M). White noise bursts (50 ms long, 5 ms squared cosine ramp) were generated with a LabView program and played through a free-field speaker. The peak sound pressure level was calibrated and set at 65 dB SPL (Bruel and Kjaer microphone and conditioning amplifier). Each noise burst train consisted of 6 noise bursts played at 6 Hz, separated by 500 ms of silence. Following the exposure period, animals were returned to regular animal husbandry rooms. Control animals ($n = 5$, 3F/2M) were housed in the regular husbandry room prior to mappings and were not exposed to noise.

Electrophysiological recording and stimuli

The right auditory cortex was mapped between P40-60 for each mouse under anesthesia using a cocktail of ketamine (100 mg/kg) and xylazine (10 mg/kg) and procedures described previously (Han et al. 2007). Animals in different groups were age-matched. Following deflection of the temporal muscle, exposure of the auditory cortex, and removal of the dura mater, we performed a coarse mapping with tungsten electrodes (FHC, Inc.) to determine the location of primary auditory cortex (AI) based on rostrocaudal tonotopy and short spike latencies. We then recorded extracellular multi-unit neural activity in putative layer 2/3 through layer 4 of the right primary auditory cortex (AI), anterior auditory field (AAF), and secondary auditory cortex (AII) (Oviedo et al. 2010) using a 4×4 silicone polytrode (NeuroNexus A4$\times$4–3mm–100–125–177). A total of 1,008 AI (266 control, 308 W1, 182 W2, 252 W3), 658 AII (182 control, 238 W1, 112 W2, 126 W3), and 854 AAF (252 control, 266 W1, 140 W2, 196 W3) multi-unit sites were sampled. The polytrode was oriented parallel to the tonotopic axis and lowered orthogonally to the cortex such that the deepest contact sites were at a depth of approximately 500 microns from the pial surface (Fig. 3.3, bottom right). The extracellular signal was obtained using a TDT amplifier connected to TDT RX5 hardware (Tucker Davis Technologies, Inc.). Spike times were calculated by thresholding the extracellular signal at 1.5 times the standard deviation of the signal-to-noise ratio, and were logged using custom MATLAB and TDT RPvdsEx software running on a Windows XP computer. Sound stimuli were presented to the contralateral ear through a cannulated speaker (EC-1, Tucker Davis Technologies, Inc.) controlled by TDT RX6 hardware, and calibrated to ensure a flat output from 4–75 kHz with less than 3% spectral distortion. Sound stimuli were pure tones (25 ms length with 5 ms squared cosine ramp) played from 4 kHz to 75 kHz in 0.2 octave steps, for a total of 22 frequencies, and white noise bursts (25 ms

length with 5 ms squared cosine ramp). Sounds were presented at 6 different loudness levels (20–70 dB SPL, 10 dB spacing) in a pseudorandom order with a 2 Hz repetition rate. For the three loudest intensities (used as input for our later model fitting), each frequency-intensity pair was repeated 15 times, for other intensities, stimuli were repeated 3 times. For each 500-ms trial, a tone pip would play at 250 ms into the trial.

3.3.2 Ising model fitting

We fit Ising models using custom MATLAB code running on 6-core linux machine (AMD Phenom II X6 1090T Processor) with 16 GB of RAM. The code is available online (`https://github.com/libertyh/ising-model/`) with sample data from a control subject in this study.

To investigate functional connectivity in the auditory cortex, we fit Ising models with three different patterns of connectivity (see Fig. 3.1A for a schematic): (1) the independent neurons model, in which recording sites were coupled to the sound stimulus condition, but not coupled to other recording sites. In this case, the predicted activity at recording sites was conditionally independent given the stimulus. (2) The no sound model, where only couplings between recording sites were considered, but there was no coupling to the sound condition. (3) The fully connected model, where the sound condition as well as couplings between recording sites were considered.

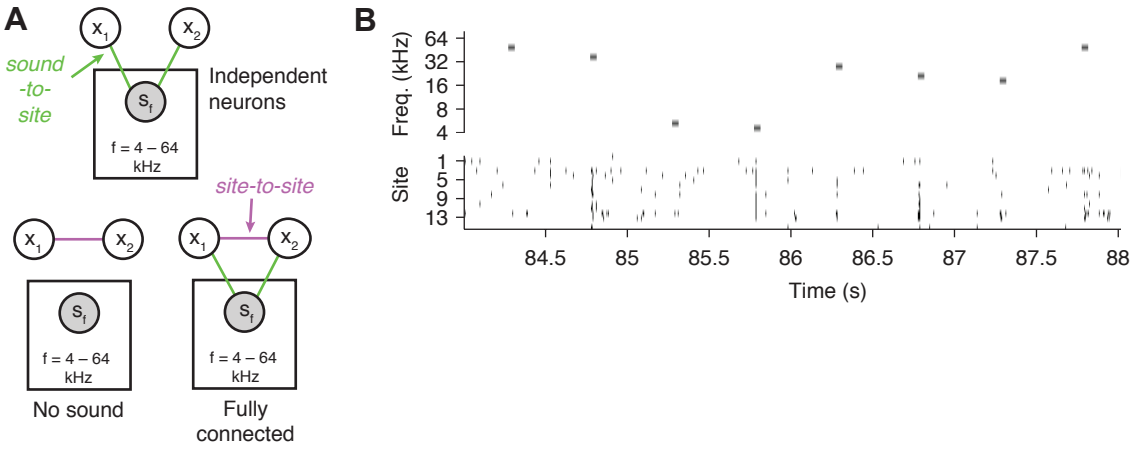

Figure 3.1: Ising model fitting for different patterns of functional connectivity. **(A)** Schematic showing simplified depiction of the three types of connectivity used for Ising model fitting. In the independent neurons model, only couplings between the sound stimulus and the recording sites were considered. In the no sound model, only couplings between recording sites were considered. In the fully connected model, all possible pairwise connections between recording sites and sound stimuli were incorporated into the model. **(B)** Input to the Ising model consisted of binary matrices representing the stimulus condition and spikes recorded for each 5 ms time bin. Pure tones from 4–75 kHz (gray rectangles) were played in pseudorandom order at a 2 Hz repetition rate. The state of each recording site over time (spike = 1, no spike = 0) is shown on the bottom half of the plot.

Matrices of sound stimulus conditions (frequency presented) and spiking data were used as input to the model (Fig. 3.1B). Spiking activity from each of the contact sites on the polytrode was

binned into 5 ms time epochs. If a spike occurred during those 5 ms, the value was set to 1, and if no spikes occurred, the value was set to 0. The sound stimulus matrix consisted of values for 23 sound frequencies (22 pure tones + 1 white noise condition) for each of t time points, with values of 1 when a stimulus at a given frequency was "on" and 0 when the stimulus was off. We set the sound matrix to 1 beginning at 15 ms after the onset of the sound and ending at 50 ms after sound onset to account for cortical response latencies. Only responses to the three highest decibel levels were used in the model. We used 10-fold cross validation to obtain robust model estimates (Kohavi 1995). For each polytrode location, the 500-ms trials were randomized, the full data matrix was split into 10 equal chunks, and each model was estimated by holding out 1 of the data chunks and training on the remaining 90% of the data.

The modified Ising model is described by the following equation:

$$p(\mathbf{x}|\mathbf{s};\mathbf{J},\mathbf{W}) \;=\; \frac{1}{\mathbf{Z}(\mathbf{s},\mathbf{J},\mathbf{W})}\exp\left(\mathbf{x}^{\mathrm{T}}\mathbf{J}\mathbf{x}+\mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{s}\right) \tag{3.1}$$

where **J** represents coupling in activity of two sites (for any sites x_i and x_j, $J_{ij} > 0$ indicates that site x_i and x_j tend to be active simultaneously, $J_{ij} < 0$ indicates that when x_i spikes, x_j remains silent and vice versa). **W** represents the sound-to-site coupling matrix, where positive values indicate that spiking increases during stimulus presentation, and negative values indicate a suppression of spiking during the stimulus.

We estimated the couplings using Minimum Probability Flow (Sohl-Dickstein et al. 2011), which has the following objective function:

$$\begin{aligned} K(\mathbf{J},\mathbf{W}) \;=\; & \frac{1}{T}\sum_{\mathbf{x}\in D}\sum_{\mathbf{x}'\notin D} g(\mathbf{x},\mathbf{x}')\exp\left(\frac{1}{2}\left[E(\mathbf{x}|\mathbf{s};\mathbf{J},\mathbf{W})-E(\mathbf{x}'|\mathbf{s};\mathbf{J},\mathbf{W})\right]\right) \\ & +\lambda\left(\sum|W|+\sum|J|\right) \end{aligned} \tag{3.2}$$

where $E(\mathbf{x}|\mathbf{s};\mathbf{J},\mathbf{W}) = -\mathbf{x}^{\mathrm{T}}\mathbf{J}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{s}$ is the energy function for the Ising model, $K(\mathbf{J},\mathbf{W})$ is the MPF objective function, and the MPF connectivity function $g(\mathbf{x},\mathbf{x}')$ was set so as to connect states which differed by a single bit flip. The value T is the number of time points used to fit the model.

3.3.3 Improving Ising model fits for neural data

One potential problem in using MPF to fit an Ising model to neural data is that spike data is typically sparse, and the majority of the training data therefore consists of zero (no spikes) or only a small number of neurons firing simultaneously (Olshausen & Field 2004; Hromadka et al. 2008). MPF functions by evaluating the relative probabilities of training data states and non-training comparison states which are in the neighborhood of the training states. MPF can have difficulty assigning correct relative probabilities to states that are far away from the training data and its neighborhood. Typically for fitting an Ising model, the neighborhood of the training data is set to be all states that differ from the training data by a single bit flip. In the case of sparse training data, states in which nearly all neurons fire simultaneously are on the opposite side of the state space and thus

very far away from the training data and the neighborhood of the training data. As a result, MPF was observed to assign erroneously high probabilities to states with nearly all neurons firing simultaneously. We propose an extension to MPF, which we refer to as MPF + bitflip, in which the neighborhood of the data states is extended to include all states which differ from the training data by a single bit flip, and also all states which differ from the training data in all bits.

Including this enhancement improves the model likelihoods for the fully connected model and for the no sound model, but worsens performance on the independent neurons model (Fig. 3.2).

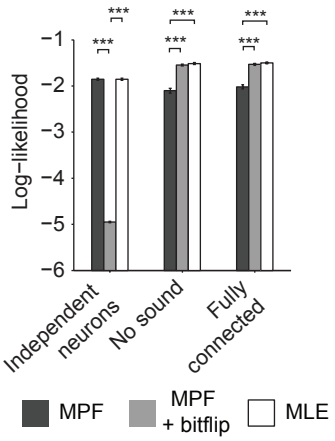

Figure 3.2: Model performance for three types of connectivity: the independent neurons model, the no sound model, and the fully connected model. Model performance was assessed by computing the log-likelihood of the Ising model couplings on a held-out dataset not used for training. We used three methods to fit couplings: Minimum Probability Flow learning (MPF), Minimum Probability Flow + bitflip (MPF + bitflip), which better fits sparsely active neural data, and Maximum Likelihood Estimation (MLE), which was used to assess "ground truth". For the independent neurons model, the original MPF solution is better than MPF + bitflip and is not significantly different from the MLE solution. For the no sound and fully connected models, MPF + bitflip couplings are significantly more likely than MPF couplings and not significantly different from the MLE solution, suggesting a good fit to the data. Error bars represent standard error of the mean (SEM). Significance was assessed using Wilcoxon rank sum tests, and p-values were Bonferroni corrected for multiple comparisons. Asterisks indicate level of significance: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

For the independent neurons model, the original MPF solution is better than the MPF + bitflip solution (Bonferroni-corrected $p = 1.3 \times 10^{-16}$, Wilcoxon rank sum test), and is not significantly different from the MLE solution ($p > 0.5$, Wilcoxon rank sum test). The poor performance of the MPF + bitflip model here arises from the inability to simultaneously assign the correct probabilities to the rare densely active states (where all neurons are firing spontaneously) and the sparsely active states. The extension to MPF helped to assign the correct probabilities to the rare densely active states at the expense of fitting the correct probabilities to the more common sparse states. Since the only terms characterizing neuronal firing in the independent neurons model are the bias terms, which describe the overall firing rate of units in each site, adding the bitflip does little to improve the model fit. For the no sound and the fully connected models, however, the likelihood of the MPF +

bitflip solution is significantly higher than that of the original MPF algorithm (Bonferroni-corrected $p = 3.2 \times 10^{-4}$ and $p = 2.1 \times 10^{-5}$, respectively, Wilcoxon rank sum tests), and is not significantly different from MLE ($p > 0.5$, Wilcoxon rank sum tests). When analyzing subsequent couplings, we chose those from the best-performing MPF models—original MPF for the independent neurons couplings, and MPF + bitflip for the no sound and fully connected models.

Sometimes the weights fitted to the training data may not generalize well to held-out data—for example, when the neural response to a stimulus is very infrequent. In this case, if there are no spikes fired in response to a tone pip in the training data, but there are one or more spikes to the same tone pip in the test set, a very large negative coupling weight would be assigned to model fit on the training data, but this would not predict the responses seen in the test set. To prevent this overfitting to training data, we regularized the MPF solution using an L1-regularization parameter $\lambda = 5.9 \times 10^{-5}$, chosen using nested cross-validation as follows. Ten values logarithmically spaced between $\lambda = 10^{-7}$ and $\lambda = 10^{-2}$ were tested by holding out 20% of the training data and training the model using the remaining 80%, repeating this 5 times, and choosing the λ with the best average log-likelihood across all sites.

Following selection of the regularization parameter, we fit the model using all of the training data, and the model log-likelihood was tested on the held out validation set. This was repeated 10 times for different validation sets, using the same regularization parameter. Coupling matrices shown in the figures are taken from the cross-validation iteration with the highest likelihood in the validation set. We evaluated model likelihoods on held out data by calculating the partition function directly for all possible spiking states.

$$\log L(\mathbf{J}|\mathbf{x}_v) = \sum_s \log \left(\frac{1}{Z} \exp \left[-(\mathbf{x}^\mathbf{T} \mathbf{J}_\mathbf{x} \mathbf{x} + \mathbf{s}^\mathbf{T} \mathbf{J}_\mathbf{s} \mathbf{x}) \right] \right) \tag{3.3}$$

3.3.4 Manipulating Ising model couplings

To test the relationship between couplings in the Ising model and receptive field parameters such as bandwidth and response magnitude, we manipulated couplings in the Ising model and generated new samples of neural activity using Gibbs sampling. The site-to-site couplings for one control site were used as the original input, and sound-to-site couplings were generated from a gaussian distribution plus noise, so that the deepest sites responded most strongly and the superficial sites (row 1) were negatively coupled to the sound. The coupling terms were then changed in the following ways. We manipulated the bias term (the diagonal elements of the **J** matrix) by adding randomly generated numbers such that the overall increase in bias was +0.61, +0.57, +0.24, or -0.05. We also manipulated the spread index by increasing or decreasing site-to-site couplings between the deep and superficial row sites such that the change in spread index was +0.05, -0.02, -0.06, or -0.10. We then generated new tuning curves by counting the number of spikes elicited at each site by a simulated sound stimulus. This was repeated 100 times (where different sets of samples were generated with Gibbs sampling (see (Sohl-Dickstein et al. 2011) for examples) to obtain error bars on the simulated tuning curves. We calculated the bandwidths of these tuning curves by fitting a gaussian

function, then calculating the frequency response width at 28% of the maximum response, which is the same threshold used to calculate bandwidths for the tone-derived receptive fields.

3.3.5 Calculation of receptive field response properties

We also analyzed basic response properties including CF distribution, tonotopy, absolute and relative bandwidth, intensity threshold, spontaneous firing rate in epochs before sound stimulation, and the incidence of distorted or multi-peak RFs. The CF was calculated as the frequency at which the maximum response was observed when collapsing across intensities. Tonotopy was calculated by performing a orthogonal least squares regression on the x coordinate (anterior-posterior axis of auditory cortex), y coordinate (dorsal-ventral axis), and CF parameters for AI and AAF separately and finding the best-fit plane. The CF predicted by the regression was plotted against the actual CF to obtain the plots shown in Fig. 3.10. Absolute frequency bandwidths were calculated after thresholding background activity at 28% of the maximum response, then calculating the width in octaves of the frequency response at each intensity level separately. Relative frequency bandwidths were calculated similarly, but starting at the threshold intensity, which was defined as the quietest intensity to elicit a response (the intensity at the tip of the frequency-intensity tuning curve). Spontaneous firing rate was calculated as the mean firing rate in the first 200 ms of the trial, before the sound stimulus onset at 250 ms.

3.3.6 Statistical tests

To test differences in coupling, we used nonparametric Kruskal-Wallis analysis of variance (ANOVA) followed by Wilcoxon rank sum tests (for comparing independent groups) or Wilcoxon signed rank tests (for comparing paired groups). Parametric tests were not used because it was determined that the couplings being tested were not Gaussian distributed (Lilliefors test).

3.4 Results

We fit Ising models for three different patterns of functional connectivity (see Methods). Depending on the connectivity modeled, the raw correlations between neural activity and external stimuli may be quite different from the couplings calculated in the Ising model. In the independent neurons model, the recovered sound coupling matrices were similar to the correlations estimated using traditional methods (Fig. 3.3, $\rho = 0.97$, $p = 0$, Spearman rank correlation).

Still, the independent neurons model had the lowest likelihood on the held-out dataset (Fig. 3.2), suggesting that it was a poor model of our population data. In the more anatomically plausible fully connected model, the indirect modulation of neural activity by sound stimuli can be "explained away" by connectivity within the neural network. In the fully connected model, couplings were distinct from the raw sound to site correlations (Fig. 3.3). These couplings were still significantly positively correlated with the raw correlations between spike activity and sound presentation ($\rho = 0.68$, $p = 0$, Spearman rank correlation), but the magnitude of the partial correlation was lower than the independent neurons model.

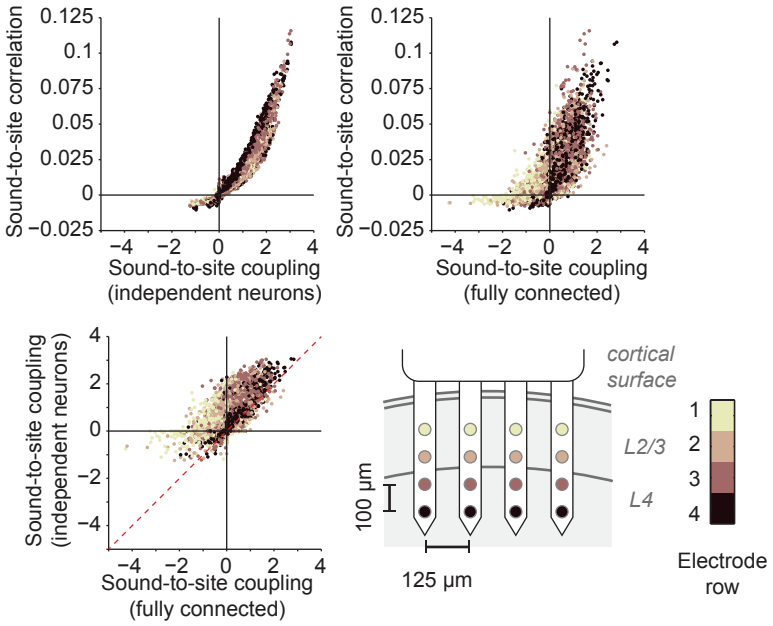

Figure 3.3: Ising model couplings for different patterns of functional connectivity. Top left: Sound-to-site correlation is highly correlated with couplings for the independent neurons model, where connectivity between recording sites is not considered ($\rho = 0.96$, $p = 0$, Spearman rank correlation). Top right: When connectivity between recording sites is taken into account in the fully connected model, sound-to-site couplings are more distinct from sound-to-site correlations ($\rho = 0.56$, $p = 9.1 \times 10^{-153}$, Spearman rank correlation). Bottom left: couplings for the independent neurons model versus the fully connected model. Bottom right: schematic of the 4×4 polytrode location. The polytrode was positioned such that the deepest row sites were approximately 500 µm from the pial surface. This meant that the bottom two rows were recorded from putative L4, and the top two rows were approximately localized to L2/3. The polytrode was oriented parallel to the tonotopic axis as determined by a coarse mapping with tungsten doublet electrodes.

3.4.1 Fully-connected Ising models recover the canonical cortical circuit

To elucidate relationships between anatomical organization and functional activity, we compared the fully connected model couplings for neural activity and sound in different layers or columns. We found that sound responses in the superficial layer were explained by connectivity to deeper layer input neurons (rows 3 and 4 on the electrode), putatively located in layer IV. In all sites tested, coupling between neural responses and sound increased with depth ($p = 4.0 \times 10^{-249}$ for main effect of depth, Kruskal-Wallis ANOVA). On average, the more superficial cortical sites showed negative coupling with the sound stimulus, particularly when the frequency of the sound played was within the receptive field (Fig. 3.4A). Pooling across all data, coupling was significantly different between all rows on the electrode, with deeper rows more strongly positively coupled with the sound stimulus (Fig. 3.4B, corrected $p < 0.001$ for all pairwise comparisons, Wilcoxon rank sum tests). This structure cannot be captured by individually correlating neural activity with the stimuli that elicit a response in that neuron (as in the independent neurons model) and is only evident when we take

into account the connectivity inherent in the network. Traditional methods of stimulus-response analysis (such as calculating classical receptive fields or spectrotemporal receptive fields for individual neurons) do not distinguish indirect responses to stimuli that are more influenced by reciprocal connectivity in the network from direct stimulus drive, especially when these measures are calculated for neurons in non-thalamorecipient layers. In addition, adding coupling terms between the sound and neural activity at each site does not significantly improve the model likelihood ($p > 0.5$, Wilcoxon rank sum test, and see Fig. 3.2), further underscoring the predominance of influences from neighboring neurons on neuronal activity over sensory input.

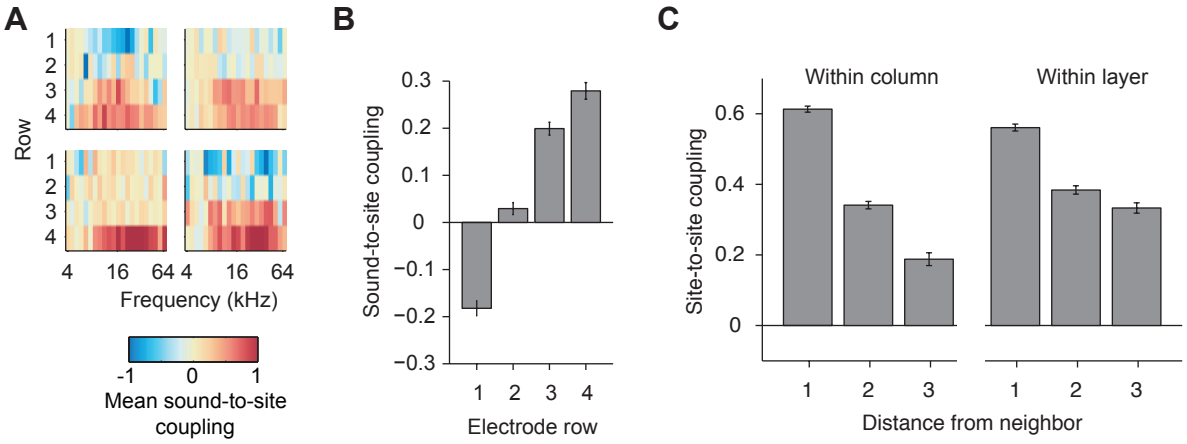

Figure 3.4: The Ising model reveals canonical functional connectivity in naive auditory cortex. **(A)** Site to sound coupling averaged across electrode sites within the same row for four representative naive control subjects in the fully connected model. Strong positive coupling was generally seen in deeper layers (rows 3–4) within frequencies corresponding to the receptive field, while some superficial layer sites showed negative coupling to the sound stimulus within the receptive field. **(B)** Average coupling between neural activity and sound within an electrode row in the fully connected model increases with depth. **(C)** Coupling between activity at pairs of recording sites at different distances from one another. Vertically within a column, site to site coupling was smaller between neurons at larger distances, and fell off more sharply than horizontal site to site coupling (within a layer), which remained strong across distances up to 3 nodes away (375 μm). In all plots, error bars indicate standard error of the mean.

Another important feature of applying this model to cortical circuits is that it can reveal the relative strength of laminar and columnar functional connectivity. By considering the spatial location of the recording sites with regard to their site-to-site coupling, we find that vertical coupling within a column in the auditory cortex falls off significantly with distance, whereas horizontal coupling within a layer is consistently strong across relatively large distances up to 375μm (3 nodes away, Fig. 3.4). In addition, coupling between distant sites (3 nodes away) in a layer was significantly greater than coupling between distant sites in a column (Bonferroni-corrected $p < 0.001$, Wilcoxon rank sum test), even though the vertical spacing between contacts on the electrode was smaller (100 μm vertically compared to 125 μm horizontally). Within a layer, coupling was more similar between sites 2 or 3 contact sites away than at the same separation within a column. It remains to be determined whether this strong horizontal connectivity is universal or is unique to auditory cortex (Sharma et al. 2000; Oviedo et al. 2010).

3.4.2 Early noise rearing reduces thalamocortical functional connectivity

The ability of the Ising model to simultaneously fit functional connectivity between cortical sites and the stimuli that modulate them makes it ideal to investigate circuit-level changes following critical period sound exposure. Sound-to-site and site-to-site functional connectivity can be used to assess relative perturbations in thalamocortical and corticocortical circuits, respectively. Assessing plastic changes in functional connectivity in addition to selectivity changes in single recording sites can thus augment our understanding of the mechanisms influencing the multiple sensitive windows making up the auditory critical period (Insanally et al. 2009).

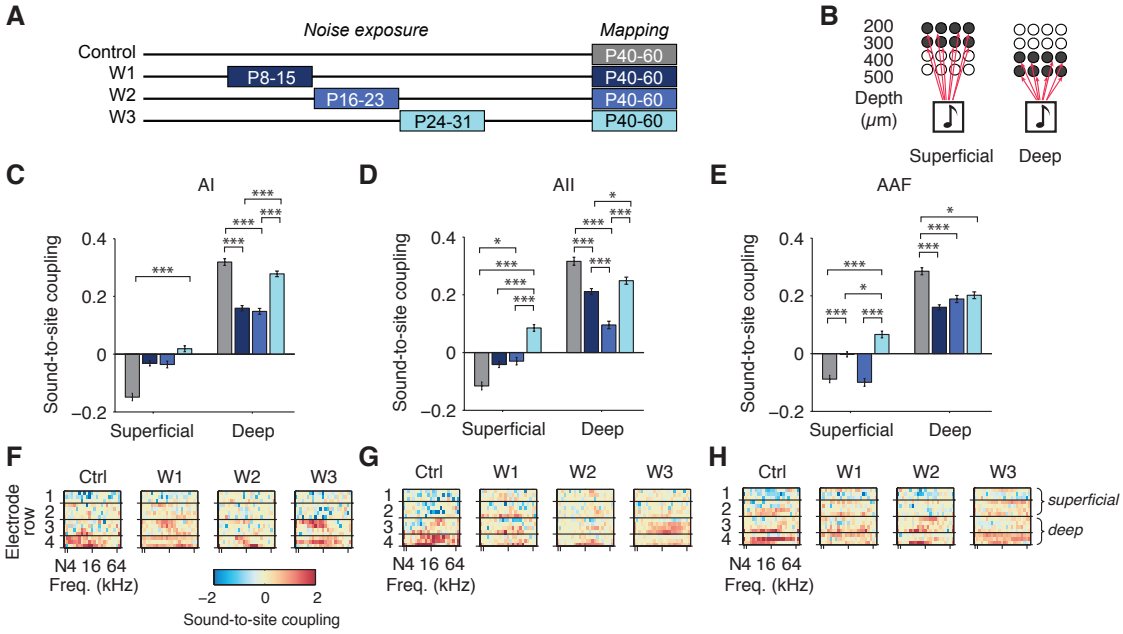

Figure 3.5: Effect of pulsed noise exposure on sound-to-site coupling in the auditory cortex. **(A)** Schematic of pulsed-noise rearing protocol. W1 animals were exposed to pulsed white noise 24h/d from P8–15, W2 from P16–23, and W3 from P24–31. Controls were reared in a normal acoustic environment. All animals were mapped between the ages of P40 and P60. **(B)** Sound-to-site coupling terms were calculated between all sound stimuli and electrode sites, and collapsed across superficial (row 1 and 2) and deep layer (row 3 and 4) sites for plotting in panes **C** through **E**. **(C-E)**. Summary data for sound-to-site coupling in controls, W1, W2, and W3 animals collapsed across superficial and deep row sites. **(C)** In AI, noise rearing during W3 significantly increases sound-to-site coupling in superficial sites. Noise rearing during W1 and W2 significantly reduces sound-to-site coupling in deep input layer sites. **(D)** In AII, noise rearing during W1, W2, and W3 increases superficial sound-to-site coupling. Significant reductions are seen in deep layer sound-to-site coupling for W1 and W2 only. **(E)** In AAF, noise rearing during W1 and W3 increases sound-to-site coupling in superficial layer sites. W1 and W2 reared animals showed significant reductions in sound-to-site coupling in deep layers. **(F-H)** Example sound-to-site coupling in AI (f), AII (g), and AAF (h) for controls, W1, W2, and W3. In general, deep layer coupling (row 3 and 4) was strongest in control and W3 animals compared to W1 and W2, which showed overall weaker amplitudes of sound-to-site coupling. Asterisks denote level of significance as assessed by Wilcoxon rank sum tests with Bonferroni correction for multiple comparisons (* = corrected $p < 0.05$; ** = corrected $p < 0.01$; *** = corrected $p < 0.001$, number of comparisons = 12).

Sound-to-site coupling can be used as a proxy for thalamocortical input to deep layers, and in superficial layers can be interpreted as additional modulations by sound that are not explained by coupling between cortical sites. To describe changes in thalamocortical functional connectivity after rearing in pulsed white noise, we analyzed the Ising model couplings between the sound stimulus and nodes in either the superficial (rows 1 and 2) or deep (rows 3 and 4) layers (Fig. 3.5B). We found a significant decrease (Bonferroni-corrected $p < 0.01$, Wilcoxon rank sum tests) in site-to-sound coupling in AI, AII, and AAF after noise exposure in early windows W1 and W2 (Fig. 3.5C).

In "primary" areas AI and AAF (Winer & Schreiner 2011), the profile of changes was relatively similar. In AII, however, the reduction in deep row sound-to-site coupling was strongest for W2 reared animals. Sound-to-site couplings in superficial row sites were significantly less negative than controls for all rearing windows, which may represent reduced inhibitory influences. These results suggest an early sensitive period for reductions in thalamocortical input to primary auditory cortex that is consistent with previous results (Chang & Merzenich 2003; Speechley et al. 2007; Insanally et al. 2009, 2010).

3.4.3 Reduced spread of functional connectivity in corticocortical circuits

Site-to-site coupling is a measure of functional connectivity between multi-unit recording sites, and in our experiments reflects corticocortical connectivity within and across L2/3 to L4 of the auditory cortex. In the Ising model, we fit site-to-site couplings J_{ij} for each pair of sites x_i and x_j. In our formulation, the diagonal of this matrix is the bias term, which represents the intrinsic firing of each site (smaller numbers indicate lower firing rates). Following noise rearing in W2 and W3, we observed increases in the bias term in AI, indicating higher intrinsic firing rate (includes both evoked and spontaneous, Fig. 3.6A). In AII and AAF, these increases were observed for late reared W3 only (Fig. 3.6B-C).

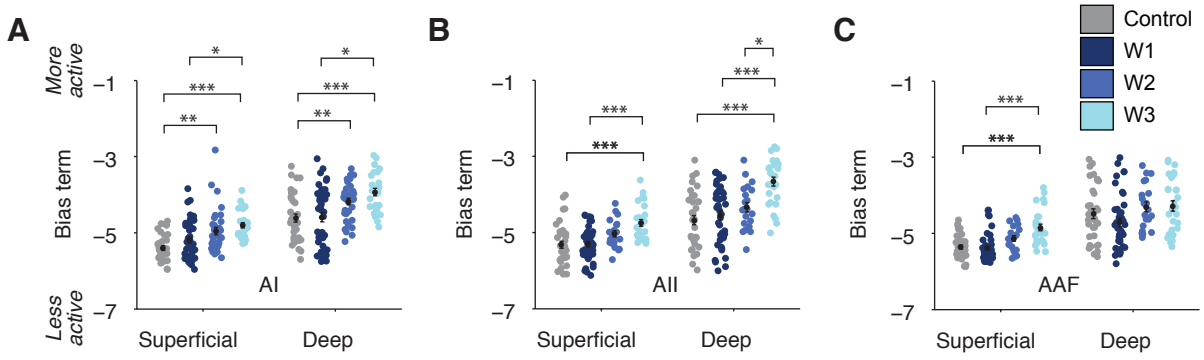

Figure 3.6: Pulsed noise rearing increases baseline activity. (**A-C**) Ising model bias term in superficial (row 1 and 2) and deep (row 3 and 4) sites for controls, W1, W2, and W3 reared animals. The more negative the bias term, the more likely a site is to be inactive (not spike). Bias terms closer to zero indicate sites with higher baseline activity. (**A**) Pulsed noise rearing in W2 and W3 increases bias in both superficial and deep sites in AI. (**B**) Pulsed noise rearing in W3 increases bias in superficial and deep layer sites in AII. (**C**) Rearing in W3 increases bias in superficial AAF sites, but no change is seen in baseline activity of deep layer sites.

In addition to increases in intrinsic activity, the W3 group exhibited decreases in the spread of site-to-site functional connectivity from deep to superficial layers. We calculated the spread as

$$\text{spread index} = 1 - \frac{\sum J_{(i,j)_v}}{\sum J_{(i,j)_v} + \sum J_{(i,j)_d}} \tag{3.4}$$

Where $J_{(i,j)_v}$ represents direct vertical coupling from deep to superficial layer sites and $J(i,j)_d$ represents off-column diagonal coupling (see Fig. 3.7A for a schematic). This measure describes how information is shared across areas of the auditory cortex that may be tuned to different frequencies. In AI, the spread index was decreased in W3 compared to controls (Fig. 3.7B, $p < 0.05$ Wilcoxon rank sum test) but did not significantly differ between any other groups. Furthermore, these changes were limited to AI and were not observed in AII or AAF (Fig. 3.7C-D).

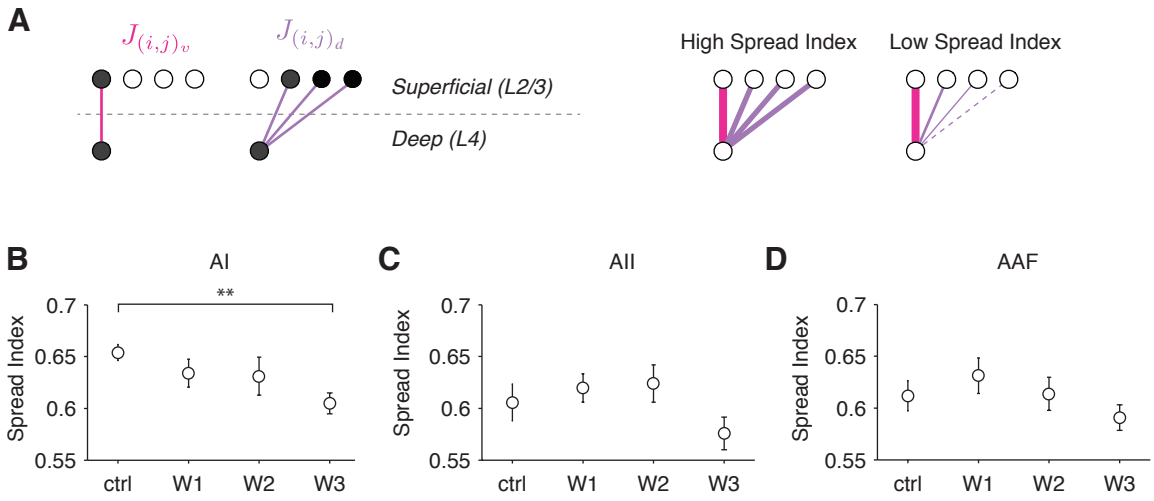

Figure 3.7: Late noise exposure decreases the spread of corticocortical connections from deep to superficial layers. **(A)** Schematic of site-to-site couplings contributing to the "Spread Index". Only site-to-site couplings from contacts from putative L4 (rows 3 and 4) to L2/3 (rows 1 and 2) are used. The Spread Index was calculated as 1 minus the ratio of the mean vertical couplings (magenta) to the mean of vertical and horizontally projecting (purple) couplings. A value of 0 indicates no functional intra-column spread of connectivity; values closer to 1 indicate strong intracolumnar spread of functional connectivity. Four sites are shown for simplicity, but in spread index calculations all possible couplings between L4 and L2/3 sites were used. **(B-D)**. Mean Spread Index in AI, AII, and AAF for different rearing groups as a function of horizontal distance from a node (see schematic below x-axis for example couplings in one shank that are 0, 1, 2, or 3 nodes away horizontally). **(B)** AI functional spread is significantly reduced in W3 animals compared to controls. **(C)** AII Spread Index did not differ across rearing conditions. **(D)** AAF Spread Index did not differ across rearing conditions. Asterisks denote level of significance as assessed by Wilcoxon rank sum tests with Dunn-Šidák correction for multiple comparisons (* = corrected $p < 0.05$; ** = corrected $p < 0.01$; *** = corrected $p < 0.001$). Error bars indicate standard error of the mean.

There were no significant differences across groups for vertical couplings (AI: $p = 0.9$ control vs. W1, $p = 0.08$ control vs. W2, $p = 0.17$ control vs. W3; AII: $p = 0.58$ control vs. W1, $p = 0.4$ control vs. W2, $p = 0.01$ control vs. W3; AAF: $p = 0.80$ control vs. W1, $p = 0.29$

control vs. W2, $p = 0.79$ control vs. W3, uncorrected p-values from Wilcoxon rank sum test, data presented as mean $\pm$ SEM), so this effect was likely driven by changes to diagonal couplings across shanks (J_d, purple couplings in Fig. 3.7A). Furthermore, no changes were seen in couplings within the same layer (Fig. 3.8), suggesting that diagonal projection patterns were differentially affected. Overall, our findings suggest that noise rearing in early windows of the critical period results in changes to thalamocortical inputs, while rearing in later windows results in changes to corticocortical functional connectivity.

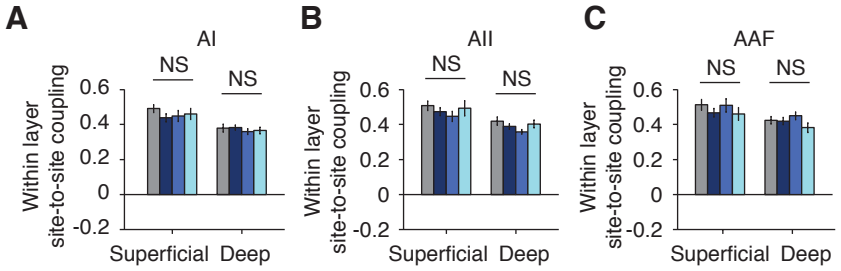

Figure 3.8: No changes to site-to-site coupling within superficial or deep layers were seen in any of the rearing groups or in any auditory areas. (Bonferroni-corrected $p > 0.05$, Wilcoxon rank sum tests).

3.4.4 Changes in response properties after noise exposure

To relate functional coupling in the Ising model to receptive field properties, we analyzed tonotopy, CF distribution, intensity threshold, absolute and threshold-relative frequency bandwidth, spontaneous firing rate, and the incidence of multi-peaked or distorted RFs. The basic gradient of tonotopy from low to high frequency in AI and high to low frequency in AAF caudorostrally was present for all rearing windows (Fig. 3.9A-B), with no significant differences in CF distribution across groups (paired Kolmorogov-Smirnov tests for each window, $p > 0.05$). However, other receptive field parameters varied across groups in a window dependent manner.

Early noise exposure was generally associated with changes in map organization and receptive field quality. In W1, tonotopy was significantly disrupted in AAF compared to all other groups (Fig. 3.10, $p = 0.009$ Wilcoxon rank sum test, significantly greater residual error to tonotopic axis compared to other groups).

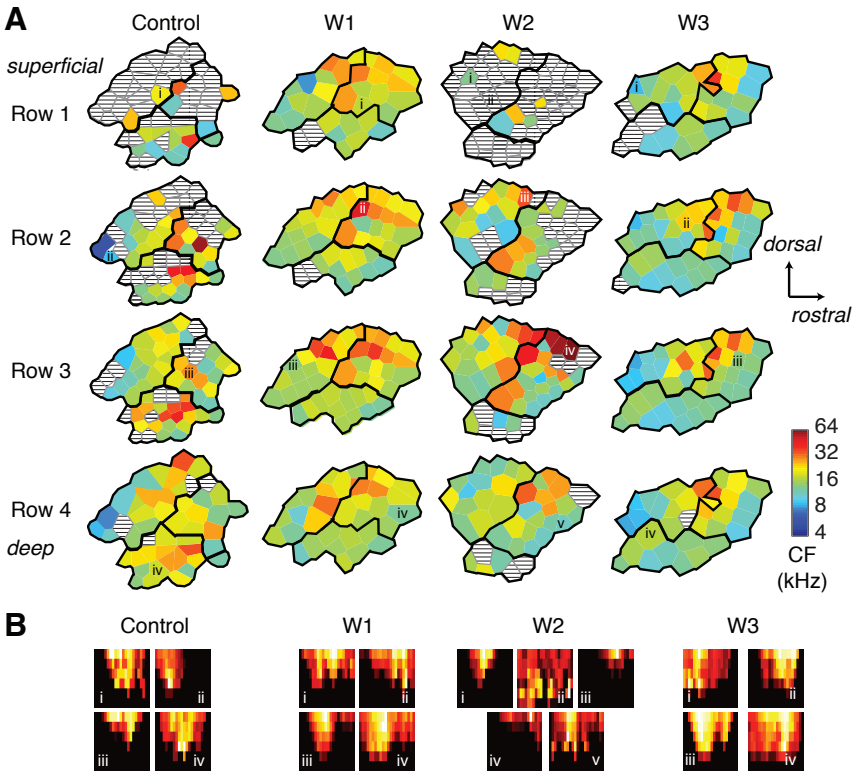

Figure 3.9: Cortical frequency selectivity maps after noise exposure in different sensitive windows. (**A**) Representative characteristic frequency (CF) maps for controls, W1, W2, and W3 noise-reared animals, split up according to row on the polytrode. Dark black outlines indicate boundaries between sites labeled as AI (more caudal), AAF (most rostral), and AII (most ventral). Hatched lines indicate sites that did not show strongly tuned responses. Responses were most robust in rows 3 and 4, corresponding to the putative thalamic input layers. (**B**) Example receptive fields from maps shown in (**A**).

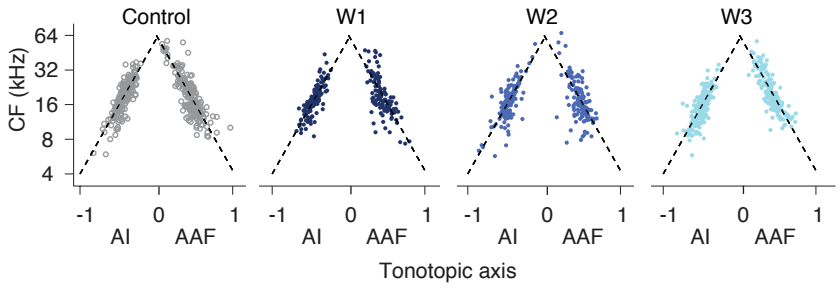

Figure 3.10: Changes in tonotopy after noise exposure in different sensitive windows. CF as a function of position on the tonotopic axis, collapsed across all subjects. Noise exposure in W1 was associated with disruptions in AAF tonotopy, as calculated by significantly larger residuals to dotted tonotopic axis line.

This disruption in tonotopy was accompanied by distorted tuning in AI. W1 subjects showed a higher incidence of multi-peaked RFs in AI compared to all other groups (Fig. 3.11A, $p < 0.01$, Wilcoxon rank sum test), indicating a less selective frequency representation. Example receptive fields and normalized tuning curves for multi-peaked (i-iii) and single-peaked (iv-v) sites from W1 subjects are shown in Fig. 3.11D.

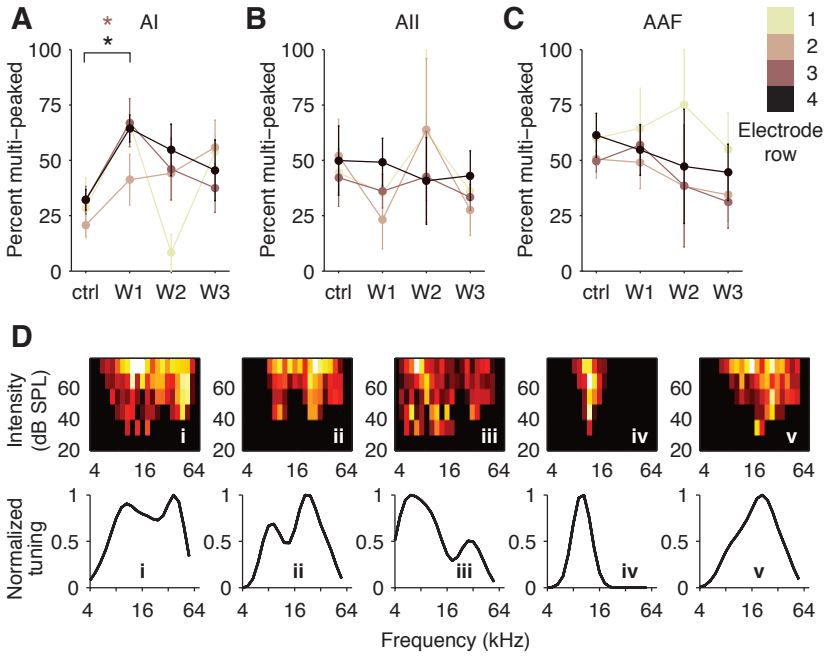

Figure 3.11: Early pulsed noise rearing in W1 (P8–15) increases incidence of multi-peaked tuning curves in AI. (A) Percentage of multi-peaked tuning curves in each rearing window for each row on the polytrode. In putative thalamorecipient input sites (rows 3 and 4), W1 animals showed a larger percentage of multi-peaked RFs compared to controls (* $p < 0.05$, Wilcoxon signed-rank test). **(B-C)**. Same plot as **(A)** but for AII and AAF. No significant changes in the incidence of multi-peaked receptive fields were seen across groups. **(D)** Example multi-peaked (i-iii) and single-peaked (iv-v) receptive fields from W1 group, AI sites.

In W2, response thresholds were significantly higher in primary auditory fields AI and AAF, indicating lower sensitivity to sound stimuli (Fig. 3.12A, 3.12C, Kruskal-Wallis ANOVA followed by post-hoc Wilcoxon rank sum test, $p < 0.05$). These increases in threshold were mainly seen in the putative thalamorecipient layers (row 3 and 4 sites on the polytrode). In AII, a non-primary area, there were no significant differences across rearing groups (Fig. 3.12B).

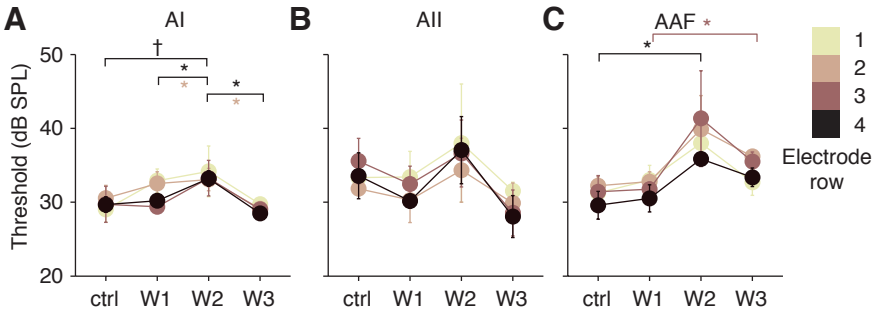

Figure 3.12: Pulsed noise in W2 (P16–23) increases receptive field thresholds in AI and AAF. Data are shown as mean ± SEM. **(A)** Thresholds in AI for each rearing group (ctrl = control group). Pulsed noise rearing increases receptive field thresholds in the putative thalamorecipient layers (row 3 and 4). **(B)** No change in threshold was seen for receptive fields in AII. **(C)** Thresholds in AAF increased in the W2 noise reared group compared to controls in row 4 sites. In row 4 sites, thresholds were significantly higher in W3 compared to W1. (†: $p < 0.1$, *: $p < 0.05$ corrected, Wilcoxon rank sum test).

These changes in threshold are particularly striking in the average receptive field across all subjects in each of the noise rearing windows (Fig. 3.13).

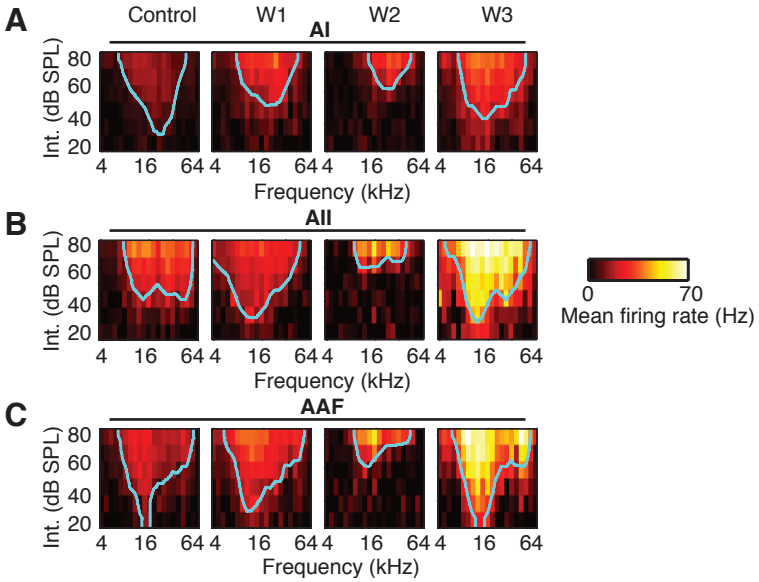

Figure 3.13: Baseline-subtracted average receptive field (RF) for each rearing condition, split across auditory areas. Blue contour lines are plotted to define the overall RF shape. Average RFs for controls, W1, W2, and W3 noise-reared groups in **(A)** AI. **(B)** AII, and **(C)** AAF.

In the late window group W3, receptive field quality was much higher, with normal tonotopy, response thresholds, and low incidence of multi-peaked RFs. In this group, bandwidths at absolute intensities in AI were significantly higher compared to controls (Fig. 3.14A, Kruskal-Wallis ANOVA followed by post-hoc Wilcoxon rank sum tests, $p < 0.05$). When comparing relative bandwidths starting at the threshold intensity, AI relative bandwidth in W3 was only significantly higher than controls in row 4 (Fig. 3.14B, Kruskal-Wallis ANOVA $p > 0.05$). In AII, both absolute (Fig. 3.14C) and relative (Fig. 3.14D) bandwidths increased in W3 compared to controls in row 3. Bandwidth changes in AI and AII cannot be explained by changes in intensity threshold across groups, since the only group with a significant change in threshold compared to other groups was W2.

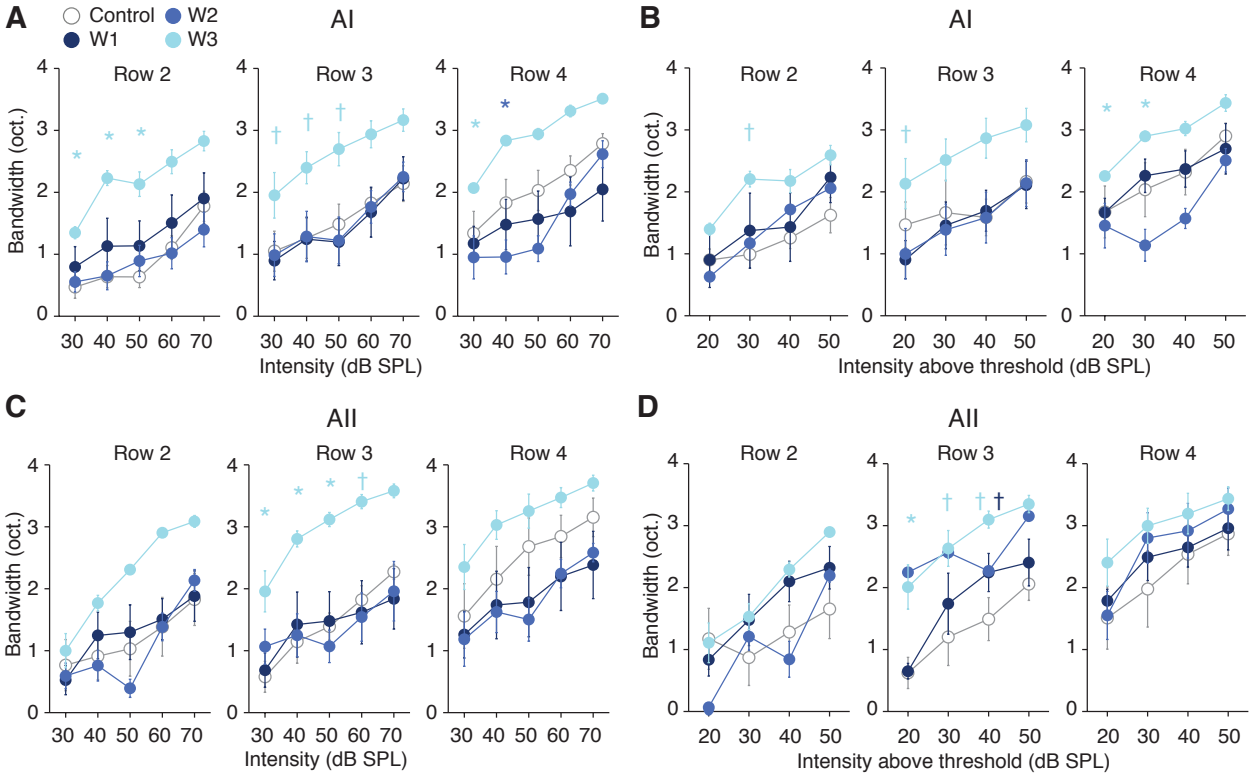

Figure 3.14: Pulsed noise in late window W3 (P24–31) increases absolute and relative bandwidths in AI and AII. Absolute bandwidths are shown in panels **(A)** and **(C)**, bandwidths relative to threshold are shown in **(B)** and **(D)**. Data are shown are mean ± SEM. Asterisks indicate level of significance from Wilcoxon rank sum test († $p < 0.1$, * $p < 0.05$).

Finally, we found that the spontaneous firing rate was significantly higher in W2 and W3 noise-reared groups compared to controls in AI and AII (Fig 3.15A,B). Increases in spontaneous firing in AAF in W3 were only at trend level and not significant (Fig. 3.15C). This is in accordance with our findings from the Ising model showing increased intrinsic activity as measured by the bias term.

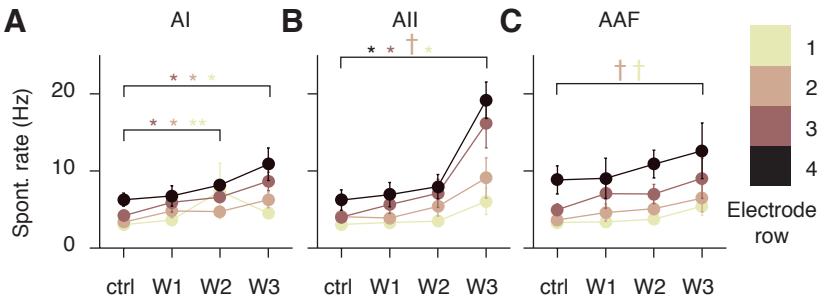

Figure 3.15: Noise rearing in late windows results in increased spontaneous firing rate in AI and AII. (**A**) Spontaneous firing rates in AI were significantly higher in all depths except row 4 for W2 and W3 groups compared to controls. (**B**) Spontaneous firing rates in AII were significantly higher for W3 compared to controls in rows 1,3, and 4, and at trend level in row 2. (**C**) In AAF, trend level increases were observed in spontaneous firing in W3 reared animals compared to controls in superficial rows 1 and 2.

3.4.5 Relationship between functional connectivity and receptive field changes

In traditional receptive field analysis, responses within a site are explained only by their relationship to an external stimulus, and do not include functional connections to other neurons. We can explore the effects of changes in intrinsic firing rate and functional connectivity on receptive field properties separately by generating simulated data from Ising model couplings. Here, we investigated whether and how the reduced spread index might be related to bandwidth changes seen in the single site analyses. To address this, we directly manipulated derived coupling matrices, regenerated spike data from these couplings using Gibbs sampling, and calculated the simulated tuning curves from the resulting data. We found that increasing the bias term increased the overall response magnitude and bandwidth of simulated tuning curves (Fig. 3.16A and C, top panel), but that manipulating site-to-site couplings to decrease the spread index by a factor similar to that seen experimentally did not significantly affect these parameters (Fig. 3.16B and C, bottom panel). Thus, it is likely that bandwidth changes observed in W3 are largely explained by increases in spontaneous firing/intrinsic firing from the bias term. The decreased spread index in W3 likely reflects disruptions in the precise timing of responses from the putative thalamorecipient to superficial layers (de Villers-Sidani et al. 2008; Zhou & Merzenich 2009).

3.5 Discussion

Neural circuits undergo experience-dependent modification during the critical period, whereby exposure to different types of external stimuli results in long-lasting and often permanent changes in representation in the brain. These changes are often examined on a single cell level or by comparing sensory responses in isolation. This neglects circuit level modifications that may underlie such changes in selectivity and representation. In our study, we use functional connectivity analyses to describe network level changes, and then compare the results to changes in individual receptive field parameters.

We use an Ising model to investigate functional connectivity because it has several advantages

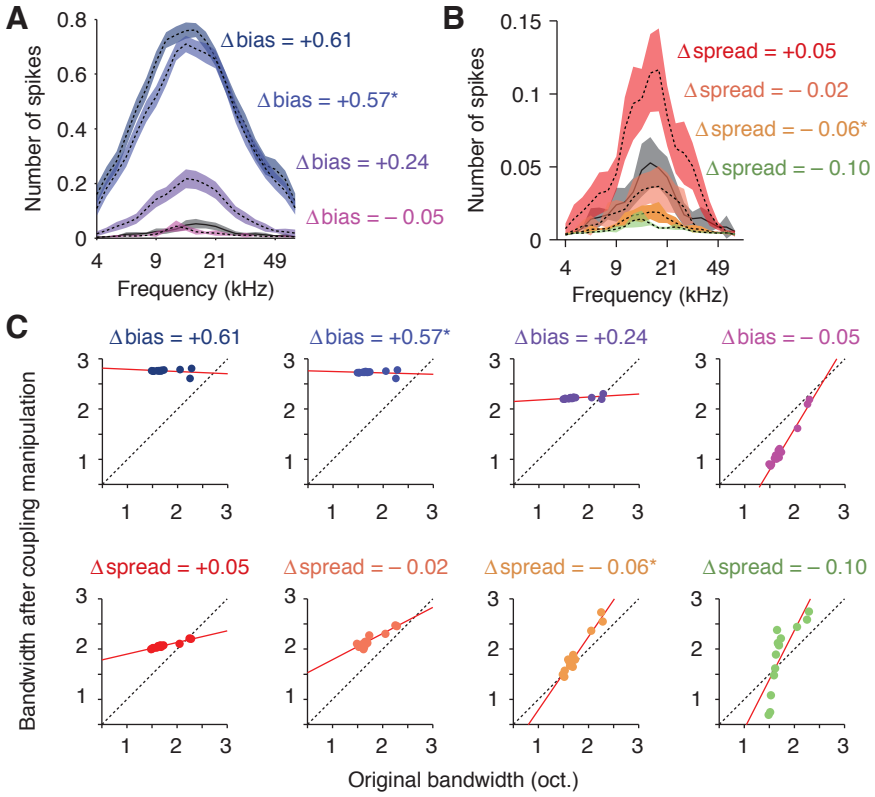

Figure 3.16: Increased bandwidth in W3 is likely a consequence of increased intrinsic excitability. (**A-B**) Simulated tuning curves generated using Gibbs Sampling on manipulated coupling matrices. (**A**) Increasing Ising model bias terms and deriving simulated activity results in tuning curves with increased amplitude and bandwidth. Average tuning curves for one site on the polytrode are shown, where the average was taken across 100 iterations of the model (see Methods). The average tuning curve created from sampling the original coupling matrix is shown in black/gray, and tuning curves resulting from manipulating the bias term by various factors are shown in shades of blue/purple. (**B**) Increasing the Ising model spread index results in tuning curves with increased bandwidth, decreased spread index generally results in small decreases in bandwidth. (**C**) Bandwidth measurements before and after coupling manipulation. Dashed line indicates the unity line, red line indicates least-squares regression line. With W3 noise rearing, we observed a $+0.57$ increase in the bias term (first row, 2nd column subplot), which in this simulation results in a significant increase in receptive field bandwidth ($p = 1.2 \times 10^{-4}$, Wilcoxon signed rank test). W3 subjects also showed a ≈ -0.06 decrease in the spread index (second row, 3rd column) which did not significantly change receptive field bandwidth in our simulation (p=0.15, Wilcoxon signed rank test).

over traditional correlation analyses. For one, all functional couplings are fit simultaneously, thus allowing us to uncover the underlying network connectivity with a lower incidence of false positive couplings arising from common inputs (Schneidman et al. 2006; Hamilton et al. 2013). Ising models have been used to recover connectivity in the canonical cortical circuit by revealing strong input to putative thalamorecipient layers and feedforward propagation of that input through intracortical vertical connectivity (Hamilton et al. 2013).

By applying the Ising model to recordings from the auditory cortex following exposure to pulsed

noise, we determine the relative effects on thalamocortical (direct sound evoked responses) and corticocortical circuitry in a principled manner. In our study, we found that early noise rearing led to decreased sound-to-site couplings in deep row sites, putatively located in the thalamorecipient layer. Effects were similar for "primary-like" fields AI and AAF, while more pronounced differences were seen between W1 and W2 in secondary area AII. These effects are consistent with observed increases in distorted RFs (in the case of W1) and increased thresholds (in W2) that we observed in the receptive field data. This is also consistent with early work showing increases in patchy or distorted RFs after pulsed noise exposure (Zhang et al. 2001) and disruptions in tonotopy, CF distribution, and threshold (Insanally et al. 2010). The finding of plasticity in thalamocortical circuits in early rearing windows is also in line with findings of earlier closure of the thalamocortical critical period compared to corticocortical (Feldman et al. 1999).

Although not addressed in this study, the reduction in corticocortical functional spread could be related to observed disruptions in GABAergic circuits. In particular, researchers have shown that parvalbumin-positive inhibitory interneurons, which are important both in the maturation of the cortex and closure of sensory critical periods (Chattopadhyaya et al. 2004; Hensch 2005; Hensch & Fagiolini 2005; Tropea et al. 2009) and in propagation of feedforward inputs through the auditory cortex (Hamilton et al. 2013), are reduced in number after exposure to band-limited noise in the critical period (de Villers-Sidani et al. 2008), and that those that remain have fewer dendritic processes. Noise exposure in early life also alters the expression of $GABA_A$ receptor subunits in the auditory cortex (Guo et al. 2012). Since parvalbumin-positive interneurons may aid in synchronizing the cortical network (Beierlein et al. 2000; Cardin et al. 2009; Sohal et al. 2009), their disruption could lead to deficits in spike timing coordination across layers and spatially distant sites in the cortex.

Though our study sheds light on the interaction between receptive field properties and circuit level modifications, couplings in the Ising model describe functional connectivity, and do not necessarily reflect synaptic connections (Roudi et al. 2009b). The mechanisms underlying the functional changes we see here could be due to changes in excitatory neurotransmission, inhibitory networks, or both. Recent work has suggested some relationship between thalamocortical synaptic plasticity and experience-dependent cortical plasticity (Blundon & Zakharenko 2013), so it may be that the changes observed in overall functional connectivity could also be observed at the synaptic level. Still, future research is needed to detail specific mechanisms.

By relating changes in stimulus selectivity to thalamocortical and corticocortical interactions, we have demonstrated that early exposure to a pulsed noise stimulus differentially affects thalamocortical functional connectivity, while late exposure affects corticocortical connections. By using a computationally tractable method where all possible pairwise interactions are fit simultaneously, we can disentangle changes to stimulus selectivity and changes to the neural circuits themselves. Thus, we provide a more complete view of the underpinnings of cortical plasticity at the level of the population.

Acknowledgments

This work was performed in collaboration with Jascha Sohl-Dickstein. This work was supported by the National Institute on Deafness and Other Communication Disorders through the National Institutes of Health (DC009259) and a National Science Foundation Graduate Research Fellowship (to L.S.H.).

CHAPTER 4

Optogenetic activation of an inhibitory network enhances feed-forward functional connectivity in auditory cortex

This paper is reproduced from work published in Neuron.

Citation: **Hamilton LS**, Sohl-Dickstein J, Huth AG, Carels VM, Deisseroth K, Bao S (2013). *Neuron* 80: 4, p1066-1076. `http://dx.doi.org/10.1016/j.neuron.2013.08.017`

4.1 Abstract

The mammalian neocortex is a highly interconnected network of different types of neurons organized into both layers and columns. Overlaid on this structural organization is a pattern of functional connectivity that can be rapidly and flexibly altered during behavior. Parvalbumin-positive (PV) inhibitory neurons, which are implicated in cortical oscillations and can change neuronal selectivity, may play a pivotal role in these dynamic changes. We found that optogenetic activation of PV+ neurons in the auditory cortex enhanced feed-forward functional connectivity in the putative thalamorecipient circuit and in cortical columnar circuits. In contrast, PV+ neuron stimulation induced no change in connectivity between sites in the same layers. The activity of PV+ neurons may thus serve as a gating mechanism to enhance feed-forward, but not lateral or feedback, information flow in cortical circuits. Functionally, it may preferentially enhance the contribution of bottom-up sensory inputs to perception.

4.2 Introduction

Neurons communicate with each other in dynamically modulated circuits. Functional connectivity, a measure of interactions between neurons in these circuits, can change gradually during learning (McIntosh & Gonzalez-Lima 1998) and formation of long-term memories, or can change rapidly depending on behavioral context and cognitive demands. While the mechanisms underlying long-

term network plasticity have been extensively documented, those underlying rapid modulation of functional connectivity remain largely unknown. At the network level, functional connectivity is affected by up-down and oscillatory states of the neural network (Gray et al. 1989). Cortical inhibition plays a key role in this process (Cardin et al. 2009; Sohal et al. 2009; Womelsdorf et al. 2007). PV-positive interneurons, which make up more than half of the inhibitory neurons in the cortex (Celio 1986), are particularly important as they provide strong feed-forward and feedback inhibition that can synchronize the cortical network (Cardin et al. 2009; Fuchs et al. 2007; Isaacson & Scanziani 2011; Sohal et al. 2009). Their precise influence on cortical networks during sensory processing, however, remains unclear. In particular, to date no studies have addressed how PV+ neurons may differentially modulate responses in different layers of the neocortex, and how the anatomical organization of the cortex affects the flow of information through these networks.

Histological studies have shown that the cortex consists of defined layers with vertical projections between those layers (Lee & Winer 2008; Linden & Schreiner 2003; Winer & Lee 2007). Functional connectivity within cortical networks has traditionally been investigated by measuring the cross-correlation between the spike trains of pairs of neurons (Douglas et al. 1989; Douglas & Martin 1991). Still, little is known about functional connectivity under sensory stimulation or about the role of inhibition in the cortical network. We combine multiple computational approaches with optogenetic activation of PV+ neurons to determine how inhibitory activity modulates network connectivity within and across layers and columns of the cortex.

4.3 Experimental procedures

4.3.1 Subjects

The University of California-Berkeley Animal Care and Use Committee approved all procedures. Subjects included 11 adult PV-Cre mice (strain B6;129P2-Pvalb$^{tm1(cre)Arbr/J}$, The Jackson Laboratory), aged ≈P100 at the time of recordings. 8 mice received an injection in the right auditory cortex at ≈P60 with 1 μL of a Cre-dependent adeno-associated viral vector carrying a double-floxed inverted copy of the light-sensitive cation channel channelrhodopsin-2 (pAAV–Ef1a–DIO–hChR2(H134R)–EYFP–WPRE–pA, 8×10^{12} viral particles/mL, UNC Vector Core) using a glass micropipette (Drummond Wiretrol, 10 μL) attached to a Quintessential Stereotaxic Injector (Stoelting) and procedures described elsewhere (Cardin et al. 2010). To control for the effect of light stimulation or heating of the cortex in general, 3 mice were injected with saline using the same protocol. A small burr hole (0.7 mm in diameter) was made over the right auditory cortex (1.75 mm rostral to lambda on the temporal ridge (Franklin & Paxinos 2008)), and virus (or saline) was delivered through a small durotomy. Each injection was performed in two stages, with 0.5 μL of virus injected at a depth of 500 μm and the remaining 0.5 μL injected at 250 μm, at a rate of 0.1 μL/s. Recordings were obtained after an infection period of approximately 1 month to ensure adequate expression of ChR2 throughout the auditory cortex, which was confirmed using GFP fluorescence goggles (BLS Ltd.).

4.3.2 Immunohistochemistry

Mice were perfused transcardially with cold 0.01M phosphate buffered saline (PBS, pH = 7.4), followed by 4% paraformaldehyde in 0.01M PBS. Brains were post-fixed in 4% PFA for 12h, and then cryoprotected in 30% sucrose for 18h. 50μm free-floating sections were cut using a cryostat (Leica CM3050). Every other section was incubated with blocking solution (10% normal goat serum in 0.01M PBS with 0.1% Triton X-100) for 45 minutes at room temperature (20°C), then incubated in primary antibody (PV 25 rabbit anti-parvalbumin, Swant, 1:4000 dilution) at 4°C overnight. The next day, slices were incubated in secondary antibody (Alexa 594 goat anti-rabbit IgG, Invitrogen, 1:200 dilution) for 1.5h at room temperature. Sections were mounted on gelatin-subbed glass slides with Fluoromount-G (Southern Biotech) and cover-slipped. Adjacent sections not stained for PV were washed in 0.01 M PBS, mounted on slides, and left to dry for 48h. They were then Nissl stained with 0.5% cresyl (w/v) for identification of individual cortical layers.

4.3.3 Quantification of virus expression

The spread of the virus was scored by hand by analyzing each 50 μm coronal section for the presence of eYFP fluorescence using a Zeiss LSM 780 34-Channel AxioExaminer fixed stage upright confocal microscope (UC Berkeley Molecular Imaging Center). Colocalization of ChR2-eYFP to PV-positive cells was analyzed by acquiring confocal images and identifying cells from each fluorescence channel by hand using ImageJ's cell counter plug-in.

4.3.4 Electrophysiological recording and stimuli

The right auditory cortex was mapped for each mouse under anesthesia using a cocktail of ketamine (100 mg/kg) and xylazine (10 mg/kg) and procedures described previously (Han et al. 2007). Following deflection of the temporal muscle, exposure of the auditory cortex, and removal of the dura mater, we performed a coarse mapping with tungsten electrodes (FHC, Inc.) to determine the location of primary auditory cortex based on rostrocaudal tonotopy and short spike latencies. We then recorded extracellular multi-unit neural activity in putative layer 2/3 through layer 4 of the right primary auditory cortex (Franklin & Paxinos 2008; Oviedo et al. 2010) using a 4 x 4 silicone polytrode (NeuroNexus A4$\times$4–3mm–100–125–177). 14/16 channels showed normal impedance measurements and were included in the analysis. A total of 350 multiunit sites (294 from ChR2-transfected animals, 56 from saline-injected controls) were used in our analyses. The polytrode was oriented parallel to the tonotopic axis and lowered orthogonally to the cortex such that the deepest contact sites were at a depth of approximately 500 microns from the pial surface (Fig. 4.1A).

The extracellular signal was obtained using a TDT amplifier connected to TDT RX5 hardware (Tucker Davis Technologies, Inc.), using a sampling rate of 25 kHz. Spike times were calculated by thresholding the extracellular signal at 1.5 times the standard deviation of the signal-to-noise ratio and bandpass filtering between 300-3000 Hz, and were logged using custom software running on a Windows XP computer. Sound stimuli were presented to the contralateral ear through an electrostatic cannulated speaker (EC1, Tucker Davis Technologies, Inc.) controlled by TDT RX6 hardware, and calibrated to ensure less than 3% spectral distortion and a flat output (< 3 dB devia-

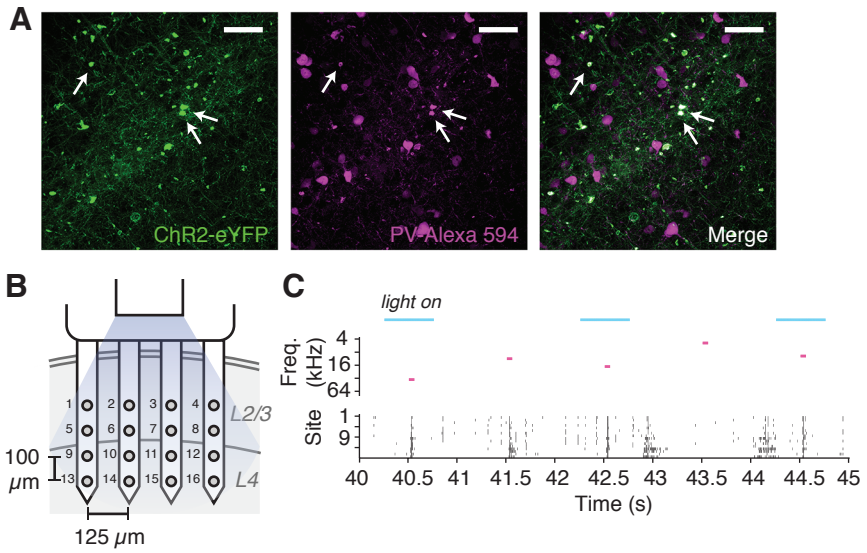

Figure 4.1: Viral expression, recording setup and responses to pure tone and optogenetic stimulation. **A.** We injected PV-Cre mice with 1 uL of a Cre-inducible adeno-associated virus (AAV) in the right auditory cortex that resulted in transfection of the light sensitive ion channel ChR2 in PV cells. Histology confirmed the colocalization of ChR2-eYFP (left) to parvalbumin-positive cells in the auditory cortex (center, immunostained with dsRed, and see merge). Approximately 58% of PV cells were transfected with ChR2 (white arrows indicate examples of colocalization). White scale bar = 50 μm. **B.** Schematic depicting recording setup. A 4x4 silicon polytrode was lowered orthogonally to the cortex such that the deepest sites were located at a depth of $\approx$500 μm. A 200 μm diameter optic fiber coupled to a 473 nm blue laser was positioned parallel to the polytrode, 1-2mm above the cortex to provide optogenetic stimulation during 50% of the trials. **C.** Light and sound stimulus conditions for example trials and corresponding spike raster plot. Input to the Ising model was a binary matrix including the light condition at each time point (blue bars represent the time during which the light was on and PV cells were being stimulated), the frequency of the pure tone stimulus that was presented at each time (represented by pink bars), and the spike data for each channel, binned at 5 ms. Sound and light conditions were randomly interleaved for each 1 second trial.

tion) from 4-75 kHz (Brüel and Kjær microphone, preamplifier, and conditioning amplifier, using SigCal32 software). Sound stimuli were pure tones generated using MATLAB (25 ms length with 5 ms squared cosine ramp, sampling rate 156.25 kHz) played from 4 kHz to 75 kHz in 0.2 octave steps, for a total of 22 frequencies. Sounds were presented at 6 different loudness levels (20 – 70 dB SPL, 10 dB spacing) in a pseudorandom order with a 1 Hz repetition rate, and each frequency-intensity pair was repeated 3 times. For the 50 dB level, stimuli were presented an additional 12 times to obtain higher resolution data at this intermediate level. For each trial of length 1 second, a tone pip would play at 500 ms into the trial. For half of the trials, we stimulated ChR2-transfected PV+ neurons using a 500 ms pulse of 473 nm blue laser light (Shanghai Laser and Optics Century Co., model BL473T3) coupled to a 200 micron optic fiber (ThorLabs, BFL37-200) beginning at 250 ms into the trial and controlled by a TTL pulse delivered by the RX5 hardware. This stimulation protocol results in the continuous spiking of the PV+ neurons throughout the duration of the light pulse (Zhao et al. 2011). The laser output was calibrated using a power meter (ThorLabs, PM100D with sensor S120C

and neutral density filter NE03A-A) to deliver light at an intensity of 1.2 mW, or ≈40 mW/mm2. This light intensity was chosen as the minimal light level that induced silencing of cortical activity throughout the light stimulation period. Photoelectric light artifacts (sharp transients locked to the onset of the light stimulus) were removed by excluding time points immediately surrounding the light onset (Cardin et al. 2010). Classical receptive fields were calculated for light-on and light-off trials separately by counting the number of spikes elicited by each frequency-intensity pair in a window defined by the peak of the PSTH. Receptive field thresholds were defined as the minimum sound intensity required to evoke a response (the intensity at the tip of the V-shaped receptive field). The receptive field bandwidths were calculated as the width of the frequency response in octaves 20 dB above the intensity threshold. Detection signal-to-noise ratio (SNR) was defined as

$$\text{SNR} = \frac{\text{\# evoked spikes - \# spontaneous spikes}}{\text{\# spontaneous spikes}} \tag{4.1}$$

for light-on and light-off epochs separately.

4.3.5 Ising model fitting

Binary matrices of the sound stimulus condition and spiking data for light-off and light-on trials separately were used as input to the model. Spiking activity from each of the contact sites on the polytrode was binned into 5 ms time epochs such that for each bin, if a spike occurred during those 5 ms, the value was set to 1, and if no spikes occurred, the value was set to 0. The bin duration of 5 ms was chosen according to the average cross-correlation between all pairs of recording sites, which showed that the mean cross-correlation was approximately three times the standard deviation of the baseline correlation at 5 ms. The sound stimulus matrix consisted of values for 22 frequency bins for each of t time points, with values of 1 when a stimulus at a given frequency was present and 0 when the stimulus was absent. Since cortical responses to sound occur with a delay relative to stimulus onset, we set the sound matrix to 1 for a window starting at 15 ms after the onset of the sound and ending at 50 ms after sound onset, corresponding to when the cross-correlation between the sound stimulus onset and neural spiking responses reached approximately three times the standard deviation of the baseline cross-correlation. Fitting separate sound-to-site couplings for each time delay relative to the stimulus onset (from 0 to 100 ms after sound onset, see Supplemental Methods) did not change our result (Fig. 4.12, 4.13). Only responses to the three highest decibel levels were used in the model (50, 60, and 70 dB). For each polytrode site, trials were randomized, the full data matrix was split into 10 equal chunks, and each model was estimated by holding out 1 of the data chunks and training on the remaining 90% of the data, and repeating this process 10 times for each possible training and validation set. This method, called 10-fold cross validation (Kohavi 1995), was used to ensure an accurate estimate of the log-likelihood that is more robust to noise in the data.

The stimulus conditioned Ising model is defined as:

$$p(\mathbf{x}|\mathbf{s};\mathbf{J},\mathbf{W}) \;=\; \frac{1}{\mathbf{Z}(\mathbf{s},\mathbf{J},\mathbf{W})}\exp\left(\mathbf{x}^{\mathrm{T}}\mathbf{J}\mathbf{x}+\mathbf{x}^{\mathrm{T}}\mathbf{W}\mathbf{s}\right) \tag{4.2}$$

where $\mathbf{x} \in \{0,1\}^N$ is the binary spike pattern for each time point, N is the number of recording sites, $\mathbf{J} \in \mathbb{R}$ is the site-to-site coupling matrix, $\mathbf{W} \in \mathbb{R}^{N\times M}$ is the sound-to-site coupling matrix, M

is the number of stimulus dimensions (in this case, sound frequencies presented), and $\mathbf{s} \in \{0,1\}^M$ is the stimulus input vector. A positive coupling value $\mathbf{J_{ij}} > 0$ indicates that sites $\mathbf{x_i}$ and $\mathbf{x_j}$ tend to be active simultaneously, while a coupling $\mathbf{J_{ij}} < 0$ indicates that when a spike occurs at site $\mathbf{x_i}$, $\mathbf{x_j}$ will be more likely to remain silent, and vice versa. Similarly, a positive sound-to-site coupling value in $\mathbf{W_{ij}}$ indicates that spiking in $\mathbf{x_i}$ tends to increase during presentation of stimulus $\mathbf{s_j}$, while a negative sound-to-site coupling value of $\mathbf{W_{ij}}$ indicates that spiking in $\mathbf{x_i}$ is suppressed during presentation of $\mathbf{s_j}$. Both site-to-site and sound-to-site couplings are unitless (much like linear regression weights, for example), with the magnitude of coupling indicating the strength of the relationship between their firing patterns. We were interested in the effect of the light stimulus condition on coupling, so separate coupling matrices $\mathbf{J}$ and $\mathbf{W}$ were trained for the light-off and light-on trials.

To estimate the couplings, we used Minimum Probability Flow Learning (MPF) (Schaub & Schultz 2012; Sohl-Dickstein et al. 2011) to minimize an L1 regularized version of the MPF objective function,

$$\begin{aligned} K(\mathbf{J},\mathbf{W}) &= \frac{1}{T}\sum_{\{\mathbf{x},\mathbf{s}\}}\sum_{\mathbf{x}'\in\mathcal{N}} \exp\left(\frac{1}{2}\left[E(\mathbf{x}|\mathbf{s};\mathbf{J},\mathbf{W}) - E(\mathbf{x}'|\mathbf{s};\mathbf{J},\mathbf{W})\right]\right) \\ &\quad +\lambda\left(||W||_1 + ||J||_1\right) \end{aligned} \tag{4.3}$$

where the sum over $\{x, s\}$ indicates a sum over all training observations, the neighborhood $\mathcal{N}(\mathbf{x})$ includes all states which differ from $\mathbf{x}$ by a single bitflip and the single state in which all bits are flipped, $E(\mathbf{x}|\mathbf{s};\mathbf{J},\mathbf{W}) = -\mathbf{x^T J x} - \mathbf{x^T W s}$ is the energy function of the Ising model, is the regularization strength, and T indicates the total number of training samples (in 5-ms binned time points).

The L1 regularization term $\lambda\left(||W||_1 + ||J||_1\right)$ was included to prevent over-fitting to training data. λ was chosen by cross validation from ten values logarithmically spaced between 10^{-7} and 10^{-2}. Cross validation was performed by holding out 20% of the training data, training the model using the remaining 80%, repeating this 5 times, and choosing the λ with the best average log-likelihood across all light conditions and all sites. The choice of λ had little effect on the log-likelihoods of the model fit for light-off trials, but there was improvement for the light-on models at intermediate λ values. Thus, we chose to use the same value of irrespective of light condition. λ was set to 5.9×10^{-5}.

Following selection of the regularization parameter, we fit the model using all of the training data, and the model log-likelihood, conditioned on the stimulus, was tested on the held out validation set. This was repeated 10 times for different validation sets, using the same regularization parameter. Coupling matrices shown in the figures are taken from the cross-validation iteration with the highest conditional likelihood on the validation set. We evaluated model likelihoods on held out data,

$$\log L = \frac{1}{T}\log p(\mathbf{x}|\mathbf{s};\mathbf{J},\mathbf{W}) \tag{4.4}$$

The normalization constant $\mathbf{Z}(\mathbf{s},\mathbf{J},\mathbf{W})$ required in the calculation of $p(\mathbf{x}|\mathbf{s};\mathbf{J},\mathbf{W})$ (Equation 4.2) was computed by exhaustive summation over all 2^{14} possible spiking states.

To test the effect of lowered baseline activity on Ising model couplings, we removed 20%, 50%, and 80% of spikes in all rows. Spikes were removed at random for each channel separately, and included both spontaneous and evoked data. We then re-ran the Ising model for the new manipulated spike data using cross-validation as before, and tested performance on a held-out set that had been manipulated similarly (20-80% spikes removed).

To test the effect of evoked activity, we removed all time points between 15 and 50 ms post-sound stimulus onset for each trial and fixed sound-to-site couplings to zero while training the model.

4.3.6 Vector Autoregression Modeling

To investigate the directionality of functional connections in the auditory cortex, we used vector autoregression (VAR) models (Lütkepohl 2005), which are linear models that relate multiple, dependent time series, such as spike trains, by a sum of linear weights. Such models allow us to predict the spiking activity of each site in the polytrode as a function of the previous spiking activity at all other sites. We fit models of the form:

$$\hat{x}_i(t) = \bar{x}_i + \sum_{j=1}^{N} \sum_{\tau=1}^{T} \beta_i(\tau, j) x_j(t-\tau) \tag{4.5}$$

where $\hat{x}_i(t)$ is the estimated response at recording site i at time t, $\bar{x}_i$ is the baseline firing rate of that site, β_i is a matrix of linear weights for the N simultaneously recorded sites over each of T time delays, and $x_j(t-\tau)$ is the response at recording site j at a given time in the past, $(t-\tau)$. T is the total number of time delays included in the analysis, and N is the total number of simultaneously recorded sites. We used delays up to 40 ms for each set of 14 simultaneously recorded sites. Those familiar with spectrotemporal receptive field (STRF) estimation will recognize this model as being essentially identical to a STRF (Aertsen & Johannesma 1981; Theunissen et al. 2001; Wu et al. 2006), with the difference being that neural activity is predicted from other activity in the network rather than by a parameterization of the external stimulus.

To solve for the VAR weights, we used ridge regression which is less prone to over-fitting than ordinary least squares. Ridge regression, also known as L2-penalized or Tikhonov regularization, minimizes the mean-square-error (MSE) between the actual and estimated response while constraining the L2 norm of the regression weights. The strength of the L2 penalty is determined by the ridge parameter, λ, where larger values of λ result in greater shrinkage of the weights (Asari & Zador 2009; Machens et al. 2002; Wu et al. 2006). In ridge regression, we minimize the following error function:

$$E(\beta_i) = ||x_i(t) - \hat{x}_i(t)||_2^2 + \lambda ||\beta_i||_2^2 \tag{4.6}$$

Where $x_i(t)$ is the true response of site i at time t, and the estimated response $\hat{x}(t)$ is given by equation [4.5] above. We estimated VAR weights using 80% of the data as a training set. Of the remaining 20% of the data, half was used for fitting the ridge parameter (10%) and half was used as a validation set to assess model performance (10%). The same recordings used in the Ising model were used in these analyses. Input to the model consisted of the binary spike trains binned at 2 ms

for each of the channels on the polytrode. Separate models were fit for light-on and light-off trials. To find the optimal ridge parameter, we tested 10 logarithmically spaced ridge parameters between 10^{-2} and 10^{5}, then selected the value that resulted in the highest average correlation on the (ridge) test set across all sites on the polytrode and both light-on and light-off models. The same ridge parameter was used for both light-on and light-off models.

Since a neuron's output is strictly positive and may scale nonlinearly with the input stimulus, we added an output nonlinearity to the model fitting (Wu et al. 2006). This nonlinearity does not change the weights of the model, but rather rescales the response predicted by the linear model to more accurately match the true response. We fit the nonlinearity as a univariate cubic spline that minimized the mean squared error between the actual and predicted responses on the training data. For both light-on and light-off models, adding the output nonlinearity significantly increased the predictive performance of the model ($p = 4.6 \times 10^{-10}$ and $p = 4.4 \times 10^{-16}$ for light-off and light-on respectively, Wilcoxon signed rank test), though these increases were quite small (0.6 ± 0.1 % increase for light-off, 1.5 ± 0.1 % increase for light-on). The increase in correlation was significantly higher for light-on over light-off ($p = 6.4 \times 10^{-13}$, Wilcoxon rank sum test), which is likely due to the overall lower firing rate during light-on trials.

VAR model validation was performed by calculating the correlation coefficient between the response predicted by the model and the actual response on the held out validation set. Significance of the correlation between predicted and actual responses was determined using resampling. The predicted response was randomly reshuffled 100,000 times, and the correlation between the shuffled prediction and actual response was computed. Reshuffling was done using 526-ms (263-time bin) segments to preserve local temporal statistics (this length was chosen to limit accidental alignment of the 1000-ms stimulation protocol across shuffled samples). The p-value of the model prediction was then computed as the fraction of the 100,000 shuffled correlations that were higher than the actual correlation.

4.3.7 Statistical tests

To test differences in coupling, we used Wilcoxon rank sum tests (for comparing independent groups) or Wilcoxon signed rank tests (for comparing paired groups) and corrected for multiple comparisons using Bonferroni correction. Parametric tests were not used because it was determined that the data being compared were not Gaussian distributed (Lilliefors test). Resampling techniques were used to obtain confidence intervals on correlation coefficients. Spearman rank correlations were used to test relationships between monotonically but not linearly related data, such as correlations and couplings. Values are reported as mean $\pm$ standard error unless otherwise stated.

4.4 Results and Discussion

We targeted expression of the light-sensitive channel channelrhodopsin-2 (ChR2) to PV+ neurons in the mouse auditory cortex (Fig. 4.1A) using a Cre-dependent adeno-associated virus (Sohal et al. 2009). One month post-transfection, we recorded neural responses with a 4 x 4 polytrode

in putative L2/3 through L4 of the primary auditory cortex (Fig. 4.1B) while playing pure tones to the contralateral ear and stimulating PV cells with blue light (Fig. 4.1C).

4.4.1 Using Ising models to recover functional connectivity in cortical circuits

Functional connectivity between the recorded sites was quantified using Ising models, which have previously been used to model neural interactions in many different systems (Ganmor et al. 2011b,a; Köster et al. 2013; Marre et al. 2009; Ohiorhenuan et al. 2010; Roudi et al. 2009b; Schaub & Schultz 2012; Schneidman et al. 2006; Shlens et al. 2006, 2009; Tang et al. 2008). The Ising model describes the coupling (a measure of functional connectivity) between pairs of recording sites and between recording sites and external stimuli based on observed population firing patterns and corresponding stimuli (Fig. 4.1B,C). By considering all pair-wise interactions simultaneously, Ising models are less prone to false positive interactions that are inherent to traditional correlation analysis (Schneidman et al. 2006). For example, in a fully connected Ising model (see Methods), the strongest coupling to sounds occurred in rows 3 and 4 (Fig. 4.2A), corresponding to the thalamorecipient layers. By contrast, traditional correlation analysis indicated strong connectivity between sounds and sites in all rows (Fig. 4.2B). This false positive connectivity between sounds and activity in rows 1 and 2 is due to the absence of site-to-site interactions in the correlation analysis. In a reduced Ising model where recording sites were coupled to sound but not to each other, which we call the independent neurons model, positive couplings between neural activity and the sound stimulus were also present in all recorded layers and did not differ across depth (Fig. 4.2C, $p = 0.55$, Kruskal-Wallis ANOVA).

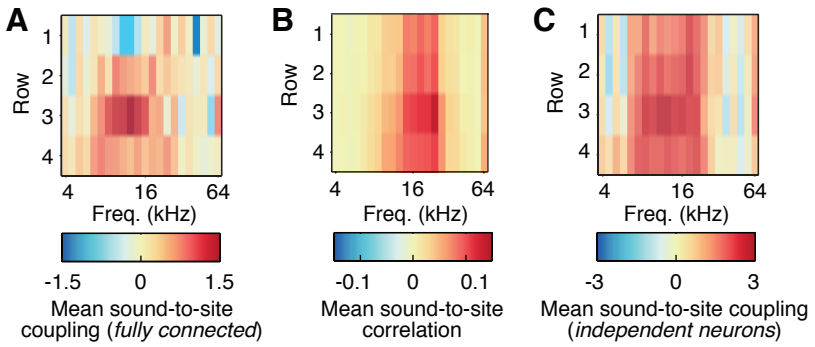

Figure 4.2: The Ising model recovers canonical cortical structure not observed with traditional correlation analysis. **A.** Mean sound-to-site coupling derived from one set of polytrode recording data using the fully connected model. When connections between sites are taken into consideration, the strongest sound-to-site couplings are seen in rows 3 and 4, the putative thalamic input layers. **B.** Mean sound-to-site correlation from the same dataset as in **A.** Neural activity on all rows of the polytrode was positively correlated with sound presentation. **C.** Mean sound-to-site coupling derived from the same dataset as in A using the independent neurons model. In this model, the coupling between neural activity and sound does not change as a function of depth (electrode row).

Furthermore, pairwise correlations were more tightly correlated with couplings in the independent neurons model than in the fully connected model (Fig. 4.3A, correlations resampled 100,000

times, with the difference in correlations significant at bootstrapped $p < 1 \times 10^{-4}$). The fully connected model showed significantly higher log-likelihood on held-out data than the independent model (Fig. 4.3B, $p = 0.013$, Wilcoxon signed rank test), suggesting a significant contribution of site-to-site interactions to neuronal activity.

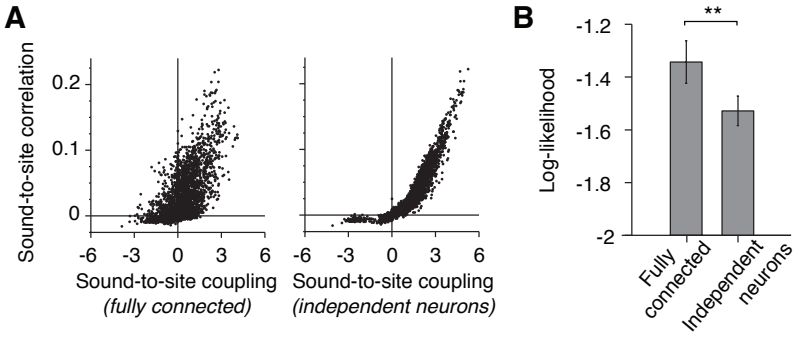

Figure 4.3: **A.** Pairwise correlations as a function of Ising model couplings for the fully connected model of connectivity for all subjects and locations (left, Spearman $\rho = 0.61$). Pairwise correlations as a function of Ising model couplings for the independent neurons model of connectivity for all subjects and locations (right, Spearman $\rho = 0.97$). When connections between neurons are considered, couplings are more distinct from correlations. **B.** Ising model performance for the fully connected model and the independent neurons model. The fully connected model showed the highest log-likelihood on held out data. Data are presented as mean $\pm$ S.E.M. Asterisks indicate level of significance: ** = $p < 0.01$.

The Ising model can discover spatial structure within the network despite no prior knowledge of spatial locations of the polytrode recording sites. In the fully connected Ising model, coupling was stronger in the vertical and horizontal than diagonal directions (Fig. 4.4A), presumably due to neuronal projections within cortical columns and layers. In addition, coupling decreased more rapidly with vertical than horizontal distance — sites up to 375 μm apart horizontally were still more strongly coupled than sites 300 μm away vertically ($p = 4.3 \times 10^{-6}$, Wilcoxon rank sum test). Such connectivity structure was much less prominent in the pairwise correlations (Fig. 4.4B; ratio of column or layer/diagonal couplings = 1.26 $\pm$ 0.03 for correlations, 2.16 $\pm$ 0.20 for couplings; $p = 0.001$, Wilcoxon rank sum test). Thus, although the model is blind to the relative locations of the recording sites, the fully connected Ising model recovered known layer and column circuitry (Linden & Schreiner 2003; Mountcastle 1957).

4.4.2 Optogenetic activation of PV+ neurons enhances functional connectivity

Using the fully connected Ising model, we analyzed how optogenetic activation of PV+ neurons influences functional connectivity in laminar, columnar, and thalamic input circuits of the primary auditory cortex. In keeping with PV+ neurons providing inhibitory input to connected pyramidal cells, we saw an overall reduction of the Ising model bias term in light-on trials, reflecting reduced firing rates in all rows (Fig. 4.5A, Bonferroni-corrected $p = 0.003$, $p = 0.0002$, $p = 8.4 \times 10^{-6}$, and $p = 8.7 \times 10^{-5}$ for rows 1, 2, 3, and 4, Wilcoxon signed rank tests). Furthermore, we found that stimulating PV+ neurons led to increases in vertical connectivity between sites within the same

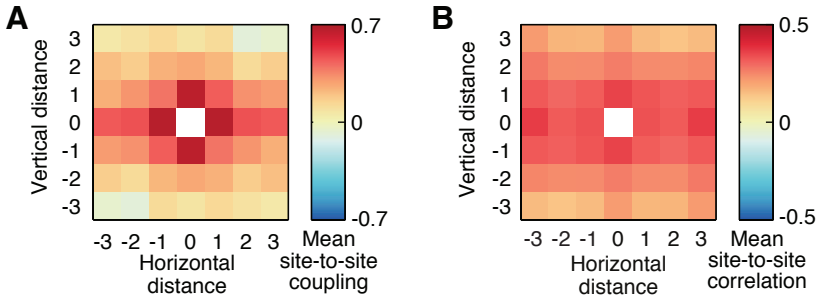

Figure 4.4: **A.** Coupling between pairs of recording sites in the fully connected model as a function of horizontal and vertical distance collapsed across all locations. Laminar and columnar structure arises from the model in the form of strong positive couplings for horizontal and vertical distances = 0, compared to couplings between sites situated diagonally (horizontal and vertical distance $\neq 0$), which are weaker. **B.** Mean correlation between sites as a function of horizontal and vertical distances collapsed across all locations. The correlation falls off strongly with vertical distance but remains more uniform across horizontal distance.

vertical column (Fig. 4.5B, Bonferroni-corrected $p = 0.01$ and $p = 1 \times 10^{-4}$ for coupling between sites within the same column, 2 and 3 rows away, respectively, Wilcoxon signed rank tests), but did not change horizontal connectivity within layers ($p > 0.05$, Wilcoxon signed rank tests).

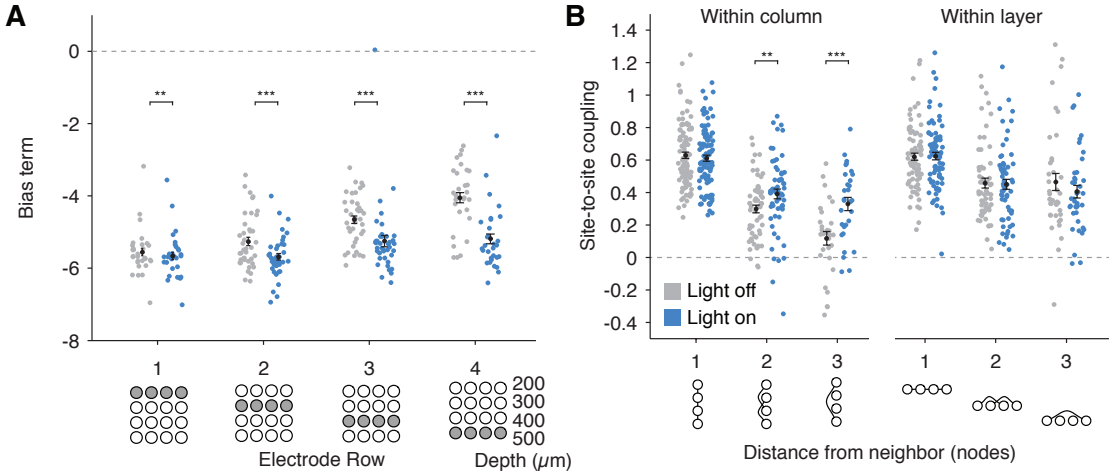

Figure 4.5: Optogenetic activation of PV+ neurons decreases average firing rate while increasing functional connectivity within columns. **A.** Bias term for the Ising model for each site in different rows on the polytrode, representing a proxy for the intrinsic firing rate of each site. Light stimulation significantly reduced the bias term in all layers, indicating an overall reduction of firing rate with PV+ neuron stimulation. **B.** Site-to-site couplings as a function of the distance between the sites. Couplings are plotted between sites 1, 2, and 3 nodes away within a column or within a layer (see node diagrams below x-axis). Light stimulation significantly increased couplings within a column at distances 2 and 3 nodes away and did not change couplings within a layer. Asterisks indicate level of significance: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, Bonferroni corrected for multiple comparisons. Errorbars and black marker depict mean $\pm$ S.E.M. See also Fig. 4.12 and 4.13.

Coupling between neural activity and sounds increased for sites in rows 3 and 4 during PV+ neuron stimulation (Fig. 4.6A,B, Bonferroni-corrected $p = 0.0003$ and $p = 8 \times 10^{-13}$ for third and fourth rows, respectively, Wilcoxon signed rank tests). These sites were likely located in the thalamorecipient input layers (layer 4 and deep layer 3). The increase in sound-to-site coupling in putative thalamorecipient layers was not an artifact of the response window selection (Fig. 4.12 and 4.13). Our findings indicate that activation of PV+ neurons results in enhanced functional connectivity specifically in thalamocortical input and cortical columnar circuits.

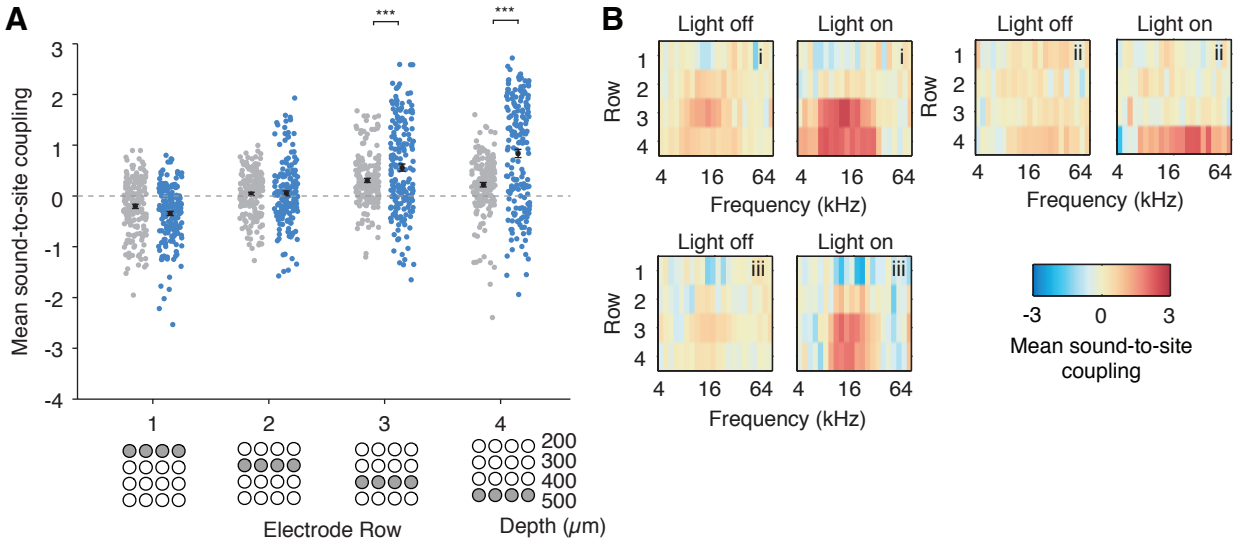

Figure 4.6: Optogenetic activation of PV+ neurons increases functional connectivity from sounds to sites located in putative thalamorecipient layers. **A.** Sound-to-site coupling averaged across sites in each row. During optogenetic stimulation, coupling between neural activity and sound stimulation increased in the putative thalamorecipient layers (rows 3 and 4). **B.** Examples of sound-to-site coupling in sites recorded from three different animals, plotted as a function of tone frequency and electrode row. In all cases, there was an increase in coupling to sounds in the putative thalamic input layers (rows 3 and 4) during light stimulation of PV+ neurons. Some locations showed decreases in coupling to sounds in the superficial layers during light stimulation, but this effect was not consistent across the population (see panel **A**). Asterisks indicate level of significance: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, Bonferroni corrected for multiple comparisons. Errorbars and black marker in **A** depict mean $\pm$ S.E.M. See also Fig. S1.

4.4.3 PV+ neuron stimulation enhances functional connectivity in the feedforward direction

While the Ising model uncovers altered functional connectivity with inhibitory neuron stimulation, it is agnostic to the direction in which these changes occur. For example, the increased coupling within cortical columns during activation of PV+ neurons could be in the feed-forward, feedback, or both directions. To address this issue, we used vector autoregression (VAR) to derive a linear model that described how activity in one site was modulated by spikes in other sites as a function of

time delay (Fig. 4.7A; see Methods for details). Unlike the Ising model, which describes dynamics within a fixed time bin, this model considers how inputs from different rows at different times affect the neural responses in a given time. Prediction of one site's activity using the population activity was significantly better during the light-on than the light-off epochs (Fig. 4.7B, Wilcoxon signed rank test, $p = 4.0 \times 10^{-10}$).

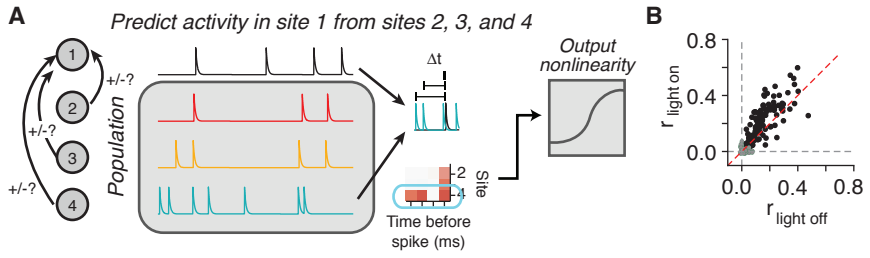

Figure 4.7: **A.** Schematic of the vector autoregression analysis. To model the interaction between the activity of sites at different depths, we fit a vector autoregression model (VAR) which predicts the activity for each site on the polytrode based on the activity of all other sites. Here, activity in site 1 is predicted from activity in sites 2 – 4. We fit a linear model describing how spikes in each site modulate spiking in the predicted site at different time delays. We predict responses based on these weights and apply an output nonlinearity to improve predictions. **B.** Prediction performance of VAR models is significantly higher for light-on epochs compared to light-off epochs ($p < 4.0 \times 10^{-10}$, Wilcoxon signed rank test). Models performing significantly better than chance are indicated in black. Gray points show VAR models excluded from analysis. The red dashed line is the unity line $y = x$.

We then examined the contribution of each site to predicting the activity of another site (i.e., the weight function in the linear model as a measure of functional connectivity; Fig. 4.8A,B). In general, neural activity was more strongly modulated by activity of sites in the same cortical layers rather than in different layers. However, these weights were not significantly altered by activation of PV+ neurons (Fig. 4.8B, diagonal subplots). By contrast, PV+ neuron activation significantly increased the weights for row 4 sites in predicting the activity of more superficial sites within a time window between 6 and 12 ms (Fig. 4.8B, far right subplots). There was also a small trend (not significant) of increased excitatory drive from row 3 to row 4, consistent with the primary input layer to auditory cortex arising in deep layer 3 and propagating information to layer 4 (Smith and Populin, 2001). Furthermore, inhibitory influences from superficial row 1 on activity in row 3 were lessened with PV+ neuron stimulation (Fig. 4.8B, first column, third row subplot), suggesting that the normal feedback inhibition from superficial layers is altered when PV+ neurons inhibit those cells. The double dissociation between the stronger baseline intra-layer influences and the light-activated increase for cross-layer influences supports our findings from the Ising model analysis that the activation of PV+ neurons specifically increases intra-column functional connectivity. The increased contribution of activity in row 4 to firing in superficial rows during light stimulation further suggests that the enhanced functional connectivity is in the feed-forward direction.

Qualitatively similar results were also observed when fitting the data in a generalized linear model (GLM) with an exponential nonlinearity (see Pillow et al. (2008)), although predictive performance of the GLM as assessed by correlation was worse than the VAR model for both light-off

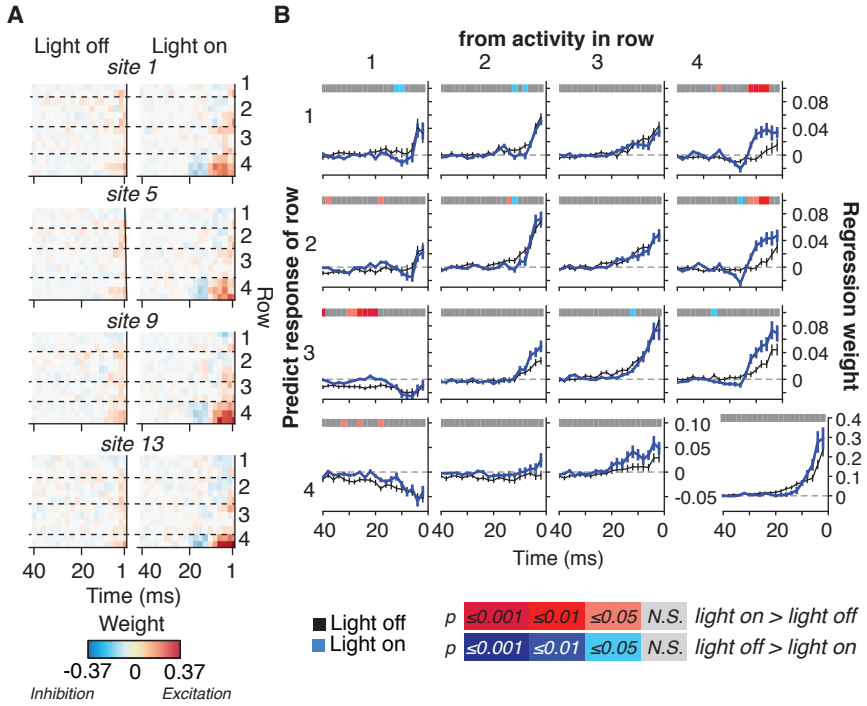

Figure 4.8: PV+ neuron activation increases functional connectivity in the feed-forward direction. **A.** Representative VAR model weights from one subject for sites on one shank of the polytrode. Red = excitatory drive, blue = inhibitory. Dashed lines separate different rows. **B.** Average VAR model weights (β from equation [4.5]) collapsed across all subjects and across contact sites in the same row show that light stimulation of PV cells results in stronger and longer excitatory drive from row 4 to rows 1 and 2. VAR model weights were higher for light-on (Bonferroni-corrected p-value indicated by color on horizontal line, Wilcoxon signed rank test) when predicting rows 1-2 from activity in row 4. Row 2 also showed significantly more suppression from the putative thalamorecipient layer before the window of strong excitation starting at 12 ms. When predicting row 3 activity from activity in superficial row 1, we found that PV+ neuron stimulation reduced the inhibitory influence of row 1 on row 3, as evidenced by regression weights closer to zero. The strong increase in drive from row 4 to more superficial layers and the small regression weights in the feedback direction shows that PV+ neuron stimulation increases drive in the feed-forward direction. Data are presented as mean regression weight across all sites and subjects $\pm$ S.E.M. See also Fig. S2.

and light-on conditions ($p < 0.0001$, Wilcoxon signed rank tests; see Supplemental Methods and Fig. S2). These results, together with the increased sound-to-site coupling in the feed-forward thalamocortical circuit, suggest that activation of auditory cortical PV+ neurons may facilitate bottom-up information flow in the feed-forward direction.

4.4.4 Activation of PV+ neurons increases detection signal-to-noise ratio in single recording sites

Previous studies have shown that optogenetic activation of PV+ neurons enhances stimulus feature selectivity and increases the signal-to-noise ratio in cortical neurons (Atallah et al. 2012; Lee et al.

2012; Sohal et al. 2009; Wilson et al. 2012). In our study, light activation of PV+ neurons induced strong suppression of spontaneous firing and weak reduction of tone-evoked responses (Mean % suppression $\pm$ SEM = 31.77 $\pm$ 0.03 % for spontaneous, 18.57 $\pm$ 0.03 % for evoked, see Fig. 4.9A,B for example peristimulus time histograms and receptive fields). This led to an increase in the detection signal-to-noise ratio (mean detection SNR $\pm$ SEM = 6.13 ± 0.73 for light-on vs. 3.17 ± 0.21 for light-off trials, $p = 0.005$ Wilcoxon signed rank test, Fig. 4.9C).

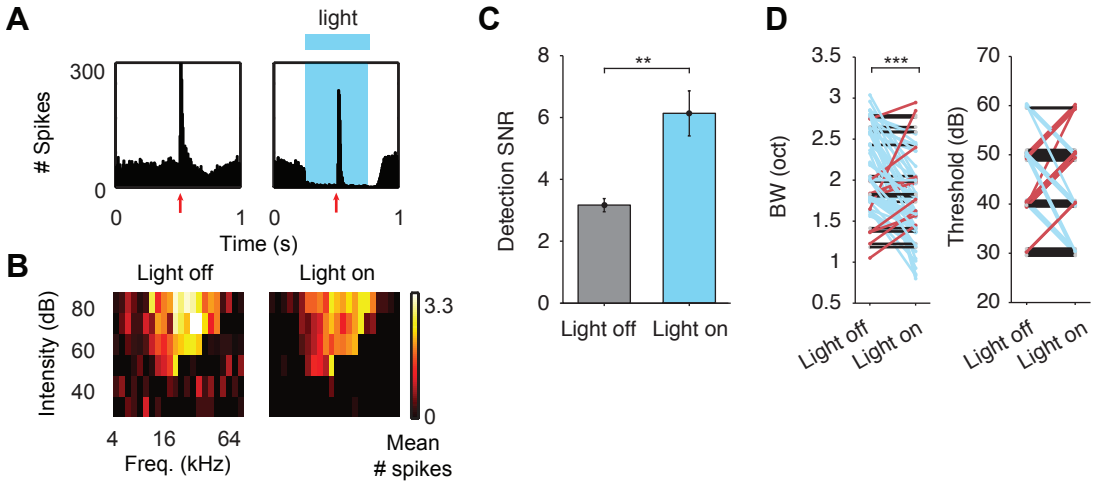

Figure 4.9: PV+ neuron activation strongly suppresses spontaneous activity, weakly suppresses sound-evoked activity, and improves detection signal-to-noise ratio of responses. **A.** Example peristimulus time histogram of multiunit activity. Trials with no light stimulation (left) showed strong responses to pure tones played at 0.5 s (red arrow). When blue light was used to stimulate PV+ neurons between 0.25 and 0.75 s (right), responses to sound were still observed, but the overall evoked response was reduced. Spontaneous activity was reduced throughout the duration of the light stimulus. **B.** Example receptive fields during "light off" (left) and "light on" (right) trials separately for the same site in A. Each pixel in the plot represents the average number of spikes evoked by a stimulus at a particular frequency and intensity level. The reduction in spontaneous rate is evident in the decreased background spikes outside of the V-shaped receptive field. **C.** Detection signal-to-noise ratio (SNR) during "light off" (gray) and "light on" (blue) trials. Optogenetic activation of PV+ neurons results in a significant increase in the SNR compared to the no stimulation condition. Data are presented as mean SNR ± S.E.M. **D.** Receptive field bandwidths 20 dB above threshold (left) and receptive field thresholds (right) for receptive fields calculated from "light off" (no stimulation of PV+ neurons) and "light on" (stimulation of PV+ neurons) trials separately. Lines are colored according to whether a reduction (blue), increase (red), or no change (black) in bandwidth was observed. We observed a significant reduction in receptive field bandwidths during PV+ neuron stimulation ($p < 0.001$, Wilcoxon signed-rank test) with no change in threshold. This indicates that stimulating PV+ neurons increases stimulus selectivity by narrowing the range of stimuli to which a site responds. Asterisks indicate level of significance: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, Wilcoxon signed-rank tests. See also Fig. S3.

In addition, PV+ neuron stimulation significantly narrowed receptive field bandwidths ($p < 0.001$, Wilcoxon signed rank test), without changing response thresholds at the characteristic frequency ($p = 0.79$, Wilcoxon signed rank test, Fig. 4.9D). In sham-injected control mice not expressing ChR2, light stimulation did not cause any significant change in response properties (Fig. 4.16).

4.4.5 Reductions in spontaneous activity alone do not account for functional connectivity changes

To test the possibility that reduced spontaneous activity and increased detection SNR (Fig. 4.9A-C) caused the observed increases in site-to-site coupling (Fig. 4.5B), we randomly removed 20-80% of spikes recorded in light-off trials to mimic the effects of PV+ neuron stimulation with light and re-ran the Ising model analysis (see Methods). The mean site-to-site coupling strength was not increased by the random reduction of spontaneous and evoked spikes (Fig. 4.10A), but rather was reduced in sites one node away within the same column ($p < 0.001$ for all comparisons, Bonferroni-corrected Wilcoxon signed rank tests). No changes to coupling between sites 2 and 3 away within the column were seen (Bonferroni-corrected $p > 0.05$, Wilcoxon signed rank tests), even with reductions in activity that were far larger than the suppression caused by PV+ neuron stimulation (≈32% suppression on average). There was also no change in sound-to-site coupling with these manipulations (Fig. 4.10B).

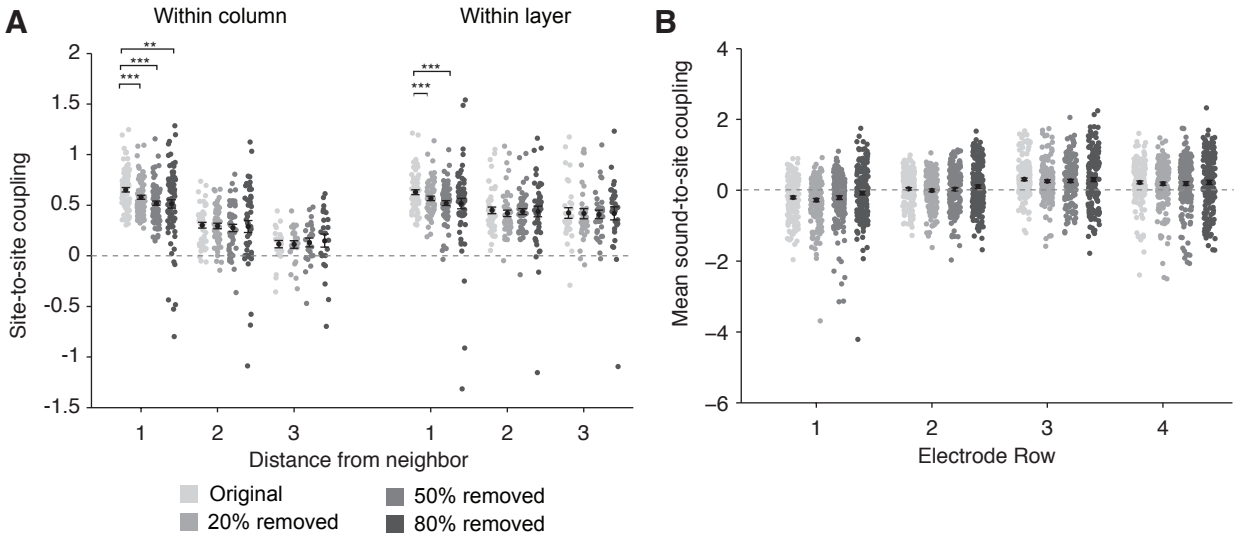

Figure 4.10: Manipulation of spontaneous and evoked activity on Ising model couplings. **A.** Random reduction of spikes in all layers during "light off" trials does not increase site-to-site couplings in the Ising model as did PV+ neuron stimulation. Instead, significant reductions in site-to-site couplings are seen in both column and layer couplings 1 node away for several levels of firing rate reduction (20, 50, or 80% reduction). **B.** Random reduction of spikes in all layers does not alter sound-to-site couplings. Data are presented as mean ± SEM.

Finally, to determine if the altered site-to-site coupling strength was due to changes in evoked activity, we removed sound-evoked spikes and re-ran the analysis with only the (unaltered) spontaneous activity. The coupling strength was still higher during activation of the PV+ neurons (Fig. 4.11, Bonferroni-corrected $p = 0.002$ and $p = 0.0002$ for sites 2 and 3 away within a column, respectively, Wilcoxon signed rank tests). These results indicate that the enhanced coupling by PV+ neuron activation was not due to the increased detection signal-to-noise ratio or reduced baseline activity. Rather, it reflects the state of the circuit connectivity, and is independent of sensory stim-

ulation and responses.

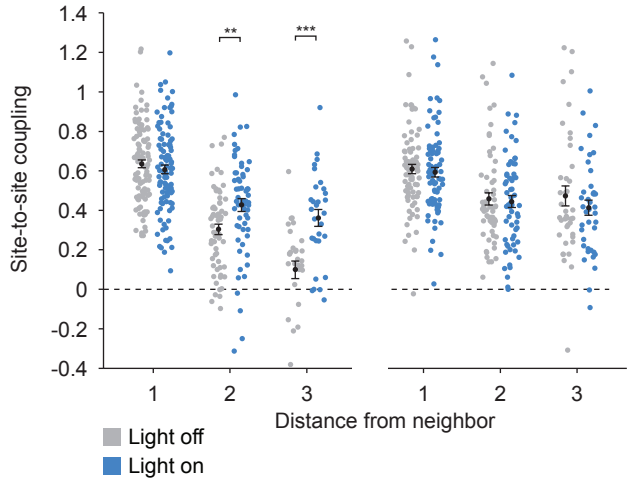

Figure 4.11: Removing sound-evoked periods from the spike trains does not change the effect of PV+ neuron stimulation on site-to-site couplings, highlighting the importance of network connectivity over stimulus input. Data are presented as mean ± SEM.

4.4.6 Possible mechanisms for the enhancement of functional connectivity by PV+ neurons

In this study, we quantified functional connectivity in the auditory cortex with coupling from the Ising model and the weight function from vector autoregression. Both measures elucidate how the activity of a neuron or the presentation of a sound stimulus drives the firing of a target neuron. The specific mechanisms underlying the modulation of functional connectivity by PV+ neurons are not investigated in the present study, but could involve the modulation of synaptic connections and changes in global network states. For example, synaptic efficacy can be rapidly altered by the prior synaptic activity (Zucker & Regehr 2002), which is likely influenced by the activity of PV+ neurons. Alternatively, by synchronizing network activity (Cardin et al. 2009; Sohal et al. 2009), PV+ neurons could set target neurons in a more excitable state when the projection neuron fires, thus enhancing their functional connectivity. The effects on column rather than layer connections may be related to anisotropic projection patterns of PV+ neurons (Packer & Yuste 2011), whereby PV+ neurons preferentially inhibit pyramidal neurons located in the same vertical columns over distances 200 μm and greater.

4.4.7 Potential limitations

While both the Ising model and the VAR models allow us to analyze the relative changes to within vs. between layer connectivity with PV+ neuron stimulation, some caution should be taken when inter-

preting these functional connections in terms of synaptic interactions. With extracellular recordings it is not possible to reconstruct the synaptic connections between recorded (or stimulated) neurons. Coupling between neurons should be considered as a functional description rather than an anatomical one. For example, researchers have found that coupling weights in the Ising model do not necessarily correspond to synaptic connections in the network (Roudi et al. 2009a). The strength of the Ising model lies in its ability to distinguish direct from indirect interactions, for example in finding direct stimulus input to rows 3 and 4, representing the thalamorecipient layer. However, the symmetric nature of Ising model couplings means that directed interactions, such as combined excitatory/inhibitory influences (cell A excites B, but B inhibits A), cannot be uncovered. The VAR model addresses some of these caveats, since it can quantify directional interactions between recording sites and describe how neuronal firing is affected in different time periods. Our model shows that strong feed-forward drive is enhanced by PV+ neuron stimulation, whereas feedback from superficial to putative thalamic input layers is not affected.

4.4.8 Implications for sensory perception

Specific enhancement of feed-forward connectivity has important implications in processing sensory information. Sensory perception is derived from both bottom-up sensory inputs and top-down stimulus expectations (Kording & Wolpert 2004; Stocker & Simoncelli 2006). Previous theoretical work indicates that sensory cortical neurons could integrate multiple sources of information by linear summation of population responses activated by each source. To achieve optimal integration under this scheme, however, the weight placed on each information source must be dynamically adjusted according to the quality of the information and task demands (Ma et al. 2006). Interestingly, the firing rate of fast-spiking neurons — likely PV+ neurons (Kawaguchi & Kubota 1998; Toledo-Rodriguez et al. 2004) — appears to increase with demand of attention to external stimuli (Chen et al. 2008; Mitchell et al. 2007). Our results show that activation of PV+ neurons preferentially emphasizes bottom-up sensory information by increasing feed-forward connectivity, without changing the weight on top-down information presumably supplied through lateral or feedback connections. Thus, PV+ neurons may play an important role in optimal integration of sensory information with top-down expectations in sensory perception. These results could inform future work on mechanisms of sensory pathologies in patients with autism and schizophrenia, both of which are associated with PV+ neuron dysfunction (Gandal et al. 2012; Gonzalez-Burgos & Lewis 2012).

4.5 Supplementary Materials

4.5.1 Supplementary Methods

To expand our analysis of sound-to-site coupling and describe changes in sound-to-site coupling over time, we fit fully-connected Ising models with 20 time delays (from 0 to 100 ms after the sound stimulus onset, in 5 ms bins) for each of 22 frequencies, for a total of 440 sound-to-site couplings per model. In this case, the Ising model formulation is identical to what was used previously, with the

exception of a change in the input matrices. For this analysis, instead of a $22 \times t$ sound input matrix, where for each of 22 frequencies, the value was set to 1 for the period between 15 and 50 ms, we have a $440 \times t$ matrix, where values are set to 1 for only 1 bin at a time, for each frequency/time delay combination separately. The site-to-site couplings are calculated as before, with all possible pairwise connections fit simultaneously with the sound-to-site couplings. Results are shown in Figure S1.

We compared the VAR model to an exponential generalized linear model (GLM) with Poisson likelihood (Pillow et al. 2008). This model is similar to the VAR model, but includes an assumed exponential nonlinearity in the fitting (Fig. 4.14). We fit the GLM on the same data as the VAR model, using gradient descent with early stopping and fitting "light on" and "light off" models simultaneously so the regularization would not differ between models, thus all weights would be comparable. We fit a GLM using maximum likelihood with the following objective function for the population log-likelihood:

$$\log L = \sum_{sp} \log x(t_{sp}) - \int x(t)dt \tag{4.7}$$

where $x(t) = \exp(\mathbf{h} \cdot \mathbf{y}_t + \mu)$ is the conditional spike rate at time t, and $\mathbf{h}$ represents the weights of the model as a spike history filter (for both the recording site's own history or the spike history of all other recorded sites), and $\mathbf{y}_t$ represents the spike train of the site to be fit as well as all other simultaneously recorded sites, and μ represents the average baseline firing rate of the site to be fit. The model was trained for each location separately using 80% of the data, with 10% of the data not used in training as an early stopping set to prevent overfitting. The remaining 10% of the data was used to calculate the predictive performance of the model, as assessed by the correlation between the predicted response and the actual response on the validation data. The results of this model were comparable to that of the VAR model (Fig. 4.15).

4.5.2 Supplementary Figures

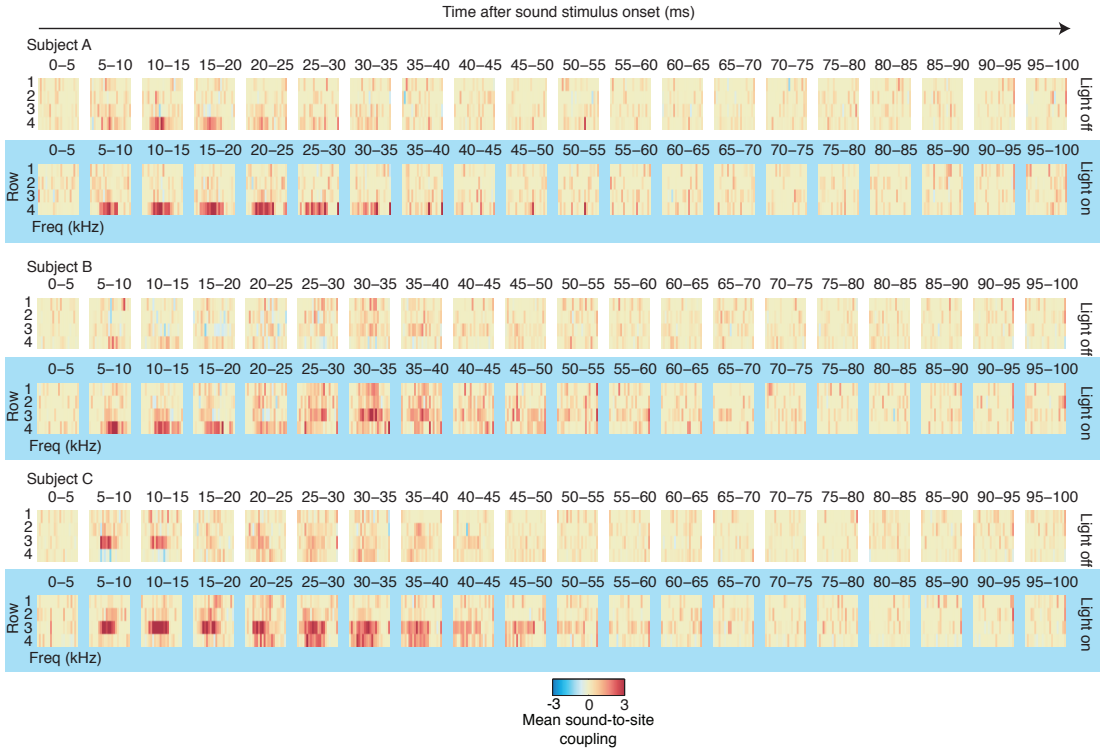

Figure 4.12: Effects of response window selection on the measurement of sound-site coupling and its modulation by PV+ neuron stimulation, related to Fig. 4.5. Sound-to-site couplings in these plots were fitted separately for 20 time windows relative to the sound stimulus onset (see Supplementary Methods). Example mean sound-to-site couplings within each row of the polytrode for "light off" (top) and "light on" (bottom, blue background) models. Strong sound-to-site coupling was observed in putative thalamorecipient layers in early windows for both "light on" and "light off" models. During PV+ neuron stimulation ("light on"), sound coupling to thalamic input layers was enhanced and can be observed for a longer period of time.

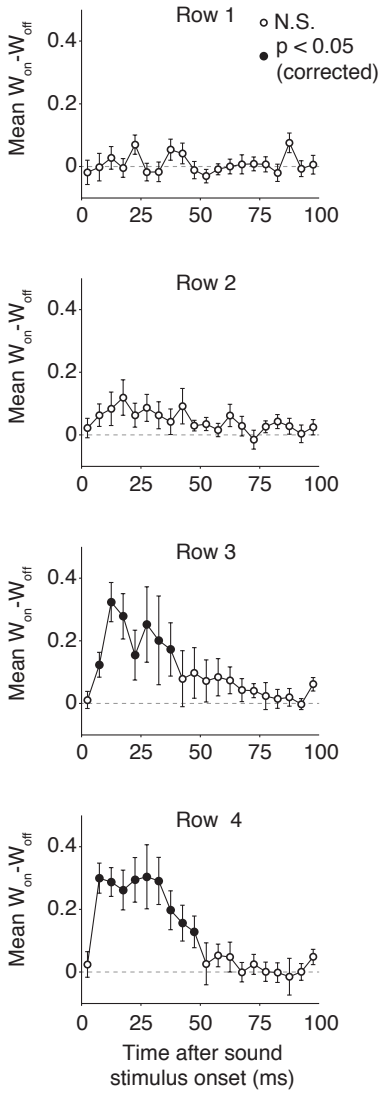

Figure 4.13: Average difference between "light on" and "light off" sound-to-site coupling for each row, collapsed across all subjects and sites. Sound-to-site couplings were significantly higher in row 3 and 4 for "light on" trials compared to "light off" trials. Filled circles show time points where the difference between "light on" and "light off" was significant (Bonferonni-corrected $p < 0.05$, Wilcoxon signed rank tests for each time point), open circles are not significant. Data are presented as mean $\pm$ S.E.M.

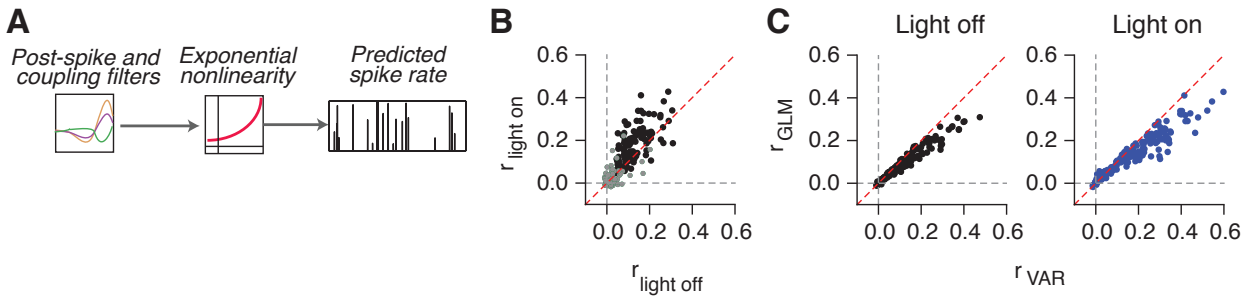

Figure 4.14: Fitting a generalized linear model shows similar results to the VAR model in Fig. 4.8. **A.** Schematic of GLM fitting. The GLM includes an exponential nonlinearity that is fit along with the post-spike and coupling filters to predict new responses on held out data. **B.** Prediction performance of exponential GLM is significantly higher when modeling spike trains during light "on" epochs compared to "off" epochs (left panel, $p < 4.3 \times 10^{-12}$, Wilcoxon signed rank test). As in Fig. 4B, we used resampling to determine which models performed significantly better than chance, indicated in black. Gray points show models that were excluded from further analysis. The red dashed line is the unity line $y = x$. **C.** GLM performance is significantly worse than the VAR model for both "light off" (black) and "light on" (blue) models ($p < 0.0001$, Wilcoxon signed rank tests). Performance was measured as the correlation between the response predicted by the GLM and the actual response from data not used to fit the model.

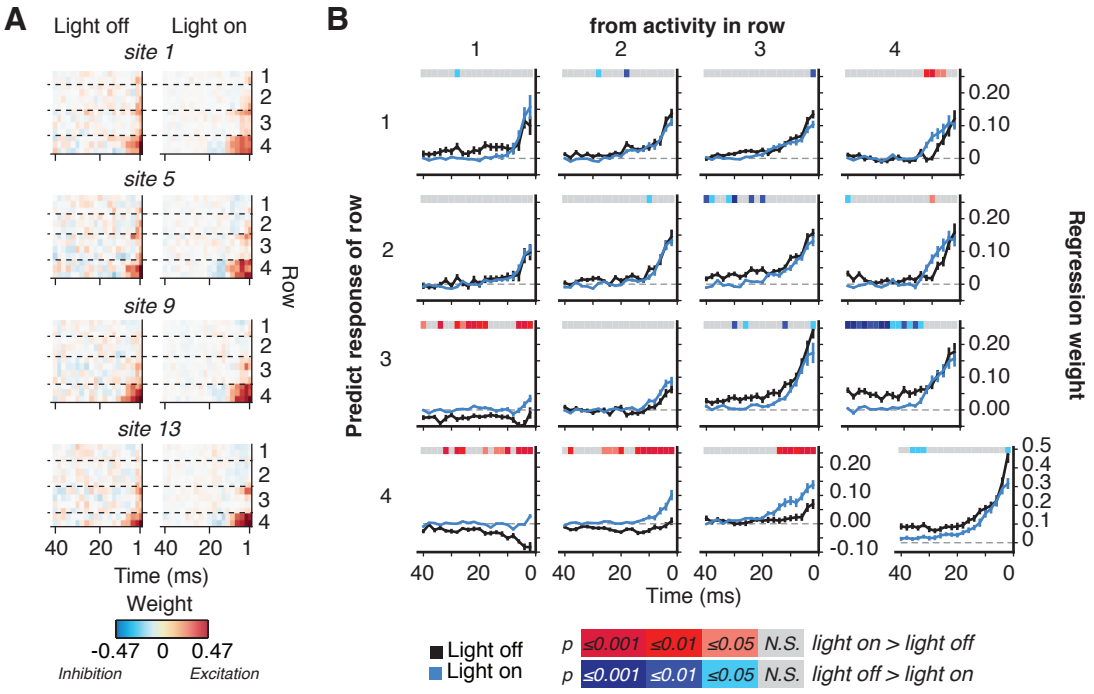

Figure 4.15: **A.** Representative GLM weights from one subject (same as Fig. 4.8C) for sites on one shank of the polytrode. Each plot shows how the activity in sites from rows 1-4 modulates the activity of sites 1, 5, 9, and 13 respectively (on the left-most shank, see Fig. 4.1B). Excitatory drive is shown in red, inhibitory drive is shown in blue. Dashed lines separate channels in different rows on the polytrode. Results are qualitatively similar to the VAR model, but with a smaller increase in excitatory drive from putative thalamic input layers during "light on" trials as well as stronger inhibition between 10 and 20 ms for "light on". **B.** Average GLM weights collapsed across all subjects and across contact sites in the same row appear qualitatively similar to VAR model weights. Data are presented as mean ± S.E.M.

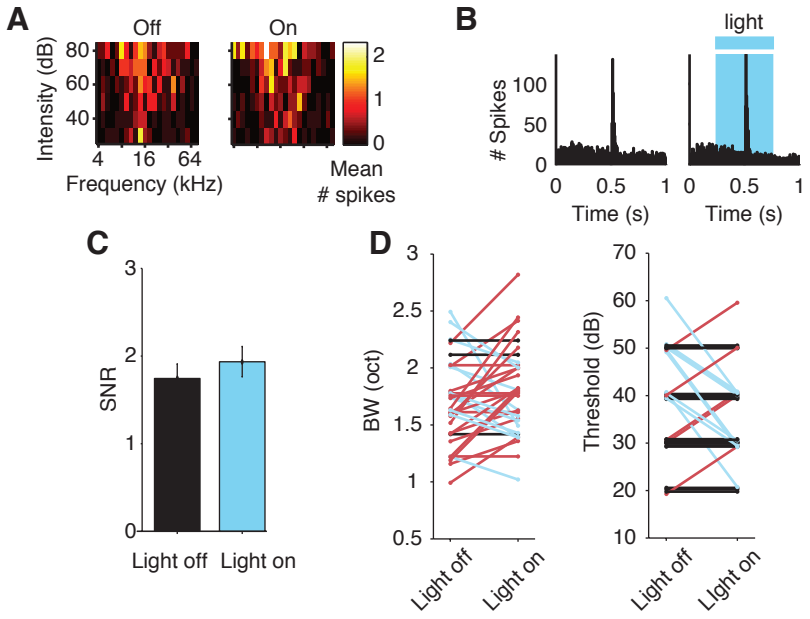

Figure 4.16: Light stimulation in the absence of ChR2 does not alter response properties, related to Fig. 4.9. **A.** Example receptive field without and with light stimulation delivered to a sham-injected control animal. **B.** Example peristimulus time histogram for "light off" (left) and "light on" (right) trials shows no significant modulation of responses by light in a representative example sham-injected control. **C.** No significant changes in signal-to-noise ratio were observed in sham-injected animals for "light off" versus "light on" trials (Wilcoxon signed rank test $p = 0.145$). Data are presented as mean $\pm$ SEM. D. No significant differences were seen for receptive field bandwidths (BW) 20 dB above threshold nor thresholds for "light off" and "light on" trials (Wilcoxon signed rank test, $p = 0.13$ and $p = 0.11$, for BW and threshold respectively).

Acknowledgments

LSH and SB contributed to the study design. LSH collected the data and performed the electrophysiological experiments. VMC and LSH performed the immunohistochemistry and histology. LSH, JSD, and AGH wrote code to fit the models and analyzed the data. KD provided the original ChR2 construct. LSH and SB wrote the manuscript. All authors discussed and commented on the manuscript. The authors would like to thank Fritz Sommer, Michael DeWeese, Daniel Feldman, David Larue, Asako Miyakawa, Robert Gibboni, Aaron Koralek, and Colleen Kirkhart for helpful comments and advice. The authors would also like to thank Seung-Hee Lee for advice on virus injection procedures and histology and Trevor Flynn for assistance in cell counting analyses. This work was supported by the National Institute on Deafness and Other Communication Disorders (DC009259), a William Orr Dingwall Neurolinguistics Dissertation Fellowship (to A.G.H), and a National Science Foundation Graduate Research Fellowship (to L.S.H.).

CHAPTER 5

Conclusion and Implications for Future Research

5.1 Summary of results

The purpose of the work described here is to elucidate the complex dynamics between incoming sensory stimuli and neuronal circuits in the auditory cortex.

In Chapter 2, I showed that natural stimuli, such as conspecific vocalizations, may be preferentially represented by the auditory cortex, and that simple organizational principles such as tonotopy may not explain responses to more complex sounds.

In Chapter 3, I showed how single-site response properties are related to functional connections between different sites in different layers and columns of the auditory cortex, and I described changes at both of these levels as a result of pulsed noise exposure during the auditory critical period for plasticity. In particular, I showed that different auditory circuits undergo experience-dependent plasticity in different temporal windows. I showed that early noise exposure preferentially affects stimulus-driven responses (presumably thalamocortical) and late noise exposure affects the spread of corticocortical connections.

In Chapter 4, I demonstrated that stimulating PV+ inhibitory interneurons enhances signal-to-noise ratio in individual cortical sites, while enhancing feedforward propagation of information flow throughout the cortex. This study is one of the first to use optogenetics with computational modeling of circuit dynamics *in vivo*, and we hope that it will provide impetus for further interrogation of cortical circuitry.

5.2 Choosing the appropriate stimulus

The results described in Chapter 2 showcase the importance of choosing appropriate stimuli to probe sensory systems. As shown here and by others, natural stimuli and non-natural stimuli can elicit different responses in the brain, and responses to one class will not necessarily predict the other (Portfors et al. 2009; Holmstrom et al. 2010; Woolley et al. 2006; Laudanski et al. 2012). Ex-

perimentalists often face a trade-off between recording responses to more different types of stimuli or to record more repeats to the same types of stimuli to obtain a more robust response, since it is obviously not possible to sample every possible auditory stimulus. The use of sounds like dynamic ripple noise was originally intended to allow a more complete sampling of the acoustic space while strongly driving cortical neurons (Kowalski et al. 1996; Depireux et al. 2001). Recently, it has been shown that even these spectrotemporally complex synthetic sounds are represented differently from natural vocalizations (Laudanski et al. 2012). Since the function of sound processing in the brain is presumably to drive behaviors, using behaviorally relevant sounds is paramount in understanding neural representations (Mizrahi et al. 2014). Still, even here our description of responses to vocalizations in isolation is somewhat "unnatural", since these vocalization sounds would naturally occur with other environmental sounds in the background. Significant advances in analyzing responses to natural stimuli (Wu et al. 2006; Calabrese et al. 2011; Ahrens et al. 2008) mean that it is no longer necessary to choose simple, synthetic sounds to drive auditory neurons. However, the choice of which stimulus to use is still a difficult one, but researchers should strive to choose the best stimulus for the question at hand.

5.3 Population dynamics can inform stimulus selectivity

In Chapters 3 and 4, I show the importance of including information about functional connections between sites when describing stimulus selectivity. Incorporating site-to-site connections significantly improves the likelihood of our models (Figs. 3.2C and 4.3), and it is perhaps not surprising that assuming neural independence goes against what we find in our data. In Chapter 4, we show in Fig. 4.11 that removing sound-evoked periods from the Ising model does not change the result of increased vertical connectivity with PV+ neuron stimulation, further highlighting the importance of site-to-site connectivity over stimulus input. Even more significantly, we find that responses to sound in superficial row sites are likely indirect and inherited from strong feedforward connections to the input layer (Fig. 3.4 and Fig. 4.2).

With the increasing use of multi-site probes and the ability to record from hundreds or more neurons at once, it is important to relate stimulus-evoked changes to underlying connectivity within the circuit. The methods described here are easily translatable to larger systems, and could provide insight into large-scale connections within and across brain areas.

5.4 Future directions

This work provides a unifying view of how circuit interactions relate to single site properties, but there are many future questions to be asked. First of all, probing a wider variety of stimuli in addition to vocalizations would provide greater insight into how sound representation is organized in the brain and what relevant parameters are extracted by neural circuits. In addition, using complex stimuli in the models described in Chapters 3 and 4 would provide insight into how complex, behaviorally-relevant information is processed by different circuits. With newer, more refined optogenetic techniques, it is now possible to record from and stimulate cells with greater spatial and

temporal precision. For example, using a polytrode where light could be delivered at each recording site separately (Chen et al. 2013; Voigts et al. 2013) would allow for more detailed dissection of the laminar and columnar circuits described in Chapter 4. Finally, determining the relevance of these perturbations in awake, behaving animals would provide insight into how increasing feedforward functional connections might influence perception. One hypothesis is that by decreasing spontaneous activity through PV+ neuron stimulation, one might reduce the influence of sensory priors on perception (Kover & Bao 2010), thus enhancing the contributions of bottom-up information rather than top-down expectations.

References

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A Learning Algorithm for Boltzmann Machines*. Cognitive Science, 9, 147–169.

Aertsen, A. M. & Johannesma, P. I. (1981). The spectro-temporal receptive field. A functional characteristic of auditory neurons. Biological Cybernetics, 42, 133–143.

Ahrens, M. B., Linden, J. F., & Sahani, M. (2008). Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. Journal of Neuroscience, 28, 1929–42.

Asari, H. & Zador, A. M. (2009). Long-lasting context dependence constrains neural encoding models in rodent auditory cortex. Journal of Neurophysiology, 102, 2638–56.

Atallah, B. V., Bruns, W., Carandini, M., & Scanziani, M. (2012). Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. Neuron, 73, 159–170.

Atencio, C. A. & Schreiner, C. E. (2008). Spectrotemporal processing differences between auditory cortical fast-spiking and regular-spiking neurons. Journal of Neuroscience, 28, 3897–910.

Attias, H. & Schreiner, C. E. (1997). Temporal low-order statistics of natural sounds. Advances in neural information processing.

Bakk, A. & Høye, J. S. (2003). One-dimensional ising model applied to protein folding. Physica A: Statistical Mechanics and its Applications, 323, 504–518.

Bao, S., Chang, E. F., Woods, J., & Merzenich, M. M. (2004). Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. Nature Neuroscience, 7, 974–981.

Barbour, D. L. & Wang, X. (2003). Contrast tuning in auditory cortex. Science (New York, N.Y.), 299, 1073–5.

Barkat, T. R., Polley, D. B., & Hensch, T. K. (2011). A critical period for auditory thalamocortical connectivity. Nature Neuroscience, 14, 1189–1194.

Beierlein, M., Gibson, J. R., & Connors, B. W. (2000). A network of electrically coupled interneurons drives synchronized inhibition in neocortex. Nature Neuroscience, 3, 904–10.

Bernstein, J. G. & Boyden, E. S. (2011). Optogenetic tools for analyzing the neural circuits of behavior. Trends in cognitive sciences, 15, 592–600.

Blake, D. T. & Merzenich, M. M. (2002). Changes of AI receptive fields with sound density. Journal of Neurophysiology, 88, 3409–20.

Blundon, J. a. & Zakharenko, S. S. (2013). Presynaptic Gating of Postsynaptic Synaptic Plasticity: A Plasticity Filter in the Adult Auditory Cortex. The Neuroscientist.

Borst, A. & Theunissen, F. E. (1999). Information theory and neural coding. Nature Neuroscience, 2, 947–957.

Boyden, E. S. (2011). A history of optogenetics: the development of tools for controlling brain circuits with light. F1000 Biology reports, 3.

Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., & Deisseroth, K. (2005). Millisecond-timescale, genetically targeted optical control of neural activity. Nature Neuroscience, 8, 1263–8.

Bryngelson, J. D. & Wolynes, P. G. (1987). Spin glasses and the statistical mechanics of protein folding. Proceedings of the National Academy of Sciences, 84, 7524–7528.

Calabrese, A., Schumacher, J. W., Schneider, D. M., Paninski, L., & Woolley, S. M. N. (2011). A generalized linear model for estimating spectrotemporal receptive fields from responses to natural sounds. PloS one, 6, e16104.

Cardin, J. A., Carlen, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L. H., & Moore, C. I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature, 459, 663–667.

Cardin, J. A., Carlen, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L. H., & Moore, C. I. (2010). Targeted optogenetic stimulation and recording of neurons in vivo using cell-type-specific expression of Channelrhodopsin-2. Nat Protoc, 5, 247–254.

Celio, M. R. (1986). Parvalbumin in most gamma-aminobutyric acid-containing neurons of the rat cerebral cortex. Science, 231, 995–997.

Chang, E. F., Bao, S., Imaizumi, K., Schreiner, C. E., & Merzenich, M. M. (2005). Development of spectral and temporal response selectivity in the auditory cortex. Proceedings of the National Academy of Sciences of the United States of America, 102, 16460–16465.

Chang, E. F. & Merzenich, M. M. (2003). Environmental noise retards auditory cortical development. Science, 300, 498–502.

Chattopadhyaya, B., Di Cristo, G., Higashiyama, H., Knott, G. W., Kuhlman, S. J., Welker, E., & Huang, Z. J. (2004). Experience and activity-dependent maturation of perisomatic GABAergic innervation in primary visual cortex during a postnatal critical period. Journal of Neuroscience, 24, 9598–611.

Chen, S., Pei, W., Gui, Q., Chen, Y., Zhao, S., Wang, H., & Chen, H. (2013). A fiber-based implantable multi-optrode array with contiguous optical and electrical sites. Journal of neural engineering, 10, 046020.

Chen, Y., Martinez-Conde, S., Macknik, S. L., Bereshpolova, Y., Swadlow, H. A., & Alonso, J. M. (2008). Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. Nature Neuroscience, 11, 974–982.

Christianson, G. B., Sahani, M., & Linden, J. F. (2011). Depth-dependent temporal response properties in core auditory cortex. Journal of Neuroscience, 31, 12837–12848.

Clopton, B. M., Winfield, J. A., & Flammino, F. J. (1974). Tonotopic organization: review and analysis. Brain research, 76, 1–20.

de Villers-Sidani, E., Chang, E. F., Bao, S., & Merzenich, M. M. (2007). Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. Journal of Neuroscience, 27, 180–189.

de Villers-Sidani, E., Simpson, K. L., Lu, Y. F., Lin, R. C., & Merzenich, M. M. (2008). Manipulating

critical period closure across different sectors of the primary auditory cortex. Nature Neuroscience, 11, 957–965.

DeAngelis, G. C., Ghose, G. M., Ohzawa, I., & Freeman, R. D. (1999). Functional micro-organization of primary visual cortex: receptive field analysis of nearby neurons. Journal of Neuroscience, 19, 4046–4064.

deCharms, C. R., Blake, D. T., & Merzenich, M. M. (1998). Optimizing sound features for cortical neurons. Science, 280, 1439–1444.

Depireux, D. A., Simon, J. Z., Klein, D. J., & Shamma, S. A. (2001). Spectro-Temporal Response Field Characterization With Dynamic Ripples in Ferret Primary Auditory Cortex. Journal of Neurophysiology, 85, 1220–1234.

Diamond, D. M. & Weinberger, N. M. (1986). Classical conditioning rapidly induces specific changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields. Brain research, 372, 357–360.

Douglas, R., Martin, K. A., & Whitteridge, D. (1989). A canonical microcircuit for neocortex. Neural Computation, 488, 480–488.

Douglas, R. J. & Martin, K. A. (1991). A functional microcircuit for cat visual cortex. J Physiol, 440, 735–769.

Douglas, R. J. & Martin, K. A. (2007). Mapping the matrix: the ways of neocortex. Neuron, 56, 226–238.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. The annals of Statistics, pp. 1–26.

Eggermont, J., Johannesma, P., & Aertsen, A. (1983). Reverse-correlation methods in auditory research. Quarterly reviews of biophysics, 16, 341–414.

Eggermont, J. J. & Kenmochi, M. (1998). Salicylate and quinine selectively increase spontaneous firing rates in secondary auditory cortex. Hearing Research, 117, 149–160.

Ehret, G. (1992). Categorical perception of mouse-pup ultrasounds in the temporal domain. Animal Behaviour, 43, 409–416.

Ehret, G. & Haack, B. (1981). Categorical perception of mouse pup ultrasound by lactating females. Naturwissenschaften, 68, 208–209.

Elhilali, M., Fritz, J. B., Bozak, D., Depireux, D. A., Simon, J. Z., Klein, D. J., & Shamma, S. A. (2002). Comparison of response characteristics in auditory cortex of the awake and anesthetized ferret. Tech. rep., DTIC Document.

Elhilali, M., Fritz, J. B., Chi, T.-S., & Shamma, S. A. (2007). Auditory cortical receptive fields: stable entities with plastic abilities. Journal of Neuroscience, 27, 10372–10382.

Escabi, M. a. & Schreiner, C. E. (2002). Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. Journal of Neuroscience, 22, 4114–31.

Feldman, D. E., Nicoll, R. a., & Malenka, R. C. (1999). Synaptic plasticity at thalamocortical synapses in developing rat somatosensory cortex: LTP, LTD, and silent synapses. Journal of Neurobiology, 41, 92–101.

Fenno, L., Yizhar, O., & Deisseroth, K. (2011). The development and application of optogenetics. Annual review of neuroscience, 34, 389–412.

Fox, K. (2002). Anatomical pathways and molecular mechanisms for plasticity in the barrel cortex.

Neuroscience, 111, 799–814.

Franklin, K. B. J. & Paxinos, G. (2008). The mouse brain in stereotaxic coordinates. (Amsterdam ;: Boston : Elsevier/Academic Press), 3rd edn.

Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nature Neuroscience, 6, 1216–1223.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? Hearing Research, 229, 186–203.

Fritz, J. B., Elhilali, M., & Shamma, S. A. (2005). Differential dynamic plasticity of a1 receptive fields during multiple spectral tasks. Journal of Neuroscience, 25, 7623–7635.

Fuchs, E. C., Zivkovic, A. R., Cunningham, M. O., Middleton, S., Lebeau, F. E., Bannerman, D. M., Rozov, A., Whittington, M. A., Traub, R. D., Rawlins, J. N., & Monyer, H. (2007). Recruitment of parvalbumin-positive interneurons determines hippocampal function and associated behavior. Neuron, 53, 591–604.

Gandal, M. J., Nesbitt, A. M., McCurdy, R. M., & Alter, M. D. (2012). Measuring the maturity of the fast-spiking interneuron transcriptional program in autism, schizophrenia, and bipolar disorder. PloS one, 7, e41215.

Ganmor, E., Segev, R., & Schneidman, E. (2011a). Sparse low-order interaction network underlies a highly correlated and learnable neural population code. Proceedings of the National Academy of Sciences of the United States of America, 108, 9679–9684.

Ganmor, E., Segev, R., & Schneidman, E. (2011b). The architecture of functional interaction networks in the retina. Journal of Neuroscience, 31, 3044–3054.

Gehr, D. D., Komiya, H., & Eggermont, J. J. (2000). Neuronal responses in cat primary auditory cortex to natural and altered species-specific calls. Hearing Research, 150, 27–42.

Gill, P., Woolley, S. M. N., Fremouw, T., & Theunissen, F. E. (2008). What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. Journal of Neurophysiology, 99, 2809–20.

Glass, I. & Wollberg, Z. (1983). Responses of cells in the auditory cortex of awake squirrel monkeys to normal and reversed species-specific vocalizations. Hearing Research, 9, 27–33.

Gonzalez-Burgos, G. & Lewis, D. a. (2012). NMDA receptor hypofunction, parvalbumin-positive neurons, and cortical gamma oscillations in schizophrenia. Schizophrenia bulletin, 38, 950–7.

Gray, C. M., Konig, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature, 338, 334–337.

Guo, F., Zhang, J., Zhu, X., Cai, R., Zhou, X., & Sun, X. (2012). Auditory discrimination training rescues developmentally degraded directional selectivity and restores mature expression of GABA(A) and AMPA receptor subunits in rat auditory cortex. Behavioural brain research, 229, 301–7.

Hackett, T. A., Barkat, T. R., O'Brien, B. M., Hensch, T. K., & Polley, D. B. (2011). Linking topography to tonotopy in the mouse auditory thalamocortical circuit. Journal of Neuroscience, 31, 2983–2995.

Hahn, M. E. & Lavooy, M. J. (2005). A review of the methods of studies on infant ultrasound production and maternal retrieval in small rodents. Behavior Genetics, 35, 31–52.

Hamilton, L. S., Sohl-Dickstein, J., Huth, A. G., Carels, V. M., Deisseroth, K., & Bao, S. (2013). Optogenetic Activation of an Inhibitory Network Enhances Feedforward Functional Connectivity in Auditory Cortex. Neuron, 80, 1066–1076.

Han, Y. K., Kover, H., Insanally, M. N., Semerdjian, J. H., & Bao, S. (2007). Early experience impairs perceptual discrimination. Nature Neuroscience, 10, 1191–1197.

Happel, M. F. K., Jeschke, M., & Ohl, F. W. (2010). Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input. Journal of Neuroscience, 30, 11114–27.

Hauber, M. E., Cassey, P., Woolley, S. M., & Theunissen, F. E. (2007). Neurophysiological response selectivity for conspecific songs over synthetic sounds in the auditory forebrain of non-singing female songbirds. Journal of Comparative Physiology A, 193, 765–774.

Henry, K. & McGinn, M. (1992). The mouse as a model for human audition. International Journal of Audiology, 31, 181–189.

Hensch, T. K. (2005). Critical period plasticity in local cortical circuits. Nature Reviews Neuroscience, 6, 877–888.

Hensch, T. K. & Fagiolini, M. (2005). Excitatory-inhibitory balance and critical period plasticity in developing visual cortex. Progress in brain research, 147, 115–24.

Holmstrom, L. A., Eeuwes, L. B., Roberts, P. D., & Portfors, C. V. (2010). Efficient encoding of vocalizations in the auditory midbrain. Journal of Neuroscience, 30, 802–819.

Holy, T. E. & Guo, Z. (2005). Ultrasonic songs of male mice. PLoS biology, 3, e386.

Hromadka, T., Deweese, M. R., & Zador, A. M. (2008). Sparse representation of sounds in the unanesthetized auditory cortex. PLoS Biol, 6, e16.

Hsu, A., Borst, A., & Theunissen, F. E. (2004). Quantifying variability in neural responses and its application for the validation of model predictions. Network: Computation in Neural Systems, 15, 91–109.

Insanally, M. N., Albanna, B. F., & Bao, S. (2010). Pulsed noise experience disrupts complex sound representations. Journal of Neurophysiology, 103, 2611–2617.

Insanally, M. N., Kover, H., Kim, H., & Bao, S. (2009). Feature-dependent sensitive periods in the development of complex sound representation. Journal of Neuroscience, 29, 5456–5462.

Isaacson, J. S. & Scanziani, M. (2011). How inhibition shapes cortical activity. Neuron, 72, 231–243.

Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. Zeitschrift für Physik A Hadrons and Nuclei, pp. 253–258.

Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of english as a second language. Cognitive Psychology, 21, 60–99.

Kaas, J. H. (1997). Topographic maps are fundamental to sensory processing. Brain research bulletin, 44, 107–12.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). Principles of neural science. (New York: McGraw-Hill, Health Professions Division), 4th edn.

Kao, M. C., Poon, P. W., & Sun, X. (1997). Modeling of the response of midbrain auditory neurons in the rat to their vocalization sounds based on FM sensitivities. Bio Systems, 40, 103–9.

Kaur, S., Lazar, R., & Metherate, R. (2004). Intracortical pathways determine breadth of subthresh-

old frequency receptive fields in primary auditory cortex. Journal of Neurophysiology, 91, 2551–67.

Kawaguchi, Y. & Kubota, Y. (1998). Neurochemical features and synaptic connections of large physiologically-identified GABAergic cells in the rat frontal cortex. Neuroscience, 85, 677–701.

Kilgard, M. P. & Merzenich, M. M. (1999). Distributed representation of spectral and temporal information in rat primary auditory cortex. Hearing Research, 134, 16–28.

Kilgard, M. P., Pandya, P. K., Vazquez, J., Gehi, a., Schreiner, C. E., & Merzenich, M. M. (2001). Sensory input directs spatial and temporal plasticity in primary auditory cortex. Journal of Neurophysiology, 86, 326–38.

Kim, H. & Bao, S. (2009). Selective increase in representations of sounds repeated at an ethological rate. Journal of Neuroscience, 29, 5163–5169.

Kim, H., Gibboni, R., Kirkhart, C., & Bao, S. (2013). Impaired critical period plasticity in primary auditory cortex of fragile x model mice. Journal of Neuroscience, 33, 15686–15692.

Kisley, M. a. & Gerstein, G. L. (1999). Trial-to-trial variability and state-dependent modulation of auditory-evoked responses in cortex. Journal of Neuroscience, 19, 10451–60.

Klemm, K., Eguíluz, V. M., Toral, R., & San Miguel, M. (2003). Nonequilibrium transitions in complex networks: A model of social interaction. Physical Review E, 67, 026120.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. International Joint Conference on Artificial Intelligence (IJCAI), 14, 1137–1145.

Kohonen, T. (1990). The Self-Organizing Map. Proceedings of the IEEE, 78, 1464–1480.

Kording, K. P. & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. Nature, 427, 244–247.

Köster, U., Sohl-Dickstein, J., Gray, C. M., & Olshausen, B. A. (2013). Higher order correlations within cortical layers dominate functional connectivity in microcolumns. arXiv preprint arXiv:1301.0050.

Kover, H. & Bao, S. (2010). Cortical plasticity as a mechanism for storing bayesian priors in sensory perception. PLoS One, 5.

Köver, H., Gill, K., Tseng, Y.-T. L., & Bao, S. (2013). Perceptual and neuronal boundary learned from higher-order stimulus probabilities. Journal of Neuroscience, 33, 3699–3705.

Kowalski, N., Depireux, D. A., & Shamma, S. A. (1996). Analysis of dynamic spectra in ferret primary auditory cortex. i. characteristics of single-unit responses to moving ripple spectra analysis of dynamic spectra in ferret primary auditory cortex. Journal of Neurophysiology, 76, 3503–3523.

Laudanski, J., Edeline, J.-M., & Huetz, C. (2012). Differences between spectro-temporal receptive fields derived from artificial and natural stimuli in the auditory cortex. PloS one, 7, e50539.

Lee, C. C. & Winer, J. A. (2008). Connections of cat auditory cortex: III. Corticocortical system. J Comp Neurol, 507, 1920–1943.

Lee, S. H., Kwan, A. C., Zhang, S., Phoumthipphavong, V., Flannery, J. G., Masmanidis, S. C., Taniguchi, H., Huang, Z. J., Zhang, F., Boyden, E. S., Deisseroth, K., & Dan, Y. (2012). Activation of specific interneurons improves V1 feature selectivity and visual perception. Nature, 488, 379–383.

Lenz, W. (1920). Beitrag zum verständnis der magnetischen erscheinungen in festen körpern. Physikalische Zeitschrift, 21, 2.

Lewicki, M. S. (2002). Efficient coding of natural sounds. Nature Neuroscience, 5, 356–363.

Li, L.-y., Li, Y.-t., Zhou, M., Tao, H. W., & Zhang, L. I. (2013). Intracortical multiplication of thalamocortical signals in mouse auditory cortex. Nature Neuroscience, pp. 1–5.

Linden, J. F. & Schreiner, C. E. (2003). Columnar transformations in auditory cortex? A comparison to visual and somatosensory cortices. Cerebral Cortex, 13, 83–89.

Liu, B.-h., Wu, G. K., Arbuckle, R., Tao, H. W., & Zhang, L. I. (2007). Defining cortical frequency tuning with recurrent excitatory circuitry. Nature Neuroscience, 10, 1594–600.

Liu, R. C., Miller, K. D., Merzenich, M. M., & Schreiner, C. E. (2003). Acoustic variability and distinguishability among mouse ultrasound vocalizations. The Journal of the Acoustical Society of America, 114, 3412.

Liu, X., Basavaraj, S., Krishnan, R., & Yan, J. (2011). Contributions of the thalamocortical system towards sound-specific auditory plasticity. Neuroscience and biobehavioral reviews, 35, 2155–61.

Lütkepohl, H. (2005). New Introduction to Multiple Time Series Analysis.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. Nature Neuroscience, 9, 1432–1438.

Machens, C. K., Wehr, M., & Zador, A. M. (2002). Spectro-temporal receptive fields of subthreshold responses in auditory cortex. In Proceedings of the 16th Annual Conference on Neural Information Processing Systems (NIPS 2002), vol. 5.

Machens, C. K., Wehr, M., & Zador, A. M. (2003). Spectro-temporal receptive fields of subthreshold responses in auditory cortex. (The MIT Press).

Machens, C. K., Wehr, M. S., & Zador, A. M. (2004). Linearity of cortical receptive fields measured with natural sounds. Journal of Neuroscience, 24, 1089–100.

Marre, O., El Boustani, S., Fregnac, Y., & Destexhe, A. (2009). Prediction of spatiotemporal patterns of neural activity from pairwise correlations. Phys Rev Lett, 102, 138101.

McIntosh, A. R. & Gonzalez-Lima, F. (1998). Large-scale functional connectivity in associative learning: interrelations of the rat auditory, visual, and limbic systems. Journal of Neurophysiology, 80, 3148–3162.

Mitchell, J. F., Sundberg, K. A., & Reynolds, J. H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. Neuron, 55, 131–141.

Mizrahi, A., Shalev, A., & Nelken, I. (2014). Single neuron and population coding of natural sounds in auditory cortex. Current Opinion in Neurobiology, 24, 103–110.

Moore, A. K. & Wehr, M. (2013). Parvalbumin-expressing inhibitory interneurons in auditory cortex are well-tuned for frequency. Journal of Neuroscience, 33, 13713–13723.

Mountcastle, V. B. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. Journal of Neurophysiology, 20, 408–434.

Müller, M. (1991). Frequency representation in the rat cochlea. Hearing Research, 51, 247–254.

Nagel, G., Szellas, T., Huhn, W., Kateriya, S., Adeishvili, N., Berthold, P., Ollig, D., Hegemann, P., & Bamberg, E. (2003). Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. Proceedings of the National Academy of Sciences, 100, 13940–13945.

Noreña, A. & Eggermont, J. J. (2002). Comparison between local field potentials and unit cluster activity in primary auditory cortex and anterior auditory field in the cat. Hearing Research, 166, 202–213.

Oberlaender, M., Ramirez, A., & Bruno, R. M. (2012). Sensory experience restructures thalamocortical axons during adulthood. Neuron, 74, 648–55.

Ohiorhenuan, I. E., Mechler, F., Purpura, K. P., Schmid, A. M., Hu, Q., & Victor, J. D. (2010). Sparse coding and high-order correlations in fine-scale cortical networks. Nature, 466, 617–621.

Olshausen, B. A. & Field, D. J. (2004). Sparse coding of sensory inputs. Curr Opin Neurobiol, 14, 481–487.

Oviedo, H. V., Bureau, I., Svoboda, K., & Zador, A. M. (2010). The functional asymmetry of auditory cortex is reflected in the organization of local cortical circuits. Nature Neuroscience, 13, 1413–1420.

Packer, A. M. & Yuste, R. (2011). Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? Journal of Neuroscience, 31, 13260–13271.

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., & Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature, 454, 995–9.

Polley, D. B., Read, H. L., Storace, D. a., & Merzenich, M. M. (2007). Multiparametric auditory receptive field organization across five cortical fields in the albino rat. Journal of Neurophysiology, 97, 3621–38.

Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. Journal of Neuroscience, 26, 4970–82.

Popescu, M. V. & Polley, D. B. (2010). Monaural deprivation disrupts development of binaural selectivity in auditory midbrain and cortex. Neuron, 65, 718–731.

Portfors, C. V., Roberts, P. D., & Jonson, K. (2009). Over-representation of species-specific vocalizations in the awake mouse inferior colliculus. Neuroscience, 162, 486–500.

Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. Science, 268, 111–114.

Recanzone, G. H. (2008). Representation of con-specific vocalizations in the core and belt areas of the auditory cortex in the alert macaque monkey. Journal of Neuroscience, 28, 13184–13193.

Roudi, Y., Nirenberg, S., & Latham, P. E. (2009a). Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. PLoS Comput Biol, 5, e1000380.

Roudi, Y., Tyrcha, J., & Hertz, J. (2009b). Ising model for neural data: model quality and approximate methods for extracting functional connectivity. Phys Rev E Stat Nonlin Soft Matter Phys, 79, 51915.

Sakata, S. & Harris, K. D. (2009). Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex. Neuron, 64, 404–418.

Sauer, B. (1998). Inducible gene targeting in mice using the cre/lox system. Methods, 14, 381–392.

Schaub, M. T. & Schultz, S. R. (2012). The Ising decoder: reading out the activity of large neural ensembles. J Comput Neurosci, 32, 101–118.

Schneider, D. M. & Woolley, S. M. (2011). Extra-classical tuning predicts stimulus-dependent receptive fields in auditory neurons. Journal of Neuroscience, 31, 11867–11878.

Schneidman, E., Berry 2nd, M. J., Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply

strongly correlated network states in a neural population. Nature, 440, 1007–1012.

Schnupp, J. W., Hall, T. M., Kokelaar, R. F., & Ahmed, B. (2006). Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. Journal of Neuroscience, 26, 4785–4795.

Schreiner, C. E. & Sutter, M. L. (1992). Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. Journal of Neurophysiology, 68, 1487–502.

Schreiner, C. E. & Winer, J. A. (2007). Auditory cortex mapmaking: principles, projections, and plasticity. Neuron, 56, 356–365.

Schumacher, J. W., Schneider, D. M., & Woolley, S. M. N. (2011). Anesthetic state modulates excitability but not spectral tuning or neural discrimination in single auditory midbrain neurons. Journal of Neurophysiology.

Seyfarth, R. M. & Cheney, D. L. (2010). Production, usage, and comprehension in animal vocalizations. Brain and language, 115, 92–100.

Sharma, J., Angelucci, A., & Sur, M. (2000). Induction of visual orientation modules in auditory cortex. Nature, 404, 841–847.

Shlens, J., Field, G. D., Gauthier, J. L., Greschner, M., Sher, A., Litke, A. M., & Chichilnisky, E. J. (2009). The structure of large-scale synchronized firing in primate retina. Journal of Neuroscience, 29, 5022–5031.

Shlens, J., Field, G. D., Gauthier, J. L., Grivich, M. I., Petrusca, D., Sher, A., Litke, A. M., & Chichilnisky, E. J. (2006). The structure of multi-neuron firing patterns in primate retina. Journal of Neuroscience, 26, 8254–8266.

Singh, N. C. & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. The Journal of the Acoustical Society of America, 114, 3394.

Smith, E. C. & Lewicki, M. S. (2006). Efficient auditory coding. Nature, 439, 978–982.

Sohal, V. S., Zhang, F., Yizhar, O., & Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nature, 459, 698–702.

Sohl-Dickstein, J., Battaglino, P., & DeWeese, M. (2011). New Method for Parameter Estimation in Probabilistic Models: Minimum Probability Flow. Physical Review Letters, 107, 11–14.

Speechley, W. J., Hogsden, J. L., & Dringenberg, H. C. (2007). Continuous white noise exposure during and after auditory critical period differentially alters bidirectional thalamocortical plasticity in rat auditory cortex in vivo. The European Journal of Neuroscience, 26, 2576–84.

Stiebler, I. & Ehret, G. (1985). Inferior colliculus of the house mouse. i. a quantitative study of tonotopic organization, frequency representation, and tone-threshold distribution. Journal of Comparative Neurology, 238, 65–76.

Stiebler, I., Neulist, R., Fichtel, I., & Ehret, G. (1997). The auditory cortex of the house mouse: left-right differences, tonotopic organization and quantitative analysis of frequency representation. J Comp Physiol A, 181, 559–571.

Stocker, A. A. & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. Nature Neuroscience, 9, 578–585.

Sugimoto, S., Sakurada, M., Horikawa, J., & Taniguchi, I. (1997). The columnar and layer-specific response properties of neurons in the primary auditory cortex of Mongolian gerbils. Hearing Research, 112, 175–185.

Šuta, D., Kvašňák, E., Popelář, J., & Syka, J. (2003). Representation of species-specific vocalizations in the inferior colliculus of the guinea pig. Journal of Neurophysiology, 90, 3794–3808.

Šuta, D., Popelár, J., Kvasnák, E., & Syka, J. (2007). Representation of species-specific vocalizations in the medial geniculate body of the guinea pig. Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale, 183, 377–88.

Syka, J., Suta, D., & Popelár, J. (2005). Responses to species-specific vocalizations in the auditory cortex of awake and anesthetized guinea pigs. Hearing Research, 206, 177–84.

Tang, A., Jackson, D., Hobbs, J., Chen, W., Smith, J. L., Patel, H., Prieto, A., Petrusca, D., Grivich, M. I., Sher, A., Hottowy, P., Dabrowski, W., Litke, A. M., & Beggs, J. M. (2008). A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. Journal of Neuroscience, 28, 505–518.

Theunissen, F. E., David, S. V., Singh, N. C., Hsu, A., Vinje, W. E., & Gallant, J. L. (2001). Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Network, 12, 289–316.

Theunissen, F. E. & Shaevitz, S. S. (2006). Auditory processing of vocal sounds in birds. Current Opinion in Neurobiology, 16, 400–7.

Toledo-Rodriguez, M., Blumenfeld, B., Wu, C., Luo, J., Attali, B., Goodman, P., & Markram, H. (2004). Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. Cerebral Cortex, 14, 1310–1327.

Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination tasks by rats. Animal Cognition, 6, 131–136.

Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2005). Effects of backward speech and speaker variability in language discrimination by rats. Journal of Experimental Psychology: Animal Behavior Processes, 31, 95.

Tropea, D., Van Wart, A., & Sur, M. (2009). Molecular mechanisms of experience-dependent plasticity in visual cortex. Philosophical Transactions of the Royal Society B: Biological Sciences, 364, 341–355.

Voigts, J., Siegle, J. H., Pritchett, D. L., & Moore, C. I. (2013). The flexdrive: an ultra-light implant for optical control and highly parallel chronic recording of neuronal ensembles in freely moving mice. Frontiers in systems neuroscience, 7.

Wallace, M. & Palmer, A. (2008). Laminar differences in the response properties of cells in the primary auditory cortex. Experimental Brain Research, 184, 179–191.

Wang, X. & Kadia, S. C. (2001). Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. Journal of Neurophysiology, 86, 2616–2620.

Wang, X., Lu, T., Snider, R. K., & Liang, L. (2005). Sustained firing in auditory cortex evoked by preferred stimuli. Nature, 435, 341–6.

Wang, X., Merzenich, M. M., Beitel, R., & Schreiner, C. E. (1995). Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. Journal of Neurophysiology, 74, 2685–706.

Watkins, P. V. & Barbour, D. L. (2008). Specialized neuronal adaptation for preserving input sensitivity. Nature Neuroscience, 11, 1259–61.

Watkins, P. V. & Barbour, D. L. (2011). Level-tuned neurons in primary auditory cortex adapt

differently to loud versus soft sounds. Cerebral cortex (New York, N.Y. : 1991), 21, 178–90.

Wilson, N. R., Runyan, C. A., Wang, F. L., & Sur, M. (2012). Division and subtraction by distinct cortical inhibitory networks in vivo. Nature, 488, 343–348.

Winer, J. A. & Lee, C. C. (2007). The distributed auditory cortex. Hearing Research, 229, 3–13.

Winer, J. A. & Schreiner, C. E. (2011). The Auditory Cortex. (New York: Springer).

Womelsdorf, T., Schoffelen, J. M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., & Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. Science, 316, 1609–1612.

Woolley, S., Hauber, M. E., & Theunissen, F. E. (2010). Developmental experience alters information coding in auditory midbrain and forebrain neurons. Developmental Neurobiology, 70, 235–252.

Woolley, S. M. N., Gill, P. R., & Theunissen, F. E. (2006). Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. Journal of Neuroscience, 26, 2499–512.

Wu, M. C., David, S. V., & Gallant, J. L. (2006). Complete functional characterization of sensory neurons by system identification. Annu Rev Neurosci, 29, 477–505.

Yu, X., Chung, S., Chen, D.-Y., Wang, S., Dodd, S. J., Walters, J. R., Isaac, J. T. R., & Koretsky, A. P. (2012). Thalamocortical inputs show post-critical-period plasticity. Neuron, 74, 731–42.

Zhang, F., Gradinaru, V., Adamantidis, A. R., Durand, R., Airan, R. D., de Lecea, L., & Deisseroth, K. (2010). Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures. Nature Protocols, 5, 439–456.

Zhang, L. I., Bao, S., & Merzenich, M. M. (2001). Persistent and specific influences of early acoustic environments on primary auditory cortex. Nature Neuroscience, 4, 1123–30.

Zhang, L. I., Bao, S., & Merzenich, M. M. (2002). Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. Proceedings of the National Academy of Sciences of the United States of America, 99, 2309–14.

Zhao, S., Ting, J. T., Atallah, H. E., Qiu, L., Tan, J., Gloss, B., Augustine, G. J., Deisseroth, K., Luo, M., Graybiel, A. M., & Feng, G. (2011). Cell type-specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. Nat Methods, 8, 745–752.

Zhou, X. & Merzenich, M. M. (2008). Enduring effects of early structured noise exposure on temporal modulation in the primary auditory cortex. Proceedings of the National Academy of Sciences of the United States of America, 105, 4423–8.

Zhou, X. & Merzenich, M. M. (2009). Developmentally degraded cortical temporal processing restored by training. Nature Neuroscience, 12, 26–8.

Zucker, R. S. & Regehr, W. G. (2002). Short-term synaptic plasticity. Annu Rev Physiol, 64, 355–405.