INVESTIGATING INFORMATION CODING IN THE AUDITORY MIDBRAIN

PhD thesis

Domonkos Áron Horváth

János Szentágothai PhD School of Neurosciences

Semmelweis University

Supervisor:	István Ulbert, D.Sc
Reviewers:	László Négyessy, Ph.D. Bence Rácz, Ph.D, habil.

Head of the Complex Examination Committee: Zoltán Benyó, D.Sc

Members of the Complex Examination Committee: József Pucsok, D.Sc.
                                                                            Gergely Szabó, Ph.D.

Budapest
2023

Table of contents

List of abbreviations

4-AP	4-aminopyridine
A/D converter	Analog-to-digital converter
AAF	anterior auditory field
AC	Auditory cortex
AI	Primary auditory cortex
AVCN	Anteroventral cochlear nucleus
CF	Characteristic frequency
CN	Cochlear nucleus
DCN	Dorsal cochlear nucleus
DNLL	Dorsal nucleus of the lateral lemniscus
FRA	Frequency response area
GABA	Gamma-aminobutyric acid
IC	Inferior colliculus
ICC	Central nucleus of the inferior colliculus
ICDC	Dorsal cortex of the inferior colliculus
ICX	External cortex of the inferior colliculus
ILD	Interaural level difference
ITD	Interaural time difference
LD	Lateral division of the central nucleus of the inferior colliculus
LLN	Lateral lemniscal nuclei
LSO	Lateral superior olive

MGB	Medial geniculate body
MNTB	Medial nucleus of the trapezoid body
MSO	Medial superior olive
PSTH	Post-stimulus time histogram
PVCN	Posteroventral cochlear nucleus
RLF	Rate-level function
SC	Superior colliculus
SNR	Signal-to-noise ratio
SO	Superior olive
SOC	Superior olivary complex
SPL	Sound pressure level
SSA	Stimulus specific adaptation
TEA	Tetraethylammonium
VNLL	Ventral nucleus of the lateral lemniscus

1. Introduction

1.1 Spatial hearing

Detection of events in space is essential for any creature. Interestingly, the notion of space cannot be encoded without using an additional dimension: time (Boring Edwin G., 1942; Jeffress, 1948; Wallach, 1940). In the avian and mammalian brain, localization of sound direction is based on the added dimension time. The location of a sound source is inferred by the brain using the differences in the neuronal signals transmitted from the two ears upon sensation of a sound. The arrival time of sound in the inner ear depends on the horizontal location of the sound source related to the ears. If a sound source is not exactly in front or behind the animal, sound will reach on of the two ears with a time delay. The temporal evolution of such time-of-arrival differences can be used to localize the sound source.

The first model for the neuronal mechanism of detecting interaural time differences (ITDs) was proposed by Jeffress in 1948 (Jeffress, 1948). The Jeffress model proposes an architecture consisting of four main elements in the auditory nervous system. First, the architecture takes temporally coded neural signals as input. These neural signals are generated in the cochlea and cochlear nuclei. The spikes generated by these structures are time-locked to the waveform of the acoustic stimulus. According to the Jeffress model, the time difference is detected by a series of neurons in the brainstem that receive excitatory input from both ears. These neurons fire action potentials, when the excitation from the two ears coincides on that neuron. Each neuron receives input from both ears but the lengths of the axons making synapse on the neuron are different. The different axonal lengths are distributed in a way that action potentials arriving from the two ears coincide when there is a difference between the arrival time of sound in the two ears. By having different lengths of the innervating axons on each neuron, the system is capable of detecting the sound arrival time delays experienced by the animal. With this structure, a series of neurons would sit in the brainstem enabling the detection of all possible time differences (Figure 1.1 A).
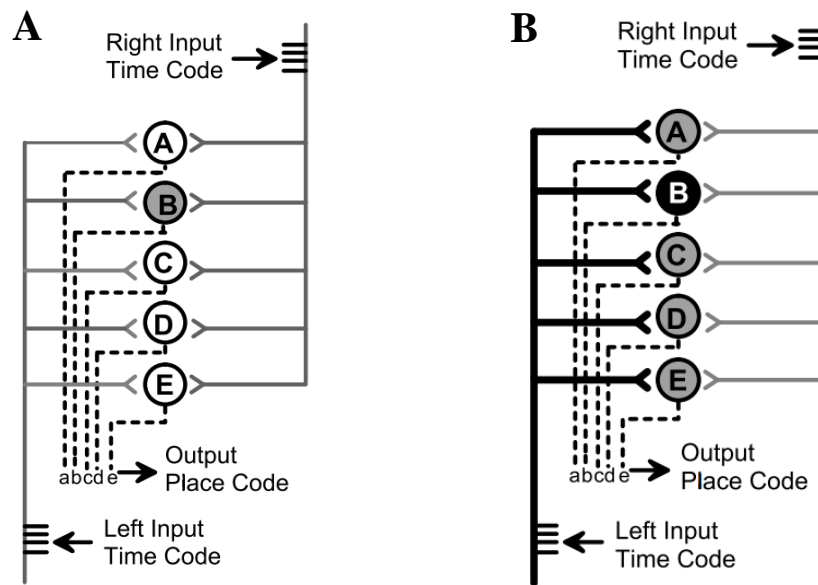
7

Figure 1.1. The original and the modified Jeffress model for detecting time delays in two ears. A: The Jeffress model for detecting time delays in the two ears. The arrangement of delay lines allows for a range of available relative delays from the two monoaural input lines that are mapped to particular places in the coincidence detector array. B: The modified Jeffress model with inhibition from ipsilateral ear. Gray lines: excitation, black lines: inhibition. (Cariani, 2011; Palmer & Kuwada, 2005)

Later, the model was extended with inhibitory inputs (Palmer & Kuwada, 2005) by finding that while contralateral inputs excite the coincidence detecting neurons, ipsilateral inputs inhibit them. Coinciding excitation and inhibition prevent the neuron from firing. Additionally, maximal neuronal firing is still enabled by the differences in the input axonal lengths from the two ears to the coincidence detectors in the brainstem (Figure 1.1 B). Anatomical evidence for the modified Jeffress model has been shown in birds (Carr & Konishi, 1990; McAlpine & Grothe, 2003) but has not yet been fully established in mammals. Also, the original and the modified Jeffress models do not account for the frequency dependence of the interaural time difference selectivity of the auditory nervous system. According to McAlpine and colleagues (McAlpine et al., 2001), neurons with lower best frequencies prefer longer interaural delays, while higher best frequency neurons prefer shorter interaural delays.

Arrival time is not the only feature of the sound that can differ in the two ears. The head size of the animal limits the detectable time delays. That is, the head size determines the upper frequency limit for detectable ITDs in a way that the smaller the head of the animal

8

is, the higher frequency ITDs are still detectable. For example, for humans the highest frequency, where ITDs are detectable, is around 1400 Hz (Brughera et al., 2013; Hartmann & Macaulay, 2014; Zwislocki & Feldman, 1956), while it is around 2800 Hz for chinchillas (Hartmann & Macaulay, 2014; Heffner et al., 1994). For frequency ranges, where ITDs are not available, interaural level differences (ILD) can be used. In addition, spectral content of the sound stimulus can also help in sound source localization. Most of the animals use all three strategies for sound source localization depending on stimulus features (Palmer & Kuwada, 2005). This gives rise to the interesting question of how the coding of ITD and ILD changes in the brainstem for single neurons and also for neuronal populations by changing time difference or level difference. The interesting aspect of this question is how the two-types of sound source information channels influence the coding precision on the level of neurons: do the changes of the two clues have the same effect on neuronal coding or do they change different aspects of neuronal response? The importance of this question is that it can shed light on the strategies the central nervous system uses to keep sound information encoding robust in the face of changing sound features, for example when the position of the animal related to the sound source is changing.

To address this question, I performed single and multiunit recordings in the inferior colliculus of anesthetized gerbils. The animals received acoustic stimuli with different time and level differences in the two ears. Responses of single neurons were then decoded using distance metrics (Victor & Purpura, 1996) Analysis of the data revealed that changing time and level difference have different effect on neuronal coding. While changes in sound intensity evoked changes in spectrotemporal filtering that influenced the overall timing of spike events but preserved their precision across trials such that the decoding of single neuron responses was not affected. In contrast, changes in interaural time difference did not trigger changes in spectrotemporal filtering but did have strong effects on the precision of spike events and, consequently, on decoder performance. However, changes in ITD had opposing effects in the two brain hemispheres and, thus, canceled out at the population level. These results were similar with and without the addition of background noise (Horvath & Lesica, 2011).

For the study about the effects of changes in interaural time and level differences on single neuronal coding, decoding methods were used (Victor & Purpura, 1996). When more single neuronal responses are recorded from the same population, population response can be characterized by the population coding performance. To obtain this measure, population decoding should be used. Thus, the next question is which the optimal decoding strategy for population responses is. The other published paper included in this thesis is about population decoder optimization.

Finding the coding performance of each individual neuron in the population is first needed for the decoder optimization. Then, for population-level decoder optimization, each neuron has to be assigned a weight that determines their contribution to the decoder. However, finding the optimal weight for each neuron is not trivial. I contributed to the development of a genetic optimization method for decoding neuronal population responses. I carried out single and multiunit recordings in the inferior colliculus of gerbils and took part in the decoding analysis. This method was also applied for data obtained from two-photon recordings in the mouse visual cortex. It was shown in both cases that genetic optimization is able to provide a superior distance metrics decoder performance to neuronal weight optimization using randomly assigned weights or weights obtained decoding each neuron individually. (Hofer et al., 2010)

In this thesis, I first summarize the anatomical and physiological properties of the inferior colliculus. Then, in the following sections, I describe my methods and results first about genetic optimization of a decoder using distance metrics, then about the time and level difference coding in the inferior colliculus. The methods and results were published in two papers.

1.2. The ascending auditory pathway

1.2.1 Auditory nuclei in the brainstem

The ascending auditory pathway originates from the cochlea and terminates in the auditory cortex, transmitting auditory information from the ear through several subcortical nuclei to the cerebral cortex. Cranial nerve VIII connects the output of cochlear hair cells to the first central auditory nucleus called cochlear nucleus. The cochlear nucleus consists of three subdivisions all having tonotopic organization, but their structure and function differ. (Figure 1.2.1).
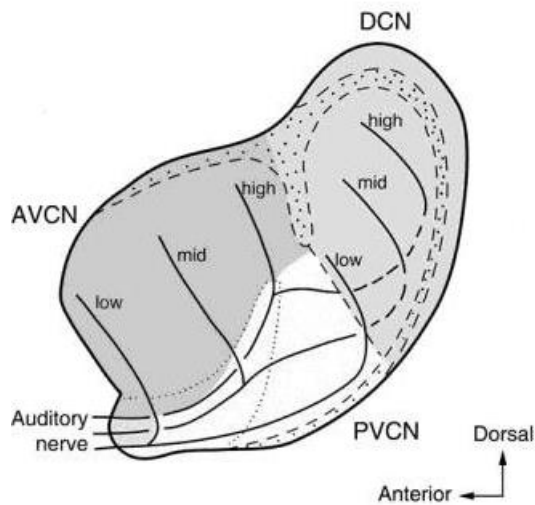
Figure 1.2.1. Sagittal section of the cat cochlear nucleus. DCN: dorsal cochlear nucleus, AVCN: anteroventral cochlear nucleus, PVCN: posteroventral cochlear nucleus. (Osen, 1969; Pickles, 2015)

Different cell types, that are morphologically different from each other, give rise to the different functions of cochlear nucleus subdivisions. Most of these cell types shows specific spiking responses to sound stimulation with a typical tuning curve (Winer & Schreiner, 2005). The different response types make up the variety of information streams emerging from the cochlear nucleus supporting parallel processing in higher order auditory nuclei, such as the superior olivary complex, the lateral lemniscal nuclei and the inferior colliculus.
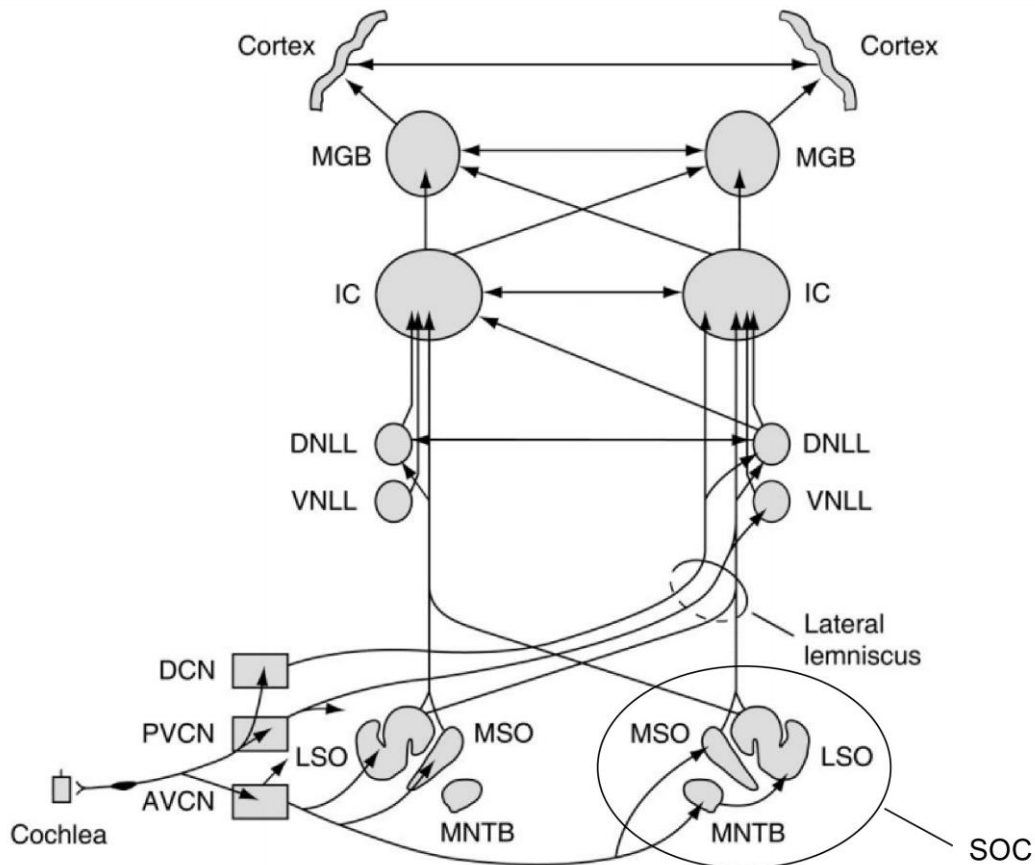
Figure 1.2.2. Major connections of the ascending auditory pathway. DCN: dorsal cochlear nucleus, AVCN: anteroventral cochlear nucleus, PVCN: posteroventral cochlear nucleus, LSO: lateral superior olive, MSO: medial superior olive, MNTB: medial nucleus of the trapezoid body, SOC: superior olivary complex, including LSO, MSO and MNTB, DNLL: dorsal nucleus of the lateral lemniscus, VNLL: ventral nucleus of the lateral lemniscus, IC: inferior colliculus, MGB: medial geniculate body. (Pickles, 2012, 2015)

The superior olivary complex is involved in binaural information extraction and integration. It comprises several nuclei that are all associated with different aspects of binaural processing. The lateral superior olive integrates monaural information from both cochlear nuclei by extracting intensity-difference sensitivity information from high frequencies and sends this information to the higher order nuclei lateral lemniscal nuclei and inferior colliculus. The precision of this binaural information extraction is determined by the inhibitory input from the medial nucleus of the trapezoid body. The medial superior olive, however, is involved in encoding delay sensitivity for low frequencies from the two ears. Neurons in the superior olive, as well as their inputs from other areas, are connected

in a highly precise manner which may reflect the precision needed for the detection of changes in stimulus features, such as interaural time and level differences and spectral information. In addition to its projections to nuclei along the ascending pathway, the superior olive sends an efferent pathway to the cochlea, supposedly modulating the electrochemical properties of outer hair cells, thus controlling the gain produced by outer hair cells in the cochlea (Liberman & Brown, 1986; Warr, 1992). Strikingly, such efferent projection to the periphery is missing in the visual or somatosensory systems. The lateral lemniscus incorporates three nuclei. These nuclei, in contrast to, for example, the IC, process purely auditory information but they are bypassed by many of the afferent pathways from the superior olive to the inferior colliculus. The dorsal nucleus of the lateral lemniscus is supposed to play an important role in binaural sound source detection in reverberant acoustic environments (Burger & Pollak, 2001; Meffin & Grothe, 2009; Pecka et al., 2007). This nucleus contains GABAergic neurons that provide a long-lasting persistent inhibition to the contralateral dorsal nucleus of the lateral lemniscus, as well as to the ICC (Faingold et al., 1993; Yang & Pollak, 1994). This feature is proposed to enable the IC to detect echoes in the binaural stimuli. Thus, the output of the dorsal nucleus of the lateral lemniscus might contribute to the perceptual phenomenon called precedence effect, when the listener's ability to localize echoes is suppressed (Blauert, 1997; Litovsky et al., 1999; Zurek, 1987). In contrast to the dorsal nucleus, the ventral and intermediate nuclei of the lateral lemniscus receive monoaural input from the contralateral ear (Merchán et al., 1997). Both contain mostly GABAergic and glycinergic cells, which project to the ICC. These neurons have very low spontaneous activity but have onset type response to auditory stimulation with precise spike timing (Covey, 1993). This suggests that neurons in the ventral and intermediate nuclei of the lateral lemniscus may transfer precise temporal information about the stimulus to the ICC (Merchán et al., 1997). Conclusively, the lateral lemniscal nuclei are additional places for the processing of different auditory information streams.

1.2.2 The auditory midbrain

The lateral lemniscal nuclei project to the next station along the ascending auditory pathway: the inferior colliculus. This nucleus is rich in diverse connections with stations along the auditory pathway, the auditory cortex and with areas associated to other sensory modalities. The IC receives inputs from all cochlear nuclei, the SO, the LLN and also

from every auditory cortical area. Besides the auditory innervations, IC also receives inputs from somatosensory nuclei located in the medulla and the pons. Most of its brainstem connections are bidirectional: the IC innervates most of the brainstem nuclei that it receives input from; and also sends ascending connections to both medial geniculate bodies (MGB). The two ICs are also highly connected with each other through commissural projections. The IC is composed of three principal divisions: central nucleus, external cortex (or lateral nucleus) and dorsal cortex. Central nucleus (ICC) is the most extensive of the three and it receives purely auditory innervation. Its neurons are organized in tonotopic layers, the so called isofrequency laminae. This division has a central role in normal hearing, as it integrates monaural inputs from ICC and binaural signals from SO and LLN. Rostrally to the ICC lies the so-called nucleus of the rostral pole of the IC. In contrast to the ICC, there is no clear topographic mapping of sound frequencies in the rostral pole, although it gets its subcortical inputs from tonotopic structures LSO and MSO. The rostral pole is also innervated by several auditory cortical areas. Main output region of the rostral pole is the superior colliculus (SC). External cortex of the IC is multisensory receiving not only auditory but also somatosensory input from the brainstem. Tonotopy has not been observed in this division. Dorsal cortex is mostly innervated by auditory cortical areas, especially the primary auditory cortex but other auditory areas such as the two other IC divisions, the MGB and brainstem auditory nuclei project to the dorsal cortex too. Although tonotopy has not yet been revealed by electrophysiological measurements in the dorsal cortex, recent 2-photon imaging results suggest the presence of a coarse gradient of characteristic frequencies. Neurons in the dorsal cortex are organized in layers but, as there is still no evidence for a tonotopic arrangement, the function of these layers is still unknown. However, stimulus specific adaptation (SSA) not purely inherited from major input primary auditory cortex is well described in this division, suggesting its role in the detection of novel sounds in the noisy acoustic environment.

1.2.3 The auditory thalamus

The thalamic structure involved in acoustic information processing is the MGB. It is innervated by subcortical auditory nuclei, mainly the IC, as well as by the auditory cortex. Its projections are almost exclusively ascending but do not only target the auditory cortex. MGB has significant projections to subcortical limbic structures, such as the amygdala

and caudate nucleus, proposing that the role of MGB is more than purely auditory. Another important feature of the MGB projections is that almost all of them are ipsilateral. The MGB comprises three divisions, namely the ventral, the dorsal and the medial division. The main sensory nucleus of the MGB is the ventral division. This nucleus receives input from the central nucleus of the IC. Structure of the ventral division resembles the IC central nucleus, as ventral division neurons with the same characteristic frequency are also arranged in tonotopic layers, except for the rostral pole in gerbils which is found to be non-tonotopic and projecting to non-primary AC areas. The rest of the ventral division sends ascending connections to primary tonotopic AC areas such as the primary auditory cortex (AI) and the anterior auditory field (AAF). Most of the neurons in the ventral division respond to binaural sound stimuli. The dorsal division is mainly innervated by the dorsal cortex of the IC but also receives input from the lateral tegmental area located near the IC. Its neurons have a wide tuning curve and long response latency, while no clear tonotopic arrangement of them is observable, just like in the dorsal cortex of the IC. Dorsal division is unique among the MGB divisions in the sense that its neurons show oscillating slow-wave activity. Projections originating from the dorsal division end in many different brain areas, including non-auditory brain structures such as the amygdala. The pattern of dorsal division projections suggests that it is involved in various computations, such as recognition of temporal patterns, spatial hearing and emotional auditory processing. The medial division is innervated by not only an auditory structure, the external cortex of the IC, but also by non-auditory brain areas such as the vestibular system. Although no clear tonotopy is found in this division, neurons are arranged along a coarse gradient of characteristic frequencies. The medial division projects to a great variety of cortical as well as non-cortical areas; it innervates all auditory cortical areas but also some non-auditory cortices and the amygdala too. This latter projection is unique among the sensory systems and it is essential for learning based on auditory cues. While the ventral division projections target layers III and IV in the cortex, medial division projections mainly end in layers I and VI, implicating that information from the MGB divisions reaches the cortex in segregated streams. The two MGBs, unlike the ICs, are not connected at all with each other; however, their auditory cortical targets are indeed connected by commissural pathways with each other.

1.2.4 The auditory cortex

The auditory cortex (AC) has several areas, with the exact number varying across species. Anyway, the AC region receiving primary sensory input from the MGB ventral division is the primary auditory cortex. Its neurons, in addition to the typical cortical organization of six layers, are arranged along two main orthogonal axes, a tonotopical and a series of bands with alternating binaural sensitivity. Auditory cortical areas, similarly to other parts of the cortex, are characterized by extensive local connections as well as by ipsilateral corticocortical and interhemispheric commissural pathways. The auditory cortex is not only innervated by various subcortical regions and densely connected with several cortical regions, but it is also the origin of pathways targeting subcortical structures. Most of these are modulating the activity of the nuclei along the ascending auditory pathway but some innervate non-auditory structures such as the striatum or the amygdala. To sum up, precise connectivity of the ascending auditory system enables fine extraction of the several aspects of auditory information that is available from the acoustic environment. However, the nuclei and pathways described in this section are only the main processing stations and connections of the ascending auditory pathway. Precise mapping of all the connections in this system would be enabled by genetic targeting of single neuronal types in each nucleus. Optogenetic targeting and neural activity control has recently been established in the gerbil (Keplinger et al., 2018), which carries the opportunity for more precise, single cell type-resolution connectivity mapping of the gerbil auditory system.

1.3. The central nucleus of the inferior colliculus

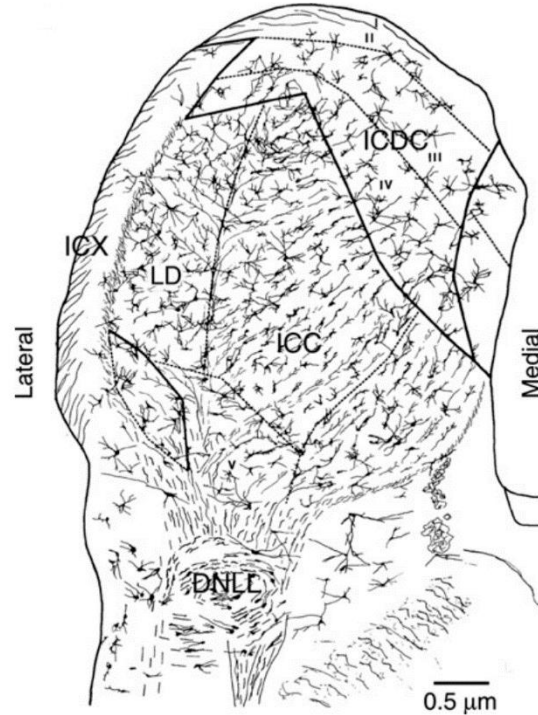### 1.3.1 Neuronal organization of the central nucleus

Figure 1.3.1: Drawing of the inferior colliculus based on Golgi-stained sections. The laminae are visible in the central nucleus. ICC: central nucleus, LD: lateral division of the ICC, ICDC: dorsal cortex, ICX: external cortex. I-IV: layers of the dorsal cortex. (Morest and Oliver 1984)

In the two studies I present in this thesis, I measured and analyzed the binaural coding properties of single neurons in the central nucleus of the inferior colliculus (ICC). To find the appropriate neurons for recordings, to appropriately design the experiments and to better understand the results, it is essential to have a good understanding of the anatomy and function of the ICC. The central nucleus is the most extensive division of the IC and it is covered by the other two IC divisions, the external and dorsal cortices (Oliver, 2005). The ICC is characterized by disc-shaped neurons with highly oriented dendritic fields. These neurons are densely packed in rows while their dendritic fields are organized parallel to each other forming the characteristic laminae of the central nucleus (Figure 1.3.1) (Morest & Oliver, 1984; Oliver, 1984, 2005).

The other characteristic cell type of the central nucleus is the stellate cell. Its dendrites radiate to all directions, forming a spherical dendritic field. These dendrites often extend

17

beyond a single lamina formed by the disc-shaped cell dendrites into the adjacent laminae (Figure 1.3.2). (Oliver, 1984, 2005)
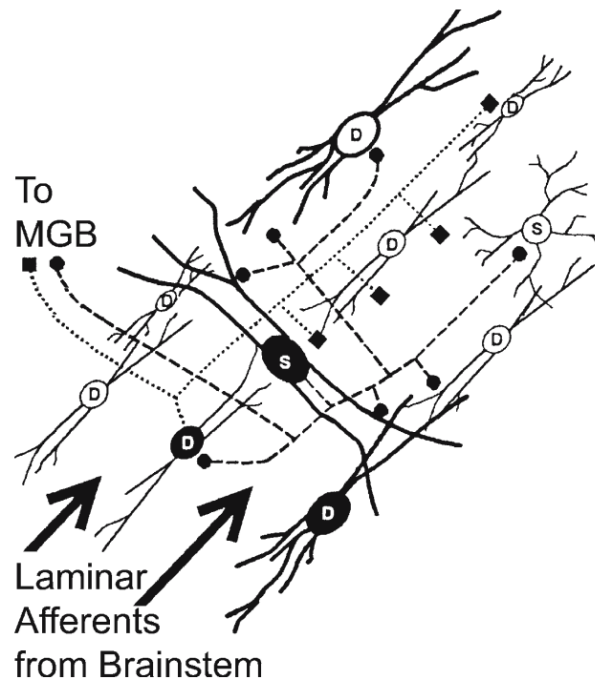

Figure 1.3.2: Afferent and efferent connections of the inferior colliculus central nucleus. D: disc-shaped neurons that align the form laminae in the central nucleus. Their axons are parallel to each other and to the afferent inputs from the brainstem tot he laminae. S: stellate cells with axons that often extend beyond a single lamina. (Oliver, 1984, 2005)

The broad distribution of their dendrites puts stellate cells in a good position to sample the input of many lemniscal axons to the inferior colliculus, whereas disc-shaped cells get a more restricted lemniscal input. Dendrites in the ICC receive input from at least seven lower level auditory nuclei. Most of the synapses ending in the central nucleus from these regions are excitatory, most of them originating from the medial superior olive, the contralateral lateral superior olive, and the dorsal and anteroventral cochlear nuclei. There is also significant inhibitory innervation provided by the dorsal and ventral nuclei of the lateral lemniscus and the ipsilateral lateral superior olive. Axons innervating the central nucleus mostly run parallel with the dendritic fields of disc-shaped neurons in the ICC laminae. In addition, axons of local cells also run in the laminae, parallel with the ascending axons (Figure 1.3.2). These local collaterals have a similar distribution as the dendrites of their parent cells. Axons of disc-shaped cells do not leave their lamina but axons originating from stellate cells also give collaterals to neighboring laminae. The two

ICs are interconnected with each other through commissural fibers that also end in laminae formed by the disc-shaped cell dendrites. (Oliver, 2005)

In the experiments presented in this thesis, we targeted the low frequency part of the ICC to record from interaural time difference (ITD) sensitive neurons. The ICC is an ideal target for recordings from neurons taking part in sound source localization because of its relatively close location to the brain surface. ITDs are perceived only for low frequency sounds, so binaural neurons in the low frequency part of the ICC play an important role in sound localization. In the following section, I shortly describe the tonotopic organization of the ICC. Knowledge about the precise tonotopic structure of the ICC was essential for me in planning and executing the recording experiments.

1.3.2 Tonotopic organization of the central nucleus

The tonotopic organization of ICC is related to the laminar structure; measured by electrophysiological recordings and protein expression studies, neurons in a lamina respond to a single frequency or, more precisely, to a small range of neighboring frequencies. Neurons in adjacent laminae can distinguish between frequencies as close as 0.28 octaves in the cat (Schreiner & Langner, 1997). Along the tonotopic axis, lowest frequencies are represented dorsolaterally, while highest frequencies are represented ventromedially. However, the existence of synaptic domains allows for the differentiation of ICC neurons not only by their characteristic frequency but also by their binaural response properties. Each excitatory innervation from an auditory brainstem nucleus to the ICC provides a different monaural or binaural input, thus enabling neurons in the laminae to code for different binaural properties, such as interaural time difference (ITD) or interaural level difference (ILD). There are three main types of ITD response of single neurons in the ICC that were described by Rose and colleagues (Rose et al., 1966) and their response characteristics were quantified by Yin and Kuwada (Yin & Kuwada, 1983). Shortly, to characterize the ITD sensitivity of a single neuron, its responses are measured and plotted for different ITDs and stimulus frequencies. The resulting curves are plotted on the same graph and the response type is determined based on these (Figure 1.3.3). For peak-type neurons, the response curves align at or near a particular frequency (Figure 1.3.3A), for trough-type neurons (Figure 1.3.3B), the response curves align at or near the trough of the curve and for intermediate-type neurons, the response curve align between the peak and the trough (Figure 1.3.3C).
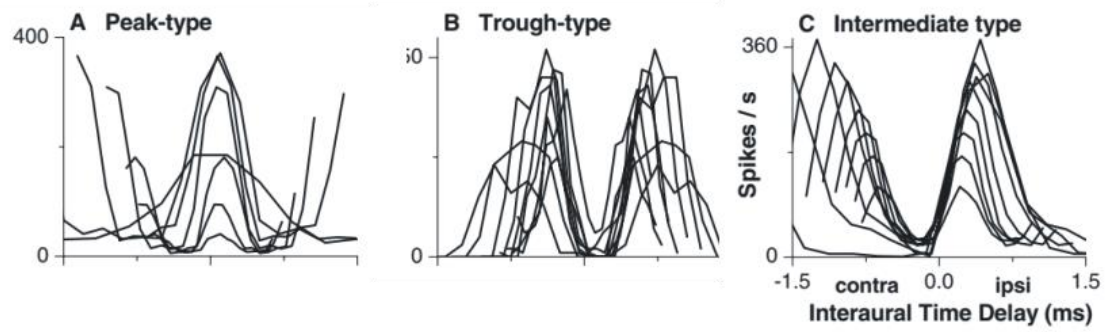
Figure 1.3.3.: ITD response types (Palmer & Kuwada, 2005; Yin & Kuwada, 1983)

In a single lamina, neurons receiving input from the medial superior olive are both ITD sensitive and show peak-type responses, whereas neurons innervated by the contralateral lateral superior olive are ITD sensitive but have a trough-type response. However, a large proportion of ICC neurons shows intermediate-type ITD sensitivity, indicating the integration of medial and lateral superior olive inputs on the same cell. Moreover, synaptic domains with monaural neurons can be also found in the ICC. These monaural inputs likely originate from the cochlear nucleus (Palmer & Kuwada, 2005).

1.3.3 Functional properties of central nucleus neurons

Neurons in the ICC are in a position to integrate inputs from several auditory brainstem nuclei such as the CN, SOC and lateral lemniscal nuclei. In addition, ICC neurons are innervated by descending axons from the AC and also by commissural axons from the contralateral IC. However, the final output of the neurons depends also on their intrinsic electrical properties. (Oliver, 2005; Sivaramakrishnan & Oliver, 2001)
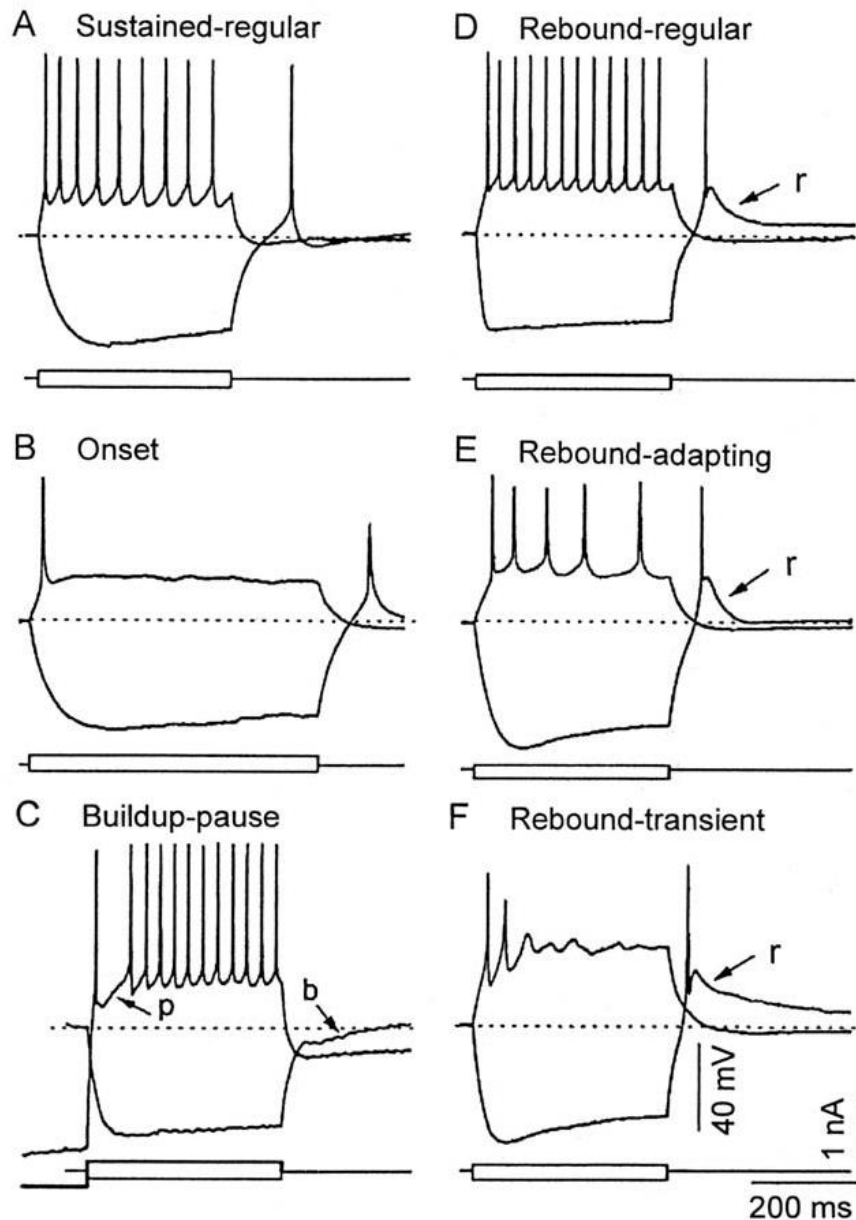

Figure 1.3.4. Intracellularly recorded firing patterns of ICC neurons. (Sivaramakrishnan & Oliver, 2001)

The intrinsic electrical properties of these cells can be described by intracellular recording techniques. ICC neurons in brain slices respond to depolarizing and hyperpolarizing current injections with various discharge patterns, which are determined by the different combinations of K^+ channels in the cell membranes. The so-called sustained-regular firing neurons (Figure 1.3.4 A) respond to a depolarizing current injection with sustained firing containing spikes with constant inter-spike intervals (Sivaramakrishnan & Oliver, 2001). Magnitude of K^+ currents sensitive to K^+ channel blockers, such as 4-aminopyridine (4-AP) and tetraethylammonium (TEA) alongside firing frequency increase similarly in response to membrane depolarization, suggesting an important role of 4-AP- and TEA-sensitive K^+ currents in sustained-regular firing (Sivaramakrishnan & Oliver, 2001). The linearity of rate-current function of this type of firing makes these cells good candidates *in vivo* for encoding of sound duration and intensity information. A population of ICC neurons shows rebound firing after hyperpolarizing current injections, that is, membrane hyperpolarization of these cells is followed by a period of depolarization that can eventually exceed action potential threshold and cause firing. The rebound depends on extracellular Ca^{2+} concentration, since reducing the Ca^{2+} concentration extracellularly eliminates the depolarization after hyperpolarizing current injection (Sivaramakrishnan & Oliver, 2001). Rebound cells can respond to depolarization in various ways (Figure 1.3.4 D, E and F). Some of them show regular firing (Figure 1.3.4 D), and, similarly to sustained-regular firing cells, delayed rectifying K^+ channels play an important role in this firing pattern. Based on their firing pattern similarity to sustained-regular firing neurons, rebound-regular cells may also contribute to sound intensity encoding. Another type of rebound cells, the rebound-adapting neurons (Figure 1.3.4 E), fire action potentials with non-uniform inter-spike intervals in response to membrane depolarization, with spikes initially close to each other and becoming more separate during the depolarization. Finally, rebound-transient cells (Figure 1.3.4 F) fire transiently in response to a depolarizing current injection, which is likely to be caused by K^+ channel blocker charybdotoxin (K^+ channel blocker)-sensitive Ca^{2+}-activated K^+ currents (Oliver, 2005; Sivaramakrishnan & Oliver, 2001). The depolarizing rebound occurring after membrane hyperpolarization may underlie several coding mechanisms observed in the ICC. Rebound mechanism, for example, can be responsible for the so called OFF response of IC cells to acoustic stimulation, that is, neurons fire after the offset

of stimulus. This response type may enable the coding of ongoing complex sounds such as amplitude-modulated tones, as well as the emergence of sound duration tuning. The onset firing cells (Figure 1.3.4 B) fire a single action potential in response to a depolarizing current injection at its onset, which may be caused by the low threshold K^+ current measured in this type of neurons (Sivaramakrishnan & Oliver, 2001). Finally, buildup-pause neurons (Figure 1.3.4 C) respond to membrane depolarization that follows a hyperpolarizing current step with delay in the onset of sustained firing or with a pause between the first and the following spikes. The fast transient and rapidly inactivating A-type K+ current may underlie this firing pattern (Sivaramakrishnan & Oliver, 2001). Due to their special response to the combination of membrane hyperpolarization and depolarization, buildup-pause neurons may encode pairs of inhibitory and excitatory stimuli or even successive excitatory stimuli. Most of the cells in the ICC were found to be sustained firing and a rebound after membrane hyperpolarization was found in more than half of them (Sivaramakrishnan & Oliver, 2001). The different firing patterns of ICC neurons were found to be not correlated with the morphology of the cells, so the two main types of ICC cells, disc-shaped and stellate cells can have any of the listed firing patterns (Bal et al., 2002; Peruzzi et al., 2000).

In vivo firing characteristics of ICC neurons are similar to discharge patterns observed in slices. Primary-like neurons, much like the auditory nerve fibers, sustain their firing during the presentation of sound, with some adaptation occurring at the beginning of the response.
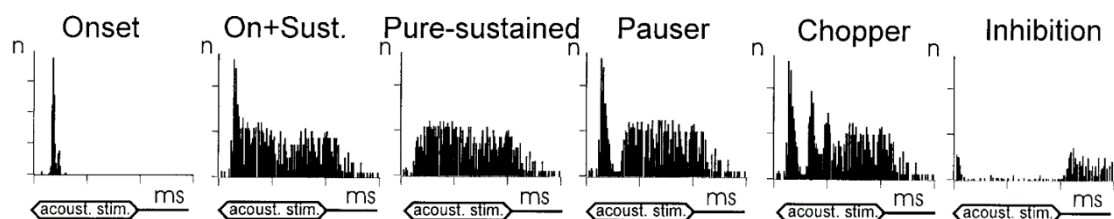

Figure 1.3.5 Typical response patterns of neurons in the ICC in vivo. (Syka et al., 2000)

Pauser firing pattern (Figure 1.3.5) is very similar to primary-like patterns, except that in pauser neurons after the response onset an abrupt and variable in length pause of silence

in firing occurs before the sustained firing during stimulus presentation (Syka et al., 2000). The two components of the pauser response can also be observed separately as responses to acoustic stimulation. Pure onset firing occurs only at stimulation onset, while late response has a delay from stimulation onset, but it is sustained during the whole stimulus duration. Some neurons in the ICC respond to sounds with ON-OFF firing, which means firing only at the onset and offset of stimulation. In addition, sound can not only excite ICC neurons, as some neurons are inhibited by stimulation (Figure 1.3.5 last panel) (Syka et al., 2000). The various response patterns of ICC neurons reflect the firing patterns of cells innervating them but are also shaped by the intrinsic properties of ICC cells. Neurons often can show more of the listed firing patterns in response to sounds with different frequencies and intensities (Le Beau et al., 1996). For example, sustained responses can be observed in response to best frequency stimulation, while stimuli away from the best frequency of the neuron or low-amplitude best frequency stimuli evoke only onset responses. Intensity change of the stimulating sound can also evoke a firing pattern change in itself, as a primary-like response to near-threshold sound turns into a pauser pattern with increasing sound intensity. In addition, tones away from the best frequency can even evoke inhibitory responses (Ehret & Schreiner, 2005).

Neuronal responses in the ICC to changes of stimulus features can vary depending on neuron type and their afferent connections. For example, when stimulus intensity increases, the number of cochlear nerve fiber action potentials increases proportionally until reaching a plateau, showing a monotonic intensity-spike count function. However, in the IC, only half of the neurons has a monotonic intensity-spike count function, while the other half responds non-monotonically, that is, the spike count drops after a peak when sound intensity is increased further. This non-monotonic response may be shaped by the interconnection of excitatory and inhibitory neurons tuned to the same tone frequency. Such circuits exist both in the ICC and in the CN, and it is assumed that both monotonic and non-monotonic response of ICC neurons are transmitted from their CN inputs (Ehret & Schreiner, 2005).

A neuron's response to sounds with different frequencies is described by its tuning curve, which indicates the firing threshold intensities of the neuron for various sound

frequencies. The frequency or frequencies evoking firing at lowest intensity are called characteristic frequency (CF). Most neurons in the ICC have sharp tuning curves with well-defined CFs and show sustained firing in response to CF stimulation. These neurons are organized according to their CFs and form the isofrequency laminae described earlier. Conclusively, the tuning curve is an important measure to characterize neural response to sound frequency changes, however, it is not able to describe neuronal firing in response to suprathreshold intensity stimulation with different frequencies. The measure frequency response area (FRA) overcomes this limitation, and it is therefore especially useful for characterizing non-monotonic responses. Based on these measures, it is most likely that information obtained by frequency discrimination is encoded by neurons with monotonic responses in the ICC (Ehret & Schreiner, 2005).

1.3.4 Neural coding in the central nucleus

As described previously, ICC neurons have to encode information retrieved by sound pitch discrimination. Their frequency coding properties are described by their response to different frequency pure tones, which is characterized by the tuning curve.
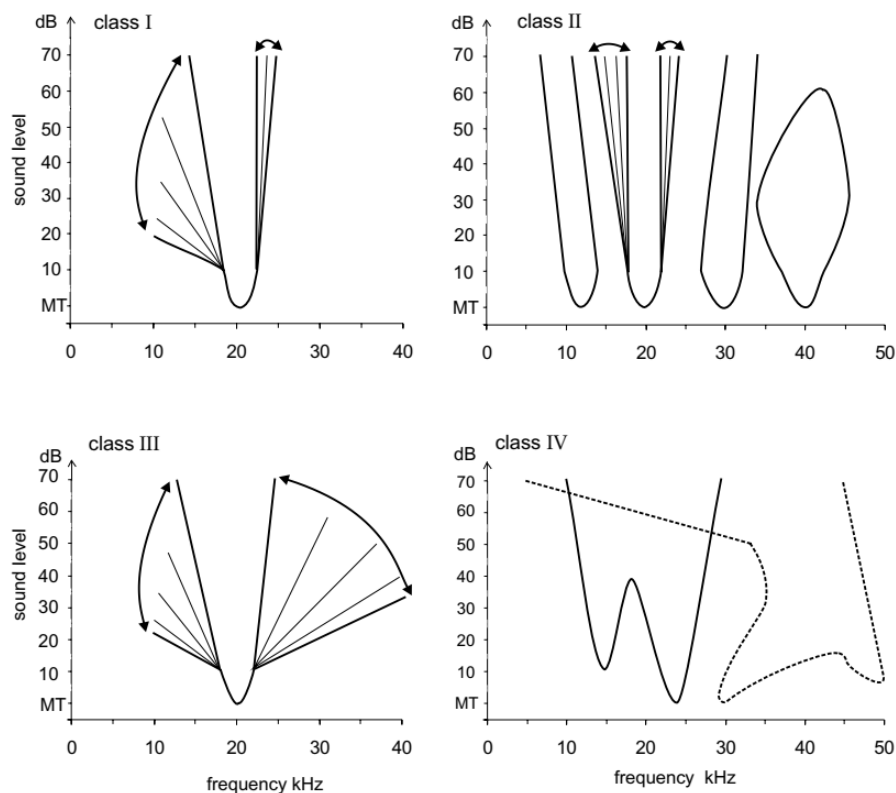

Figure 1.3.6 Typical tuning curves in the central nucleus of the inferior colliculus. (Egorova et al., 2001; Ehret & Schreiner, 2005)

ICC tuning curves differ from auditory nerve tuning curves in the more important role of inhibition in the ICC shaped by the various intrinsic and innervating inhibitory connections in this nucleus. From the interaction of excitation and inhibition in response to pure tone stimulation emerge four major classes of tuning curves of ICC neurons. Class I neurons (Figure 1.3.6) have tuning curves that are very similar to auditory nerve tuning curves. These are strongly asymmetric with steep high frequency and shallow low frequency slopes, have asymmetric inhibitory areas below and above CF with higher inhibitory threshold below CF and the inhibitory areas only partially overlap with the excitatory receptive field. In contrast, Class II is dominated by inhibition, as the inhibitory areas are symmetrical, the inhibitory thresholds below and above CF are similar and the inhibitory areas always overlap, often completely, with the excitatory receptive field. Still shaped by inhibition, excitatory tuning curves of class II neurons can be either symmetrical, skewed toward the low or high frequency side or even closed, this latter meaning an upper response threshold. Class III is characterized by weaker inhibition that makes excitatory tuning curves symmetric and shallow, as well as inhibitory areas small. Finally, Class IV neurons have complex tuning curves and can have several excitatory receptive fields or even multiple CFs. The proportion of the four classes is similar in mammal species, with about one-fourth of ICC neurons belonging to Class II, less than 10 % belonging to Class IV and low-inhibition classes I and III sharing the rest of the neurons. These four classes cover the whole population of ICC neurons, as virtually all of them are frequency tuned, with their sharpness of tuning differing from that of auditory nerve fibers (Egorova et al., 2001; Ehret & Schreiner, 2005).

Neurons in the ICC are organized tonotopically along two frequency gradients. The first gradient spans from dorsal and dorsolateral to ventromedial ICC with lower frequencies represented dorsolaterally and higher frequencies represented ventromedially. This gradient is orthogonal to the cellular laminae described before, as well as to their afferent inputs. The previously described isofrequency laminae emerge from this frequency gradient (Figure 1.3.7 A). The second gradient spans within the isofrequency laminae, with low frequencies represented dorsomedially and high frequencies represented ventrolaterally. The two gradients together form so called frequency-band laminae that

are larger than the synaptic domains formed by the anatomical organization of ICC neurons (Figure 1.3.7 B).
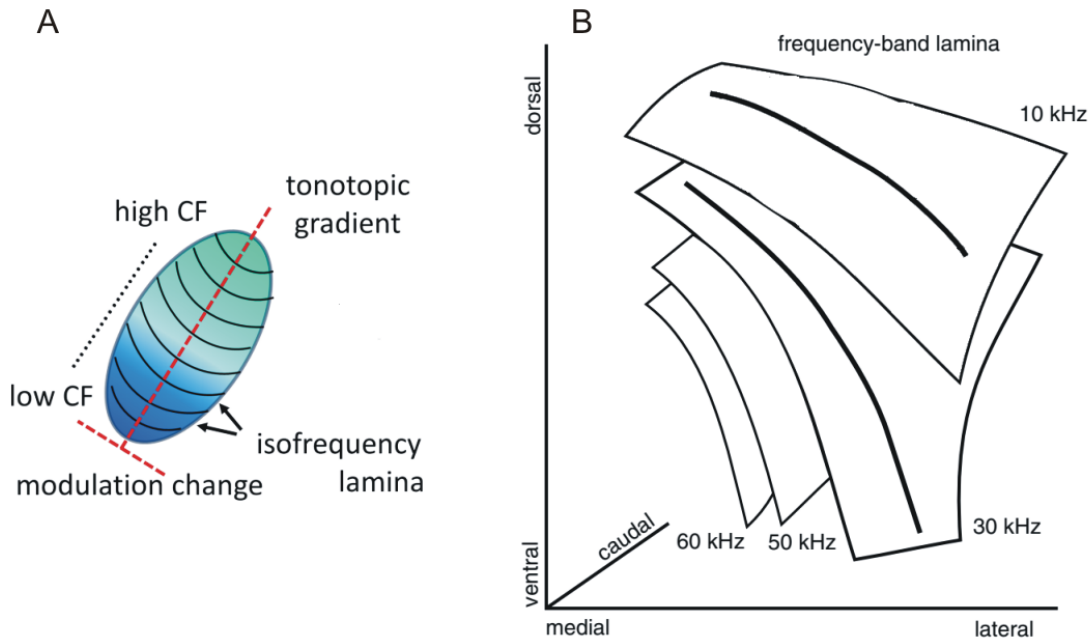

Figure 1.3.7 Schematics illustrating the tonotopic organization in the central nucleus. A: The main tonotopic gradient of the central nucleus and the isofrequency laminae. (Lyzwa & Wörgötter, 2016) B: Isofrequency-band laminae in the central nucleus. Thick lines: gradient of frquency increase from medial to lateral within a lamina. (Schreiner & Langner, 1997; Ehret & Schreiner, 2005)

Sound intensity coding in ICC neurons, as already described, depends on the rate level function of neurons. About half of them have monotonic rate level functions, often with a plateau above a certain intensity, while the other half have non-monotonic rate level functions. Receptive fields and rate level functions of ICC neurons are likely shaped by the same factors, such as the interaction and combination of excitatory, inhibitory and facilitating inputs to the ICC, since receptive fields often have intensity dependent components (Ehret & Schreiner, 2005). In addition, ICC cells also encode other sound features such as periodicity and are organized in a periodotopic map that is orthogonal to the main tonotopic axis of the ICC (Langner & Schreiner, 1988). According to recent studies, periodotopy emerges first in the IC in the auditory system, which then also appears in other higher order auditory brain areas such as the primary auditory cortex.

This feature adds another important aspect to the role of the IC in the ascending auditory pathway. (Rees & Langner, 2005)

1.3.5 Coding of binaural sounds and spatial location in the central nucleus

The inferior colliculus and especially the central nucleus lie in a perfect position to process binaural inputs, as the main brainstem binaural nuclei, the medial and lateral superior olive send their axons here (Figure 1.3.8) (Schofield, 2005). Additionally, the inferior colliculus receives monaural inputs from the cochlear nuclei, thus extending the processing capabilities of this nucleus (Cant, 2005).
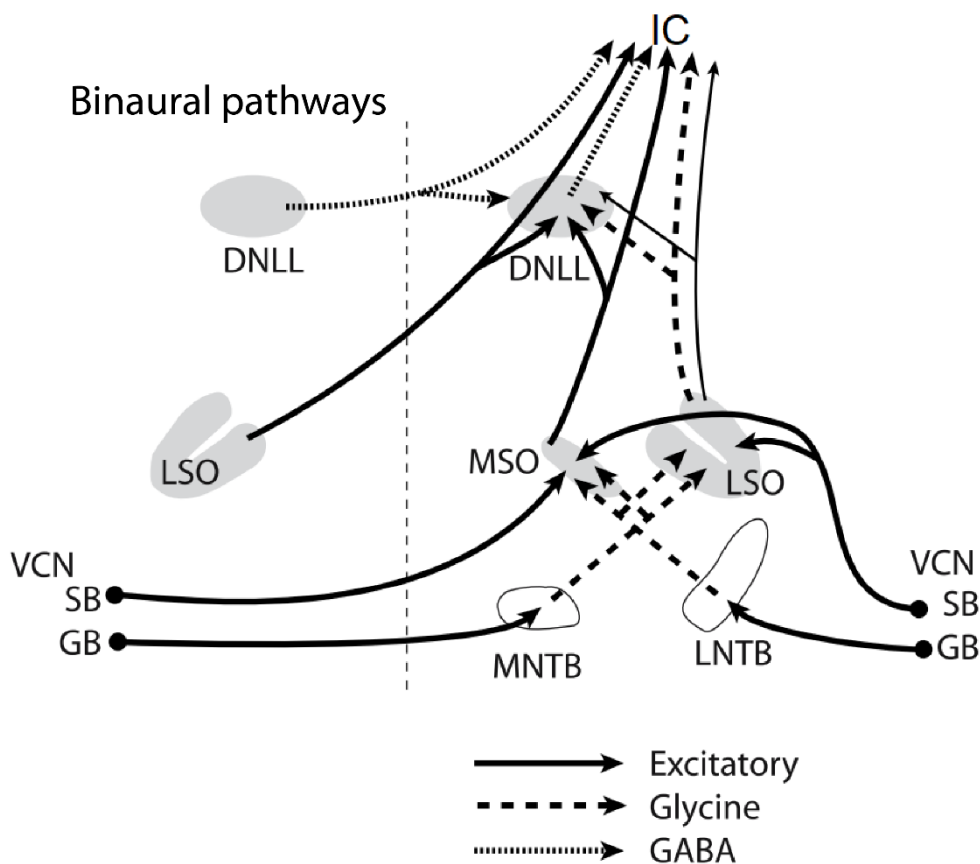
Figure 1.3.8. Binaural inputs to the inferior colliculus. (Schofield, 2005)

Neurons in the central nucleus are therefore sensitive to binaural cues, such as ITD, ILD or spectral cues. As previously described, the ITD function of central nucleus cells is shaped by their afferent inputs, as well as by intrinsic connections in the central nucleus. Most central nucleus neurons prefer contralateral ITDs, that is the sound arrives first in the contralateral ear and then in the ipsilateral one (Palmer & Kuwada, 2005). However, this ITD preference is frequency dependent, namely neurons with high characteristic

frequency prefer low or even 0 ITDs while cells with low characteristic frequency tend to pick higher ITDs. The central nucleus ITD functions' peak often lies outside the physiological range of the animal but the steepest slope of the ITD function matches the preferred ITDs of the cells, allowing for the rapid detection of ITD change (Palmer & Kuwada, 2005). There are more possibilities for the anatomic arrangement of the ITD detection system. According to the classical model proposed by Jeffress (Jeffress, 1948), ITD sensitive neurons are arranged along their preferred ITD and act as so called coincidence detectors, which means that they receive input from the two ears by axons of different lengths according to the preferred delay of each cell and fire when input arrives from the two ears coincidentally. Such a structure was confirmed in birds, however, in mammals there is still no clear evidence for the existence of these coincidence detectors (McAlpine & Grothe, 2003). In contrast, studies suggest another method for ITD detection in mammals that is shaped by the interaction of excitation and inhibition in the brainstem nuclei lateral and medial superior olive and the medial nucleus of the trapezoid body. It is most likely that in rodents this later mechanism is responsible for ITD detection and the ITD sensitivity in the central nucleus is formed by the afferent connections from the above-mentioned brainstem nuclei (McAlpine & Grothe, 2003). As it is seen from the summary above, ITD sensitivity to static sounds is formed already in the SO, where neurons are insensitive to motion. Meanwhile, ICC neurons are clearly sensitive to moving sound sources, so this stage of the auditory pathway adds an important aspect to sound source detection. Sensitivity to the ITD of high frequency sounds is limited by the size of the head, but ITDs of higher frequency sounds can also be detected in the case when the ITD of the envelope of amplitude modulated high frequency sounds is detected (Leakey et al., 1958; Palmer & Kuwada, 2005). However, the main mechanism for detecting interaural differences of high frequency sounds is the interaural level difference (ILD) detection. ILD sensitivity in the ICC is determined by input from the LSO, the nucleus in the brainstem where ILD sensitivity first develops in the brain (Palmer & Kuwada, 2005). ILD sensitive neurons in the LSO receive excitation from the ipsilateral ear and are inhibited by input from the contralateral ear. This pattern is reversed in the ICC because the input from high frequency LSO to ICC is crossed. As all ICC neurons are excited and inhibited similarly by input from the two ears, it is postulated that ILD is encoded by a population code,

that is, more and more ICC neurons get in excitation as a sound source moves from the ipsilateral to the contralateral hemisphere (Schofield, 2005).

1.3.6 Spectral and temporal filtering in the central nucleus

In the early auditory pathway, spike patterns generally reflect the fine structure of the sound waveform (for neurons with low preferred frequencies) and/or its amplitude envelope (for neurons with high preferred frequencies). The overall spike rate of auditory neurons, adaptive adjustments in dynamic range notwithstanding (Dean et al., 2005; Wen et al., 2009), typically increases with increasing mean intensity, though it may saturate or decrease at high intensities. In addition to modulating spike rate, changes in intensity can also have an indirect effect on the timing of spike patterns by evoking changes in the way in which sounds are filtered. For example, temporal filtering for high-frequency sounds is adapted to changes in intensity such that the system is always optimized for the current operating regime; for soft sounds, temporal filtering is low-pass so that resources are focused on low modulation frequencies where the signal-to-noise ratio (SNR) in natural sounds is typically highest, whereas for loud sounds, temporal filtering is bandpass, so that the redundancy in natural sounds at low modulation frequencies can be reduced (Lesica & Grothe, 2008b; Nagel & Doupe, 2006; Rees & Møller, 1987). Such changes in spectral and/or temporal filtering can help ensure that the flow of information in the periphery is robust to changes in intensity and may provide a substrate for invariant responses in the cortex (Billimoria et al., 2008; Sadagopan & Wang, 2008). Space is not represented topographically within the brain areas of the early auditory pathway, but is instead encoded directly in neuronal responses such that the spike rate evoked by a given sound is dependent not only on its intensity, but also on its spatial location. In the inferior colliculus (IC), for binaural neurons with low preferred frequencies, spike rate varies with the interaural time difference (ITD), typically as a monotonic function within the range of ITDs corresponding to realizable azimuthal angles (Groh et al., 2003; Hancock & Delgutte, 2004; Lesica et al., 2010; McAlpine et al., 2001). It is unknown whether, as described above for intensity, ITD-dependent changes in spike rate are accompanied by changes in spectral and/or temporal filtering that help to maintain the flow of information. Furthermore, although changes in intensity and ITD may have similar effects on the spike rate of a given neuron, they are certain to have different effects on spike rates across the entire population. A change in intensity will cause, on average, the same change in spike

rate in both hemispheres of the brain, whereas a change in ITD will have opposing effects on the two hemispheres, which, in terms of ITD sensitivity, are mirror images of each other.

1.4. Decoding of neural responses

Neural responses can be analyzed in several ways. In sensory experiments, neural responses are evoked by sensory stimulation and should be analyzed regarding the stimulus. With the advent of precise recording and fast analysis techniques, decoding of the responses of neurons to sensory stimuli became one of the most frequently used methods for analysis (Brown et al., 2004; Pouget et al., 2000). Decoding means for each single trial the prediction of which stimulus evokes the particular neural response (Quian Quiroga & Panzeri, 2009). Several different decoding approaches have been developed in the past years. Most of them divide the trials in two groups: training set and validation set. Training set is used to optimize the decoder, while validation set is used to validate the performance of the decoder. It is important to exclude trials in the training set from the validation set because using the test trials for validation can lead to overfitting and provide artificially high values of decoder performance (Aljadeff et al., 2016). To achieve this separation of training and test sets, a widely used method is the so called 'leave-one-out' method when each trial is predicted based on all the other trials. The different decoding algorithms use various mathematical methods to decode the neural responses, such as Bayesian approach, nearest-neighbor estimation or support vector machines. In all cases, performance of the decoder has to be quantified which is achieved by computing the relative number of correctly decoded trials (Quian Quiroga & Panzeri, 2009). Based on this, each prediction can be regarded as a Bernoulli variable (having two values, correct or incorrect). Consequently, the probability that a sequence of trials was decoded correctly follows the Binomial distribution, which enables computing the statistical significance of decoder performance in the following way. The probability of getting k hits by chance from n trials based on the Binomial distribution is

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

where $p=1/K$ is the chance of getting a hit by chance and K is the number of stimuli. Then, the p-value for statistical significance can be computed by adding up the probabilities of getting k or more hits by chance from the n trials:

$$p - value = \sum_{j=k}^{n} P(j)$$

This quantification enables researchers to use decoding as a powerful tool for neuronal signal analysis. The power of decoding neural signals lies in its capability of predicting the stimulus that caused the single-trial neural response. However, the amount of information extracted by decoding will always be less than the information available in the neural responses (Quian Quiroga & Panzeri, 2009). Another issue that has to be considered when using decoding of neural responses is the dimensionality problem. This means that explicit estimation of the probability distribution described above may be difficult due to the limited amount of data and high dimensional responses (Quian Quiroga & Panzeri, 2009). Experiments recording simultaneously a high number of cells yield such data: a high number of single units for which the relevant feature of firing is spike timing and not just the number of spikes. One of the successful methods to overcome this dimensionality problem is the application of distance metrics for decoding.

1.4.1 Distance metrics of spike trains

The variability of neural responses has to be quantified in order to be able to mathematically compare two single trials. A useful tool for this quantification is the application of distance metrics (Victor, 2005). A recording of neural activity can be conceptualized as a sequence of discrete events in continuous time, the neural spikes that are action potentials fired by single neurons. Such a structure of events is called a point process (Perkel et al., 1967) and various distance metrics can be constructed for it (Victor, 2005). By definition, a metric is a function that associates a real nonnegative number analogous to distance with each pair of elements in a set such that the number is zero only if the two elements are identical, the number is the same regardless of the order in which the two elements are taken, and the number associated with one pair of elements plus that associated with one member of the pair and a third element is equal to or greater than the number associated with the other member of the pair and the third element (Rolewicz, 1987). This definition is applicable for the construction of distance metrics for spike trains

too. The most straightforward way to construct a metric for spike trains is to define a cost-based metric, that is, the distance of two spike trains is the minimal cost of transforming one spike train to the other (Victor, 2005). For this, the costs of transformation steps have to be defined, which is done by applying simple rules to the basic operations of inserting, deleting and shifting a spike in a spike train. In the most widely used type of spike train metrics, the so-called spike time metrics (Figure 1.4.1), inserting or deleting a single spike has a cost of 1. The other rule of spike time metrics states that the cost of moving a spike in time (shifting the spike) is proportional to the amount of time that it is moved. By applying these rules, the distance between two spike trains will be the minimum total cost of the above defined steps transforming one spike train into the other (Victor, 2005).
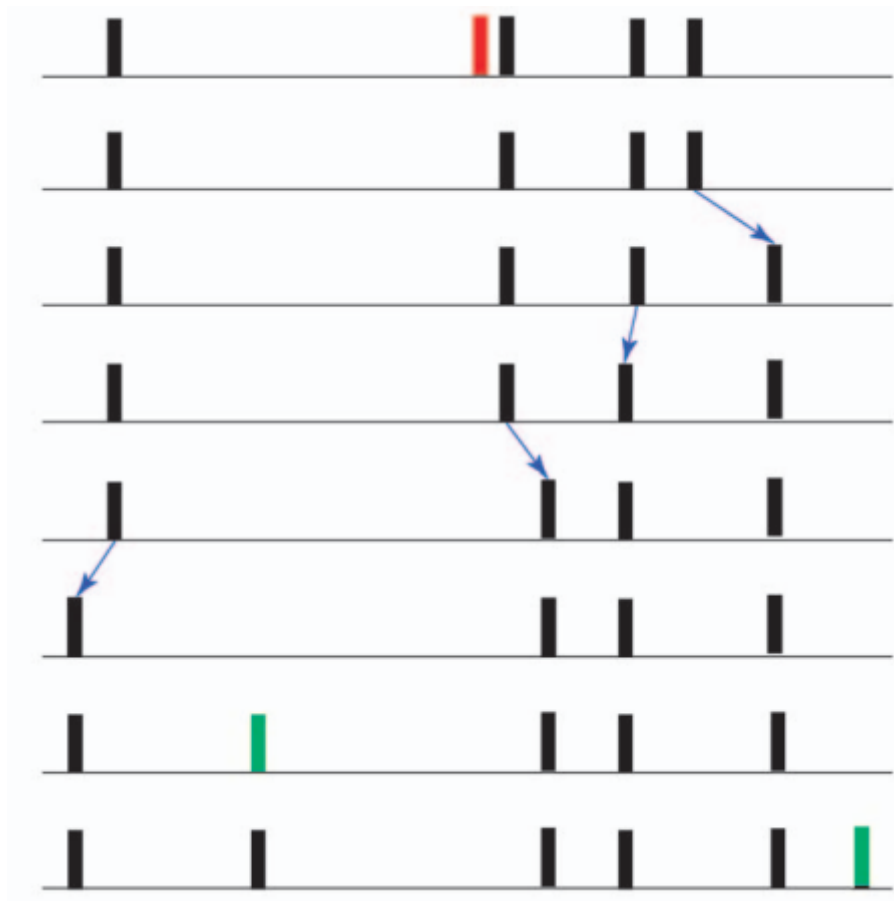

Figure 1.4.1. Distance metrics of spike trains. Red: deleting a spike. Green: inserting a spike. Blue arrows: shifting a spike. (Victor, 2005)

1.4.2 Decoding of neural population responses with distance metrics

The introduced distance metrics provide an intuitive method for decoding neural responses, as responses to the same stimulus should be closer to each other, while responses to different stimuli should be farer (Hofer et al., 2010). In addition, it can also be related to the classical approach of decoding under simplifying assumptions: If each response is represented as a point in multi-dimensional space and the distribution of the responses evoked by repetitions of the same stimulus within that space is assumed to be Gaussian, then the log likelihood that a response was evoked by a particular stimulus (assuming that all stimuli are equally likely) is proportional to the square of its distance from the average of all responses evoked by that stimulus. When the application of distance metrics for decoding began to be a routine, most studies focused on the question if decoding performance is influenced by cell identity, that is, whether the performance differs if all spikes are assumed to come from a single cell (Aronov et al., 2003; Houghton & Sen, 2008). However, the advent of multiple channel recordings, advanced spike sorting methods and high computational power computers enables researchers to study the influence of multiple identified neurons on decoder performance. In this section I will describe our study (Hofer et al., 2010) investigating how varying the influence, or weight, of each cell affects population decoding performance. The choice of weight for each cell is a complex problem and should be based not only on how informative the response of each cell is individually, but also on the correlations between the responses of each cell and the others in the population. This problem can be illustrated through a simple example of averaging: For a series of measurements in which the noise in each measurement is independent and of equal magnitude, standard averaging yields the optimal estimate of the underlying signal. If the magnitude of the noise varies across measurements, then some measurements will be more reliable than others, and a weighted average based on this reliability will yield the optimal estimate. However, if the noise in a fraction of the measurements is correlated, then averaging across those measurements will be less effective in reducing the noise (in the extreme of identical noise, averaging has no effect) and a weighted average that favors the uncorrelated measurements may provide the optimal estimate, even if the correlated measurements are individually more reliable. As with averaging, the optimal weights for population decoding with distance metrics are dependent on both the individual reliability of cells in the population and the correlations between them. In the following section, based on our study (Hofer et al., 2010), I will

describe an algorithm to find these optimal weights and demonstrate its utility by decoding experimental responses with a variety of correlation structures.

2. Specific aims

2. 1 Decoding neuronal population activity with distance metrics

In the classical approach of decoding, the probability distribution of the stimulus, conditioned on the observed responses, is involved to determine which stimulus was the most likely. However, for experiments where data are limited and the responses are high dimensional, the explicit estimation of the stimulus probability distribution may be difficult. An intuitive and practical way to overcome the dimensionality problem of decoding is using distance metrics (van Rossum, 2001; Victor & Purpura, 1996). Several studies applied distance metrics successfully for answering various sensory neuroscience questions (Victor, 2005), however, these used distance metrics for analyzing single neuronal responses. Previous attempts to apply distance metrics for decoding of neuronal population activity focused on the extent to which decoding performance is dependent on cell identity, i.e. whether performance differs if all spikes are assumed to come from a single neuron (Aronov et al., 2003; Houghton & Sen, 2008). Our aim was to investigate, how varying the influence, or the weight, of each cell affects population decoding performance. The motivation came from the observation that the choice of weight for each cell is a complex problem because it should be based not only on how informative the response of each cell is individually, but also on the correlations between the responses of each cell and the others in the population. In the chapters related to this study, I will describe an algorithm for finding these optimal weights and demonstrate its utility by decoding experimental responses with a variety of correlation structures.

2.2 Revealing the effects of interaural time and intensity difference on the coding of low-frequency sounds

The aim of this study is to describe both the single cell and the population level coding of low frequency sounds, while changing ITD and intensity. Spike train metrics was applied as a decoding method to reveal the changes in coding. This method enables the preservation of spike timing. A combination of 3 different sound intensity levels and 5 ITD values were used for auditory stimulation. When decoding responses to these stimuli, special care must be taken to make sure that the changes in the responses are not due to overall spike rate changes, but they are real changes in spike timing. Careful analysis was carried out to reveal if changes in intensity and ITD affect neural responses in different ways. Intensity changes can be studied both along the positive and negative

slopes of the RLF, and this part also shows the difference in the effects of changes along the positive and negative slopes of the RLF.

To sum up, the aim of this part is to demonstrate that, at least at moderate intensities, the auditory system employs different strategies at the single neuron and population levels simultaneously to ensure that the coding of sounds is robust to changes in other stimulus features.

3. Materials and methods

3.1. Experimental surgical procedures

Data obtained from the acoustic stimulation experiments on anesthetized animals was used to test the performance of the decoding algorithm described in (Hofer et al., 2010) and to perform the analysis steps for the ITD and ILD change effects on information coding (Horvath & Lesica, 2011). Adult Mongolian gerbils (Meriones unguiculatus) of both sexes were used for anesthetized neural activity recording experiments. A total of 24 animals were used for the two studies. All experiments were approved according to the German Tierschutzgesetz (AZ 211–2531– 40/01 and AZ 211–2531– 68/03). First, the animals were intraperitoneally injected (0.5 ml / 100 g body weight) a physiological saline solution containing 20% ketamine and 2% rompun (xylazine). During surgery and recordings, the same solution was infused continuously at a rate of ~0.1 ml/h. Constant body temperature of 37-39°C was maintained using a thermostatically controlled heating blanket. Skin and tissue covering the upper part of the skull was cut and gently pushed aside laterally, and a small metal rod was mounted on the frontal part of the skull using UV-sensitive dental-restorative composite material (Charisma, Heraeus Kulzer). The rod was used to secure the animal's head in the stereotaxic frame while leaving the ear canals free for auditory stimulation. Earphones and probe-tube microphones were inserted into the ear canal and were held by custom-made earphone-holders. Then, the animal was transferred into a sound-attenuated chamber and mounted in a custom-made stereotaxic holder. A craniotomy was made over the inferior colliculus, 1.3-2.6 mm lateral from the midline and 0-1 mm caudal from bregma. The dura mater overlying the cortex was carefully removed, and either glass electrodes filled with 1 M NaCl (5-15 MOhm) were inserted into the central nucleus of the inferior colliculus (2-4 mm below brain surface) (in case of the publication (Hofer et al., 2010)) or a multielectrode microdrive (Thomas Recording) was used to advance seven independently moveable microelectrodes into the central nucleus of the inferior colliculus (in case of the publication (Horvath & Lesica, 2011)). For this latter publication, all recordings were made in the low-frequency lamina of the rostrolateral quadrant of the IC, where inputs from the medial superior olive (MSO) are clustered (Cant & Benson, 2006) and cells are likely to be ITD sensitive. For both publications, extracellular action potentials were amplified and fed into a computer via an A/D converter (RX5, Tucker Davis Technologies). Clear isolation of single units was achieved using offline spike sorting in MClust software (Fraley & Raftery, 1999, 2002).

Only those units with an isolation distance > 10 were included in these studies (Schmitzer-Torbert et al., 2005). Experiments typically lasted 10-14 hours. After recordings, the animal was killed without awakening by an injection of 0.1 ml of barbital.

3.2 Acoustic stimulation

Acoustic stimuli were generated at a 50 kHz sampling rate by TDT System III (Tucker Davis Technologies). The generated stimuli were converted to analog signals (RP2-1, TDT), attenuated (PA5, TDT) and delivered to electrostatic speakers (EC1, TDT or ER2; Etymotic Research) that were coupled to the tubes which were inserted in the ears along with microphones (ER10B; Etymotic Research). Speakers were calibrated to have a flat frequency response (5 dB SPL from 0.1 to 10 kHz) after coupling to the ears at the beginning of each experiment. For the publication (Hofer et al., 2010), stimuli were 8 different instances of randomly generated Gaussian white noise. In total, each instance was played 20 times in a random order. For the publication (Horvath & Lesica, 2011), at each recording site, a sequence of sounds with various frequencies, intensities, and ITDs were presented to characterize basic response properties. First, 100 ms pure tones of various intensities and frequencies were presented, separated by 150 ms periods of silence, to determine the frequency response area (Figure 3.2.1 A). Tones were presented binaurally with zero ITD and had a rise/fall time of ± 5 ms. Next, eight repeated presentations of a 250 ms token of frozen Gaussian noise at ITDs ranging from - 2 to 2 ms were presented, separated by 500 ms periods of silence, to compute noise delay functions shown in Figure 3.2.1. B. From the noise delay functions, the best ITD (the one evoking the highest spike rate) was extracted (Figure 3.2.1 C). The noise was filtered to contain only frequencies between 200 and 4000 Hz and had a rise/fall time of 5 ms. The intensity of the noise was 50 dB SPL. Finally, the sounds used for the main decoding analysis were presented: 20 repeated presentations of eight different 250 ms noise tokens (filtered as above) at five different ITDs (- 135, - 67.5, 0, 67.5, and 135 µs) and three different intensities (43, 63 and 83 dB SPL), separated by 500 ms periods of silence. The same tokens were then presented a second time with added broadband background noise (different on every trial) with a signal-to-noise ratio of 0 dB.
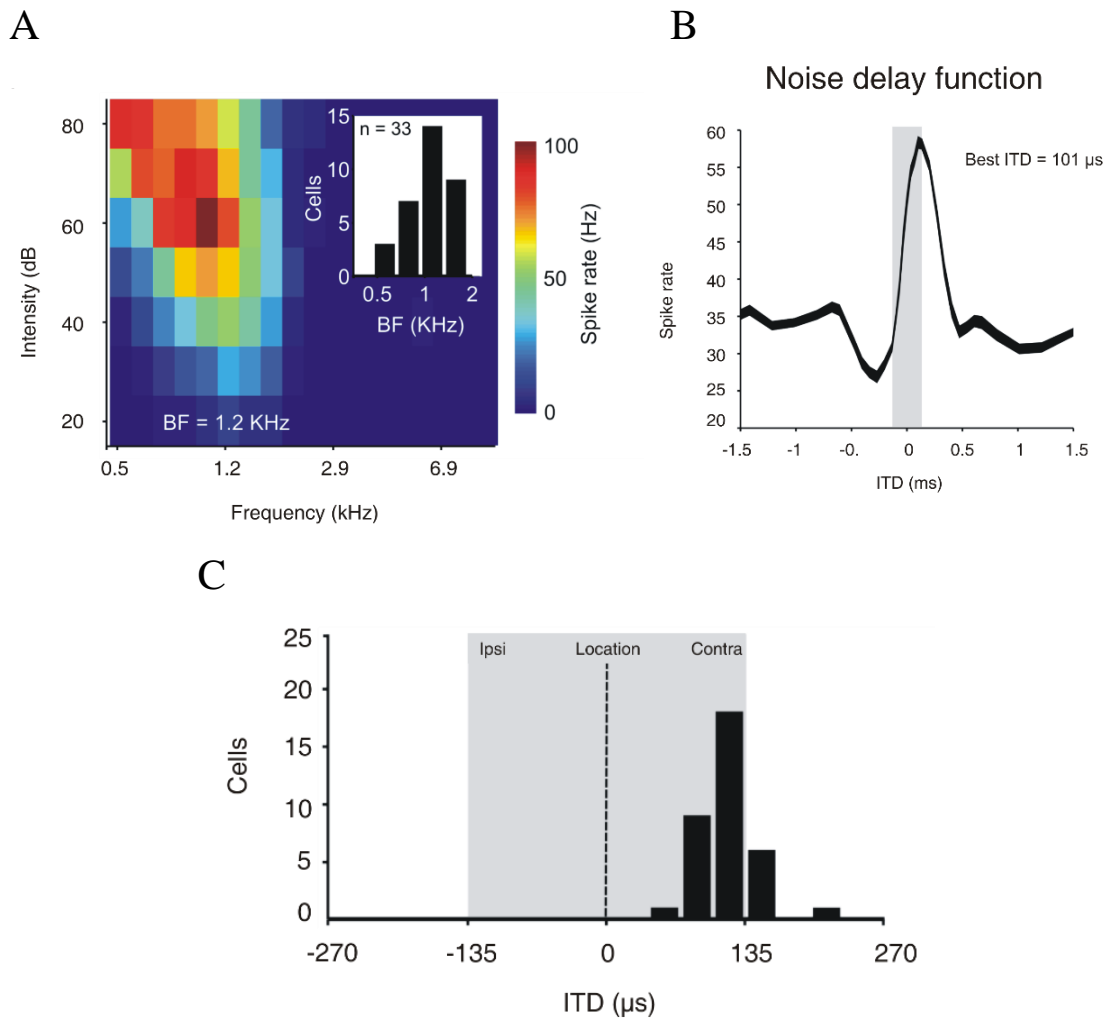
Figure 3.2.1: Response properties of IC neurons. A: frequency response area (FRA) from responses to tone bursts with varying frequency and intensity, for a typical neuron. For our population of neurons, the best frequencies (BFs) were clustered between 500 and 2000 Hz, with a median value of 1120 Hz (see inset). B: noise delay functions (NDFs) of the neurons from responses to binaural noise bursts with different ITDs for a typical neuron (the thickness of the line indicates the standard error of the mean response across 8 trials). C: From the NDFs, we extracted the best ITD, i.e. the ITD that evoked the highest spike rate. For our population of neurons, the distribution of best ITDs was clustered around the edge of the physiological limit for gerbils, corresponding to a sound located contralateral to the recording site (135 μs; (Maki & Furukawa, 2005)) with a median value of 101 μs.

3.3. Decoding using distance metrics

We define the set of responses from cell p in response to I trials of S different stimuli as r^{pSI}, where $S = \{1, 2, \ldots, S\}$ and $I = \{1, 2, \ldots, I\}$. To decode the response evoked by

trial i of stimulus s, r^{psi}, we remove it from the set and infer which stimulus evoked it, $\hat{s}(r^{psi})$. Assuming we have a metric for quantifying the distance between two responses, $d(r^{psi}, r^{ps'i'})$, then we can determine the average distance from r^{psi} to the responses evoked by all trials of a given stimulus s', $\bar{d}_{s'}(r^{psi}) = \langle d(r^{psi}, r^{ps'i'}) \rangle_{i'}$, with trial i excluded to avoid overfitting when s = s', and choose the stimulus for which this average distance is minimal, $\hat{s}(r^{psi}) = \arg\min_{s' \in S} \bar{d}_{s'}(r^{psi})$ [note that in the equation for $\bar{d}$, an exponent can be introduced inside the expectation to bias the result toward larger or smaller values]. This approach is easily extended to decode the responses r^{PSI} from a population of cells P = {1, 2, . . . , P}. To decode the responses from the population of cells P in response to trial i of stimulus s, r^{Psi}, we choose the stimulus for which a weighted sum of the average distances for each cell is minimal, $\hat{s}(r^{Psi}) = \arg\min_{s' \in S} \sum_{p \in P} w^p \bar{d}_{s'}(r^{psi})$. The central question in this section is how to choose the weights $w = [w^1, w^2, . . ., w^P]$ so as to maximize decoding performance.

3.4 Optimization of decoder weights

After decoding the spike trains for every trial of every stimulus as described above, we measure overall performance as the percent of spike trains that were correctly decoded and denote this quantity as PCp for a single cell p, and PCP(w) for the population P with weights w. The standard approach to finding the optimal set of weights, i.e. the set of weights that maximize PCP(w), is to calculate the gradient dPCP(w)/dw and use it as a guide toward a local, and hopefully global, maximum. However, for the particular problem considered here, analytical specification of the gradient was not possible and algorithms that calculated the gradient numerically performed very poorly. Fortunately, there is another class of algorithms known as 'evolutionary' that do not require knowledge of the gradient. These algorithms operate iteratively, choosing the best of several candidate solutions on each iteration until performance saturates. While there are many evolutionary algorithms that may be suitable for this particular problem, we chose to implement two of the most common, genetic and particle swarm. As illustrated in the examples below, the performance of these algorithms was similar. However, the genetic algorithm was superior in that it required less computation time and is easily implemented via the Genetic Algorithm and Direct Search Toolbox in Matlab (The Mathworks, USA), and, thus, we describe only its implementation in detail here. Details of the particle swarm

algorithm can be found in (Kennedy et al., 2001). The genetic algorithm for optimization begins by creating a population of y random vectors of length p, drawn with uniform probability from the interval [0, 1], and computing PCP(w) for each vector. Next, the population is evolved in three steps: First, e 'elite' vectors, those with the highest PCP(w), are moved to the next generation. Next, x 'crossover' vectors are created by random recombination between two 'parent' vectors from the current population, with the probability of a particular vector being chosen as a parent proportional to its PCP(w). Finally, u 'mutant' vectors are created by adding random noise $n \sim N(0, \sigma)$ to a parent vector, with parent vectors chosen as above. The standard deviation of the noise $\sigma = 1$ for the first generation and is decreased linearly with each successive generation such that $\sigma = 0$ if the algorithm runs to completion. The algorithm stops after either completing V evolution generations or when the change in the highest PCP(w) over the past G generations is less than ε. The set of weights with the highest PCP(w) after the completion of the algorithm is denoted $w_{genetic}$. For the examples in this study, $y = 25$, $e = 2$, $x = 18$, $u = 5$, $V = 100$, $G = 25$, and $\varepsilon = 10^{-5}$, in accordance with the suggested default parameters for the ga function in Matlab. As a baseline for comparison with $w_{genetic}$, we also computed the optimal weights w_{swarm} via particle swarm optimization, and used two other simple weighting schemes: $w_{equal} = 1$, where all cells are weighted equally, and $w_{percorr} = [PC1, PC2, \ldots, PCp]$, where the weights are determined by the overall performance of each cell when its responses are decoded individually. To prevent overfitting, it is important to exclude the responses to be decoded when optimizing the weights. For all optimizations, we split the responses into successive training sets (95% of responses) and testing sets (5% of responses) such that all responses were included in the testing set exactly once.

3.5. Decoding spike trains for revealing the effects of ITD and ILD change on coding

We decoded spike trains (i.e., used the spike trains to infer the sound that evoked them) using the metric introduced by (Victor & Purpura, 1996), which measures the distance between two spike trains as the overall cost of the set of operations required to transform one spike train into the other, with possible operations including the insertion of a spike, the deletion of a spike, and the time-shift of a spike (software available at http://www-users.med.cornell.edu/~jdvicto/spkdm.html). By changing the cost of time-shifting a spike relative to deleting the spike at one time and inserting it at another, the metric can

be used to evaluate the distance between spike trains at different timescales. The implementation of the metric followed the one described in (Victor & Purpura, 1996). We also decoded spike trains using the metric of (van Rossum, 2001), but, as decoder performance was similar with both metrics, only results from decoding with the Victor and Purpura metric are shown in the Results, below. Decoding using this metric was performed in three steps. First, a single spike train was removed from the full set of all spike trains. Second, the distance between the removed spike train and each of the remaining spike trains in the set was computed. Third, the removed spike train was assigned to the sound for which its average distance to the remaining spike trains evoked by that sound was smallest. This process was repeated for all spike trains in the set to obtain an overall percentage correct. For population spike trains, the distances for individual neurons were summed before decoding. The sound tokens used for the decoding analysis were 250 ms in duration, but, in all cases, responses to the first 50 ms were discarded because many neurons responded strongly to the onset of all of the different tokens and, thus, decoding token identity based on this portion of the response was not possible. For testing the significance of tuning to intensity, ITD, and token identity in single neurons as described below, responses to the remaining 200 ms of sound were used. For testing the effects of changes in ITD and intensity on decoding of token identity in single neurons, the duration that yielded decoder performance of ~ 50% correct for the base condition was determined individually for each neuron, and the same duration was used for the ITD and SPL conditions. For testing the effects of changes in ITD and intensity on decoding of token identity in populations, responses from 50 to 65 ms after sound onset were used for the analysis without background noise and responses from 50 to 200 ms were used for the analysis with background noise.

3.6. Evaluating the significance of tuning

The significance of each neuron's tuning to sound intensity, ITD, and token identity was determined by comparing decoder performance on the actual responses to performance after randomly reassigning the stimulus value associated with each response. Decoding was performed on 100 different sets of randomized responses and the significance threshold was defined as 2 SDs above the mean percentage correct for the randomized sets. For evaluating the significance of ITD and intensity tuning, responses to all tokens for each intensity and ITD were combined and decoding was based only on spike rate. To

be included in the full analysis comparing the effects of changes in ITD and intensity on the decoding of token identity, a neuron had to be significantly tuned to changes in ITD at all intensities and significantly tuned to changes in intensity at all ITDs. For evaluating the significance of tuning to token identity, decoding was performed at a range of timescales as described above. To be included in the full analysis, a neuron had to be significantly tuned to token identity at all ITDs and intensities for at least one timescale.

3.7. Calculating signal-to-noise ratio

SNR of responses (in 1 ms bins) was calculated as described by (Borst & Theunissen, 1999). First, the signal spectrum was obtained by computing the power spectrum of the response after averaging across all trials. Next, to obtain the noise power, the response from each trial was subtracted from the average response and the power spectrum of this difference was computed. These difference spectra were averaged over all trials to yield the overall noise spectrum. The SNR at each frequency was defined as the ratio of the power of the signal and noise spectra at that frequency and the total SNR was defined as the ratio of the sum of the power of the signal and noise spectra over all frequencies (i.e., the ratio of the variances of the signal and noise).

4. Results

4.1 Decoding with distance metrics

4.1.1 Decoding experimental spike trains

To illustrate the utility of the optimization algorithm under experimental conditions, we decoded spike trains from P = 34 cells recorded in the inferior colliculus of anesthetized gerbils in response to the presentation of I = 20 trials of S = 8 different sounds (different instances of Gaussian white noise). We decoded the spike trains using the metric introduced by (Victor & Purpura, 1996), which measures the distance between two spike trains as the overall cost of the set of operations required to transform one spike train into the other, with possible operations including the insertion of a spike, the deletion of a spike, and the time-shift of a spike. By changing the cost of time-shifting a spike relative to deleting the spike at one time and inserting it at another, the metric can be used to evaluate the distance between spike trains at different timescales. The details of the implementation of the metric are not given here but can be found in (Aronov et al., 2003; Victor & Purpura, 1996). A 10 ms segment of the set of spike trains r^{pSI} for a typical cell is shown in Figure 4.1.1 along with the decoding performance for the individual responses of each cell as a function of the response timescale parameter of the decoder (decoded responses were 100 ms in duration). The median best timescale, i.e., the timescale that yielded the best decoding performance, across the sample of cells was 2 ms (black arrow) and, for simplicity, we fixed the response timescale at this value for all decoding of these responses. The distribution of significant pairwise correlation coefficients for the population is shown in Figure 4.1.2 ($P = 34$ cells). Correlation coefficients were estimated after converting the spike trains to binary vectors with a temporal resolution of 2 ms. Only significant correlations ($p < 0.05$) are shown. The total correlation (ρ_{total}) was computed directly from the responses, the signal correlation (ρ_{signal}) was computed from the responses after randomizing the trial order, and the noise correlation (ρ_{noise}) was computed as the difference between ρ_{total} and ρ_{signal}. The total correlation (ρ_{total}) was significant between approximately half of the cell pairs (262 of 561).
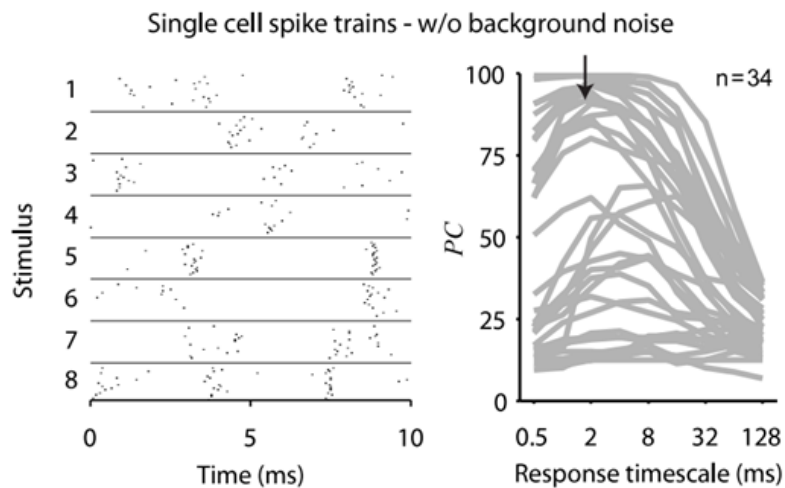
Figure 4.1.1. Left: raster plot showing the spike trains of a neuron in the inferior colliculus of an anesthetized gerbil in response to I = 20 trials of S = 8 different sounds. Right decoder performance (percent correct, PC) as a function of decoder response timescale for the responses of each individual cell. The black arrow indicates the population median best timescale, 2 ms.
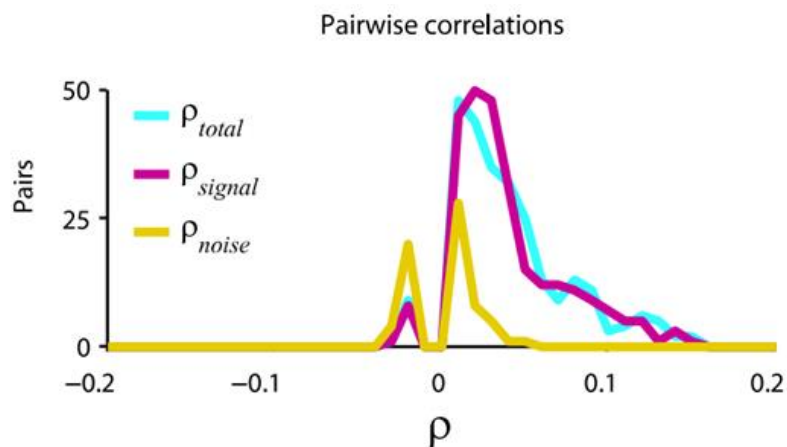

Figure 4.1.2. Histograms of the correlation coefficients between pairs of cells (P = 34). Correlation coefficients were estimated after converting the spike trains to binary vectors with a temporal resolution of 2 ms. Only significant correlations (p < 0.05) are shown. The total correlation (ρ_{total}) was computed directly from the responses, the signal correlation (ρ_{signal}) was computed from the responses after randomizing the trial order, and the noise correlation(ρ_{noise}) was computed as the difference between ρ_{total} and ρ_{signal}.
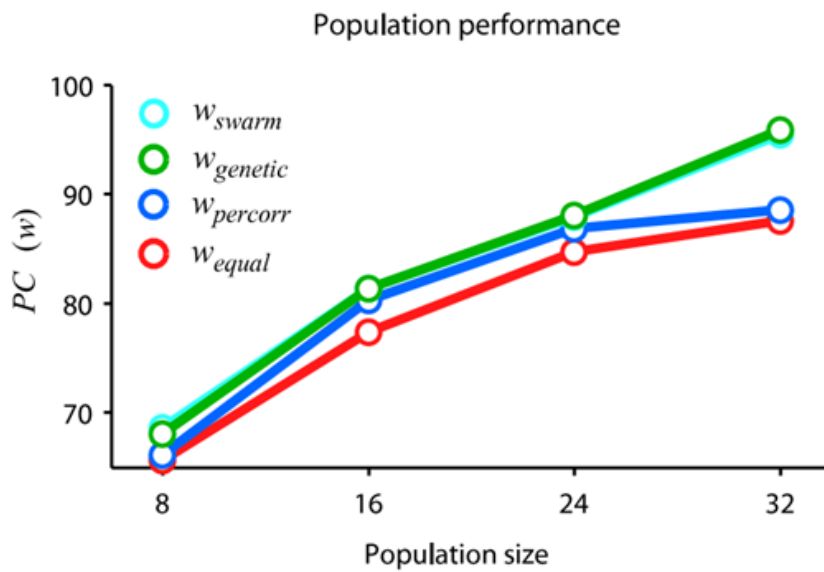
Figure 4.1.3. The decoder performance PC$_P$(w) for subpopulations of increasing size for four sets of weights: w$_{equal}$ (all weights equal), w$_{percorr}$ (weights based on individual performance), w$_{genetic}$ (weights optimized with genetic algorithm), and w$_{swarm}$ (weights optimized with particle swarm algorithm). The circles and bars indicate the mean and standard error of the performance for 100 different random subpopulations.
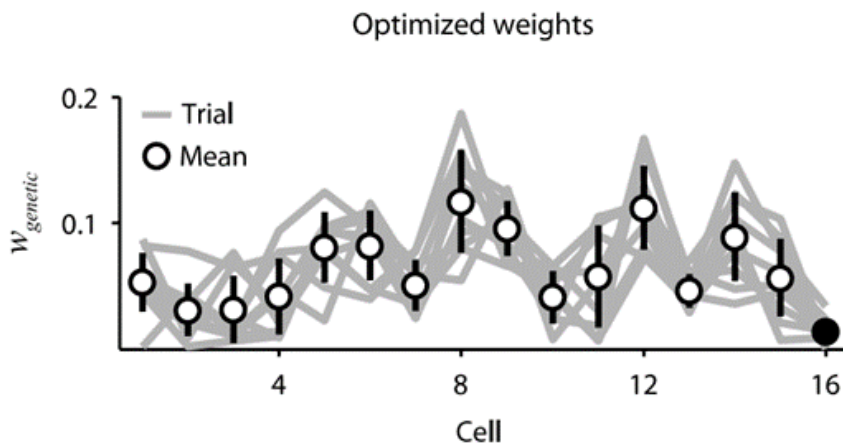

Figure 4.1.4 The weights w$_{genetic}$ resulting from 10 different optimizations with random initial values in the population y for a particular subpopulation of 16 cells. The lines indicate the weights for each individual optimization and the circles and bars indicate the mean and standard deviation. The dummy cell is indicated by the filled black circle.

Correlations between cells can have both signal and noise components: signal correlations arise from correlations in the stimulus itself and/or similarity in preferred stimulus features (frequency, orientation, etc.), while noise correlations arise from shared inputs that contribute to the trial-to-trial variability in responses. In this set of responses, most of the total correlation was due to signal correlations (in fact, because most of the cells were not recorded simultaneously, the population is expected to contain few noise correlations). The decoder performance $PC_P(w)$ for subpopulations of increasing size is shown in Figure 4.1.3. for four sets of weights: w_{equal} (all weights equal), $w_{percorr}$ (weights based on individual performance), $w_{genetic}$ (weights optimized with genetic algorithm), and w_{swarm} (weights optimized with particle swarm algorithm). For population decoding, only 10 ms segments of the responses were used in order to increase the difficulty of the decoding task. While the decoder performance was similar for all sets of weights for small population sizes, genetic and particle swarm optimization provided a performance increase of approximately 10% for large populations. To determine whether optimization produced global optima, we used the genetic algorithm to find the optimal weights for a given set of responses using 10 different initial populations y. In each set of responses, we also included a 'dummy cell' for which the stimulus identity associated with each response was randomized. As illustrated in Figure 4.1.4. for a particular subpopulation of P = 16 cells, the genetic algorithm converges to approximately the same set of optimal weights $w_{genetic}$ independent of the initial values in the population y and the weights associated with the dummy cell (filled black circle) were always near zero. These results suggest that the set of weights produced by the genetic algorithm is indeed the global optimum and that the algorithm is successful in minimizing the contribution of uninformative cells. For the same cells, we also recorded responses to the same sounds in the presence of background noise. Because a different background noise was added on each trial, and this noise was the same for all cells, the background noise served to reduce signal correlations and introduce noise correlations. The set of spike trains r^{pSI} for a typical cell is shown in Figure 4.1.5. along with the decoding performance for the individual responses of each cell as a function of the response timescale (decoded responses were 100 ms in duration). While the background noise resulted in considerable variability in the spike trains evoked by the same stimulus, the median best timescale was again 2 ms and we fixed the response timescale at this value for all decoding of these

responses. The distribution of significant pairwise correlation coefficients for the population is shown in Figure 4.1.6. Again, the total correlation (ρ_{total}) was significant between approximately half of the cell pairs (299 of 561), but for this set of responses, most of the total correlation was due to noise correlations. The decoder performance for the population $PC_P(w)$ for subpopulations of increasing size is shown in Figure 4.1.7. for w_{equal}, $w_{percorr}$, $w_{genetic}$, and w_{swarm}. For population decoding, only 30 ms segments of the responses were used in order to increase the difficulty of the decoding task. As in the previous example, performance was similar for all sets of weights for small population sizes, but optimization provided a substantial performance increase for large populations and, as shown in Figure 4.1.8., optimizations with different initial values in the population y produced similar sets of weights with values near zero for the dummy cell. Taken together, the results in Figures 4.1.1-8. demonstrate that the optimization algorithm was effective for decoding population spike trains under experimental conditions when the responses contained both signal and noise correlations.
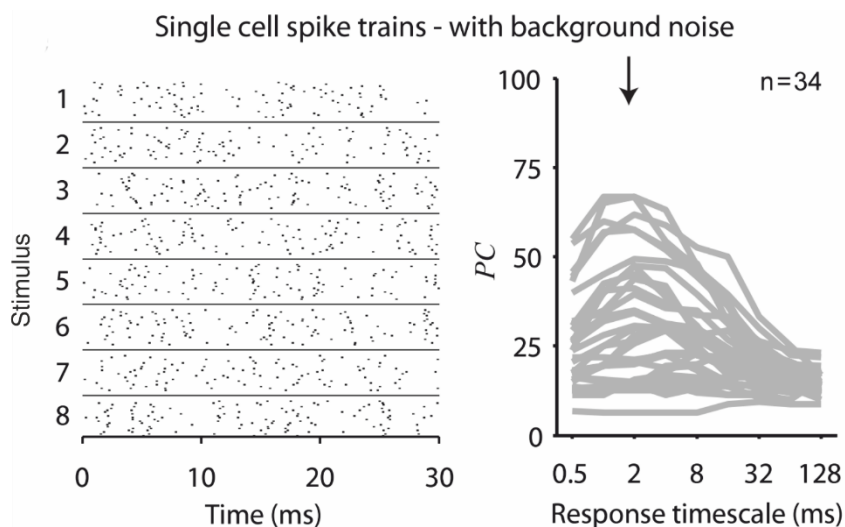

Figure 4.1.5. Left: raster plot showing the spike trains of a neuron in the inferior colliculus of an anesthetized gerbil in response to I = 20 trials of S = 8 different sounds. Right decoder performance (percent correct, PC) as a function of decoder response timescale for the responses of each individual cell. The black arrow indicates the population median best timescale, 2 ms. A different random background noise was added to each sound on each trial.
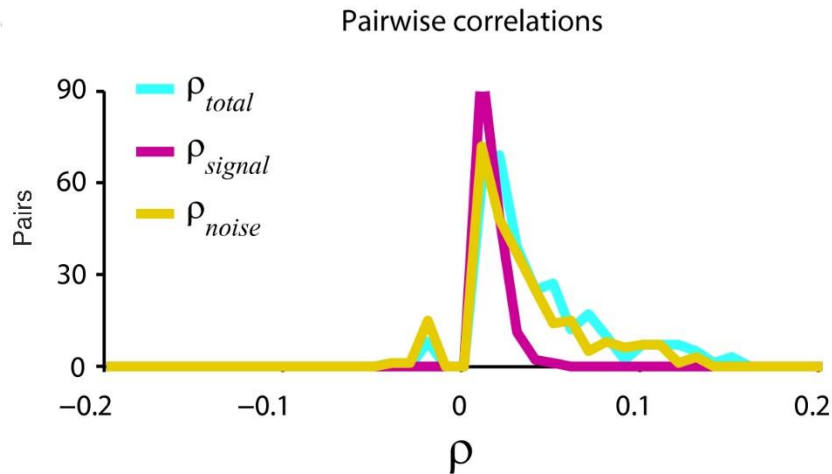
Figure 4.1.6. Histograms of the correlation coefficients between pairs of cells (P = 34) with added background noise.
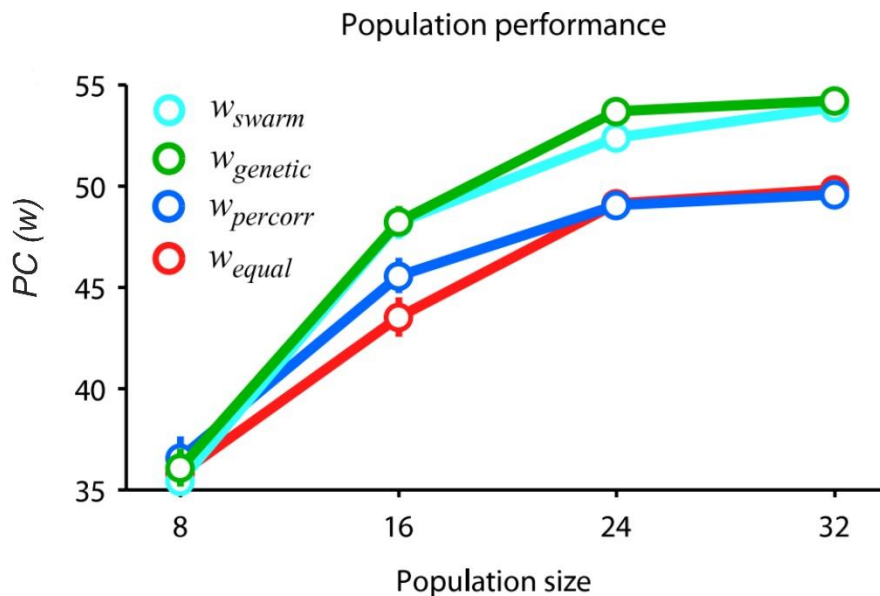

Figure 4.1.7. The decoder performance $PC_P(w)$ for subpopulations of increasing size for four sets of weights: w_{equal} (all weights equal), $w_{percorr}$ (weights based on individual performance), $w_{genetic}$ (weights optimized with genetic algorithm), and w_{swarm} (weights optimized with particle swarm algorithm). The circles and bars indicate the mean and standard error of the performance for 100 different random subpopulations. Background noise was added to each sound on each trial.
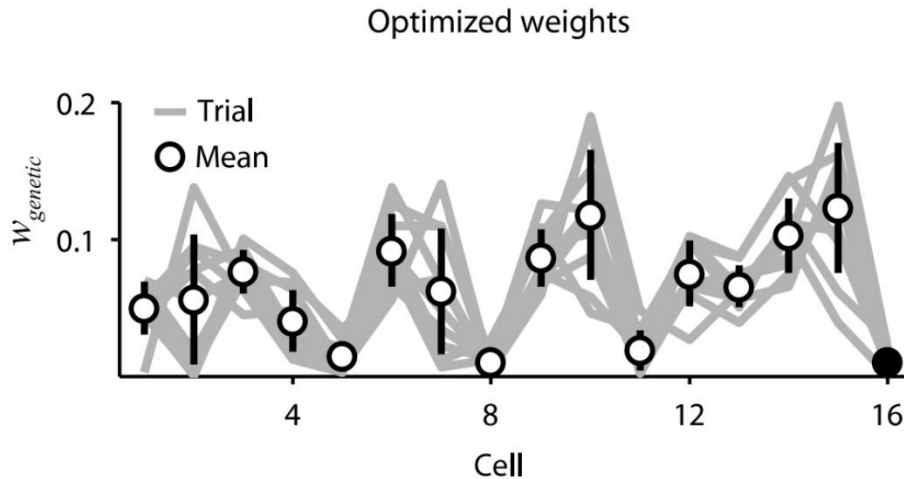
Optimized weights

Figure 4.1.8. The weights w$_{genetic}$ resulting from 10 different optimizations with random initial values in the population y for a particular subpopulation of 16 cells. The lines indicate the weights for each individual optimization and the circles and bars indicate the mean and standard deviation. The dummy cell is indicated by the filled black circle. Background noise was added to each sound on each trial.

4.1.2. Decoding experimental calcium signals

To further illustrate the utility of optimization under experimental conditions, calcium signals (relative change in indicator fluorescence) were also decoded from P = 37 cells recorded in the visual cortex of anesthetized mice in response to the presentation of I = 18 trials of S = 8 different oriented sinusoidal gratings (each grating was displayed for 2 s at 50% contrast and drifted at a rate of 2 Hz; calcium signals were sampled at 15 Hz). Although I did not analyze these signals, that come from an experiment described in (Mrsic-Flogel et al., 2007), the following results add notable information to the decoder optimization. The set of calcium signals r^{pSI} for a typical cell is shown in Figure 4.1.9. The top image gives an overview of the dynamics and reproducibility of the signals as the orientation of the grating changed (the order of the orientations was the same on each trial), while the lower plots show the signals in detail for two particular orientations. The distribution of significant pairwise correlation coefficients for the population (computed at a timescale of 66 ms) is shown in Figure 4.1.10. The total correlation (ρ_{total}) was significant between most of the cell pairs (560 of 666) and contained both signal and noise components (all cells were recorded simultaneously). We decoded the calcium signals

51

using the Euclidean distance metric $d(r^{psi}, r^{ps'i'}) = |r^{psi} - r^{ps'i'}|$. The decoder performance PC$_P$(w) for subpopulations of increasing size is shown in Figure 4.1.11. for w$_{equal}$, w$_{percorr}$, and w$_{genetic}$, w$_{swarm}$. As with spike trains, performance was similar for all sets of weights for small populations, but optimization improved performance for large populations and, as shown in Figure 4.1.12., optimizations with different initial values in the population y produced similar sets of weights with values near zero for the dummy cell. These results demonstrate that the optimization algorithm was effective for decoding not only population spike trains, but also population calcium signals with signal and noise correlations.
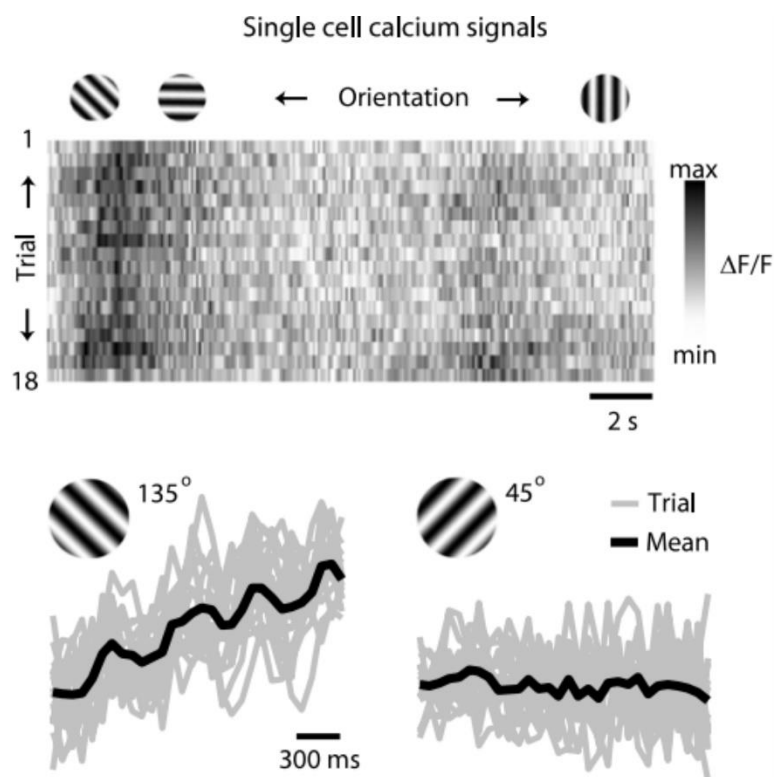

Figure 4.1.9. Top: an image showing the calcium signal (relative change in indicator florescence) of a neuron in the visual cortex of an anesthetized mouse in response to I = 18 trials of S = 8 different oriented gratings. Bottom: the calcium signals for the same neuron in response to two particular orientations. Gray lines indicate the signal for each trial and the black line indicates the mean.
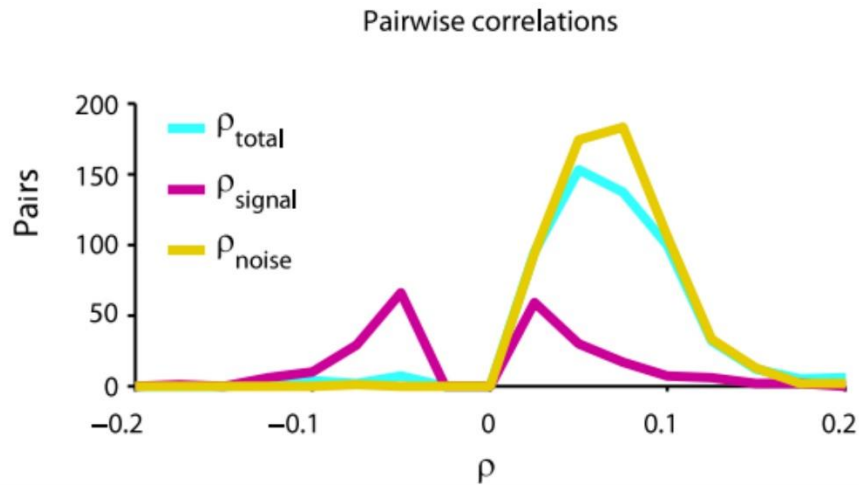
Pairwise correlations

Figure 4.1.10. Histograms of the correlation coefficients between pairs of cells (P = 37).
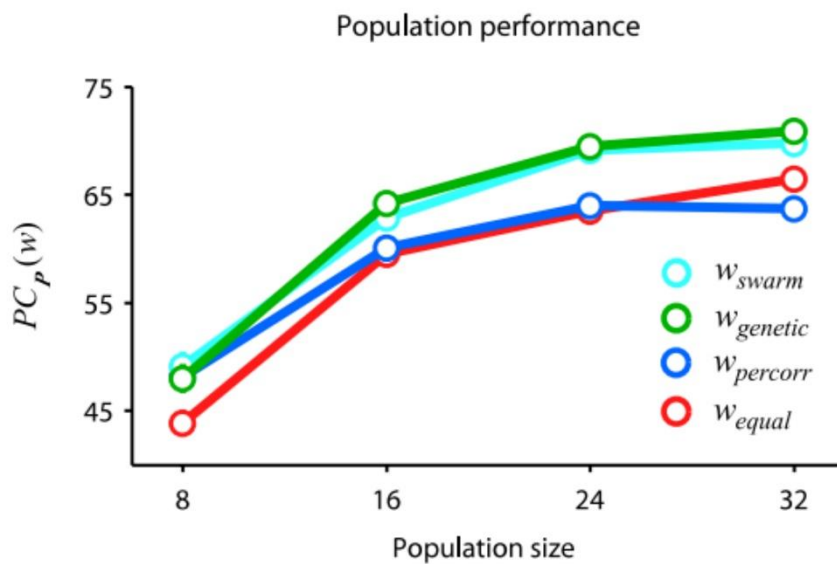
Population performance

Figure 4.1.11. Decoder performance $PC_P(w)$ for subpopulations of increasing size for four sets of weights: w_{equal}, $w_{percorr}$, $w_{genetic}$, and w_{swarm}.
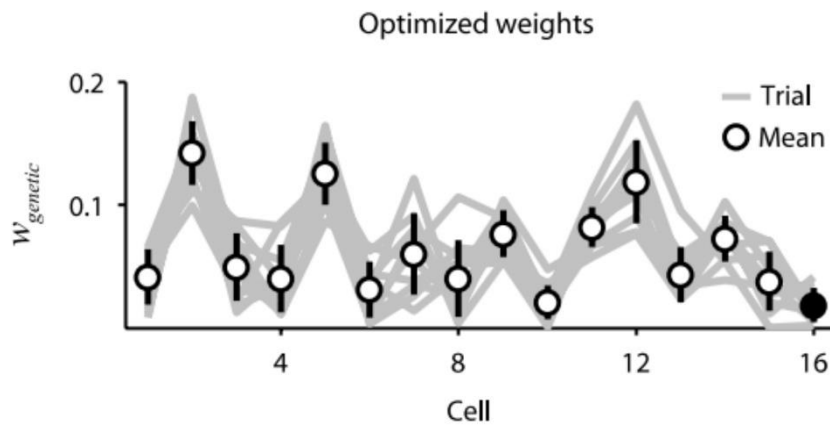
Figure 4.1.12. The weights $w_{genetic}$ resulting from 10 different optimizations with random initial values in the population y for a particular subpopulation of 16 cells. The lines indicate the weights for each individual optimization and the circles and bars indicate the mean and standard deviation. The dummy cell is indicated by the filled black circle.

4.2. Revealing the effects of interaural time and intensity difference on the coding of low-frequency sounds

To investigate the influence of intensity and ITD on the ability of auditory neurons to encode low-frequency sounds, we made extracellular single-unit recordings from the central nucleus of the IC in anesthetized gerbils using a multielectrode array. Recordings were made in the low-frequency lamina of the rostrolateral quadrant of the IC, where inputs from the MSO are clustered (Cant & Benson, 2006) and cells are likely to be ITD-sensitive. Because these cells are sensitive only to low frequencies, ITD is the only available cue for azimuthal angle (Maki & Furukawa, 2005). Of our original population of 55 neurons, we analyzed only the 33 that were significantly tuned to changes in intensity, ITD, and sound token identity. All of these neurons had significant sustained responses to broadband binaural sounds (spike rates between 50 and 100 ms after sound onset were greater than spontaneous spike rates; Wilcoxon rank-sum tests, $p < 0.05$). The distributions of preferred frequencies and ITDs for the population are shown in Figure 3.2.1.

4.2.1. Changes in intensity and ITD influence the precision and timing of spike events

To determine the effects of changes in intensity and ITD on the ability of single neurons to encode low-frequency sounds, we analyzed responses to 20 repeated trials of eight different sound tokens (Gaussian noise bandpass filtered between 200 and 4000 Hz) at three different intensities (43, 63, and 83 dB SPL) and five different ITDs (evenly spaced between ±135 µs, spanning the physiological range for gerbils (Maki & Furukawa, 2005); positive ITDs indicate that the sound reached the ear contralateral to the recording site first), for a total of 15 different intensity/ITD combinations. The different tokens reliably evoked different spike patterns, as illustrated in the responses of a typical neuron to sounds presented at 63 dB SPL with 0 µs ITD shown in Figure 4.2.1.
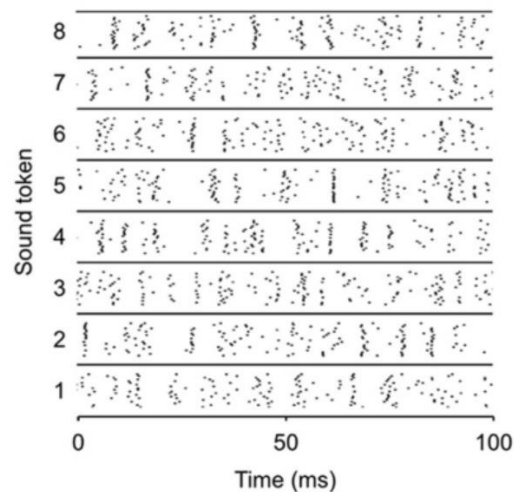

Figure 4.2.1. A raster plot showing the spike trains recorded from a typical neuron in response to 20 repeated presentations of eight different sound tokens presented at 63 dB SPL with 0 µs ITD.

To study the effects of changes in intensity and ITD beyond those that result from changes in overall spike rate, we analyzed only responses from those neurons for which we found a decrease in intensity and a negative change in ITD that caused approximately the same decrease in spike rate relative to an arbitrary base condition (the base condition could be any intensity/ITD combination and was chosen independently for each cell) (Figure 4.2.2.).
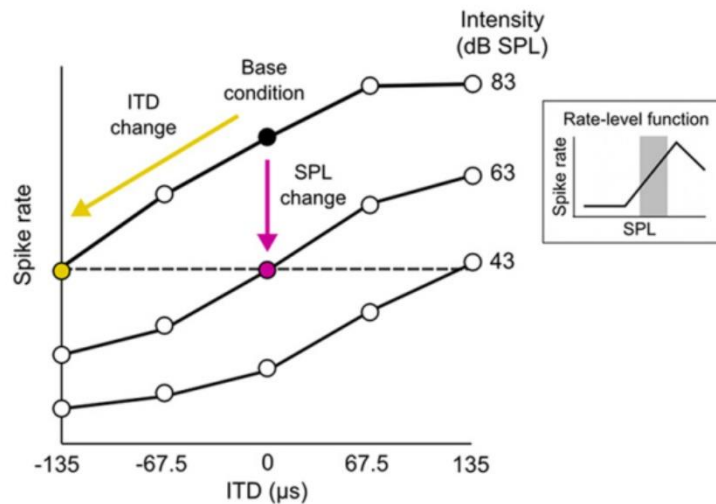
Figure 4.2.2. A schematic diagram depicting the relationship between the three stimulus conditions: the base condition, the ITD change (ΔITD; a decrease in ITD) condition, and the intensity change (ΔSPL; a decrease in intensity) condition. Only those cells for which the ΔSPL condition could be defined by a change in intensity along the positive slope of the RLF were analyzed.

Because our sampling of the space of possible intensity/ITD combinations was relatively sparse, only 19 neurons satisfied this criterion (the reductions in spike rate for the intensity change (ΔSPL) and ITD change (ΔITD) conditions relative to the base condition for these neurons were not significantly different; paired Wilcoxon test, $p = 0.08$; median reduction was 29% for ΔSPL and 30% for ΔITD). The requirement that a decrease in intensity cause a decrease in spike rate ensured that the analysis was restricted to the range of intensities corresponding to the positive slope of the rate-level function (RLF; the function relating sound intensity to overall spike rate) (Figure 4.2.2, inset), even for neurons with nonmonotonic RLFs (analysis of responses from the negative slope of the RLF are presented below). The responses of a typical neuron to one sound token for the base, ΔSPL, and ΔITD conditions are shown in Figure 4.2.3, left. The changes in intensity and ITD had similar effects on the spike rate, but they had different effects on the timing of spikes within the response. Relative to the base condition, the change in intensity caused a change in the overall timing of events, but had little impact on precision of spike timing across trials, whereas the change in ITD caused a decrease in precision of spike timing across trials but left the overall timing of spike events largely unchanged.
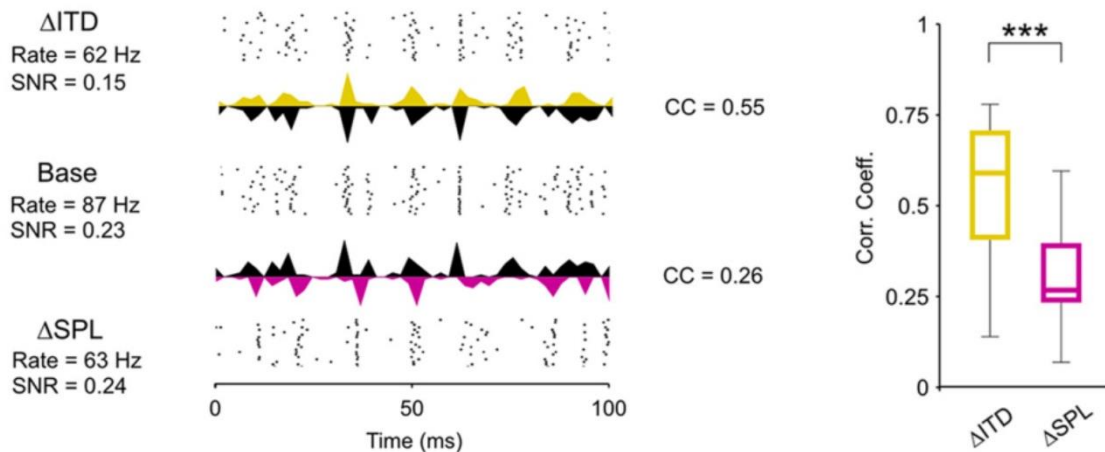
56

Figure 4.2.3. Left: Raster plots and PSTHs showing the responses of a typical neuron to the same sound token for the three conditions. The mean overall spike rates and response SNRs are shown for each condition, and the CCs between the PSTHs for the base condition and each of the two change conditions are shown. Different PSTHs extend upward and downward from the same axis for ease of visual comparison. Right: Boxplots showing the distribution of CCs between the PSTHs for the base condition and each of the two change conditions for a sample of 19 neurons. In each plot, the central mark indicates the median, the edges of the box indicate the 25th and 75th percentiles, and the error bars extend to the most extreme values. The results of paired Wilcoxon tests comparing the medians of the distributions are indicated. ***p < 0.001

To quantify the effects of changes in intensity and ITD on the overall timing of events, we measured the correlation coefficient (CC) between the peristimulus time histograms (PSTHs) (in 1 ms time bins) for each of the change conditions and the base condition. As shown in Figure 4.2.3, right, across our sample of neurons, the CC between responses for the ΔITD and base conditions were significantly larger than those between responses for the ΔSPL and base conditions (paired Wilcoxon test, p < 0.001). To quantify the effects of changes in intensity and ITD on precision, we measured the SNR of the responses. SNR, a measure commonly used to describe the precision of spike trains in early sensory systems, compares the power in the part of the response that is repeatable from trial to trial (the PSTH) with that which is variable from trial to trial (the deviation from the PSTH on each trial).

As shown in Figure 4.2.4, across our sample of neurons, the change in ITD caused a significant decrease in SNR relative to the base condition, whereas the change in intensity had no significant effect (paired Wilcoxon tests, $p < 0.001$ for ITD and $p = 0.18$ for SPL). These results suggest that changes in intensity and ITD have different effects on the timing of spikes: a change in intensity causes a change in the overall timing of spike events, whereas a change in ITD causes a change in the precision of spike timing across trials.
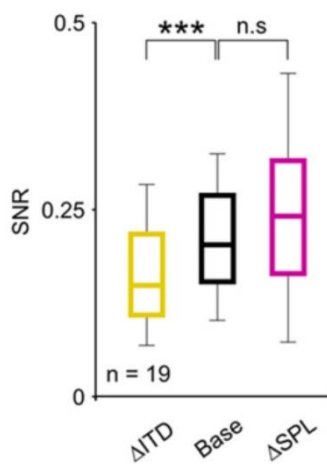
*Figure 4.2.4. Boxplots showing the distribution of response SNRs for each of the three stimulus conditions. In each plot, the central mark indicates the median, the edges of the box indicate the 25th and 75th percentiles, and the error bars extend to the most extreme values. The results of paired Wilcoxon tests comparing the medians of the distributions are indicated. n.s., Not significant., ***p < 0.001*

4.2.2. Changes in ITD, but not intensity, influence decoder performance

To determine the impact of the observed effects of changes in intensity and ITD on coding, we used a decoder to infer which sound token evoked each response for each stimulus condition. The performance of the decoder for a given condition reflects how well information about token identity is encoded in the spike trains for that condition; if the spike trains evoked by a given token are similar to each other, but different from the spike trains evoked by the other tokens, then the decoder will correctly assign the spike trains to the tokens that evoked them (note that this approach is different from training the decoder for one condition and testing its performance for a different condition to examine the degree of invariance in how the information is encoded (Billimoria et al., 2008)). Because the tokens were exactly the same for each condition, the difference in the performance of the decoder for the three stimulus conditions provides a direct measure of the effects of changes in intensity and ITD on coding. The decoder was based on a

metric that computes the distance between two spike trains at a specified timescale (Victor & Purpura, 1996). To decode a given spike train, the decoder measured its distance to all of the other spike trains evoked by each sound token and chose the token for which the mean distance was smallest. This decoder is not designed to mimic the function of a neuron in any particular downstream auditory area, but simply to serve as tool for assessing how well information about token identity is encoded in IC responses. The performance of the decoder at different timescales for a typical neuron for the base, ΔSPL, and ΔITD conditions is shown in Figure 4.2.5, left.
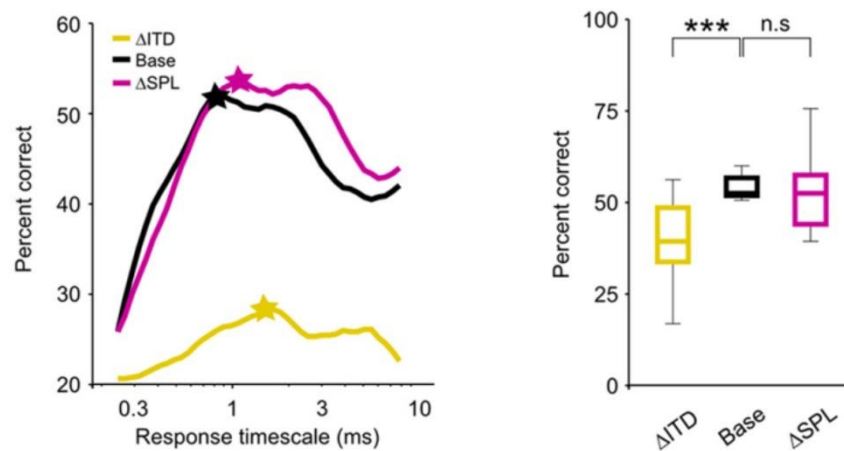

Figure 4.2.5. Left: Decoder performance as a function of response timescale for a typical neuron under the three stimulus conditions. The stars indicate the time scale corresponding to the best performance. Right: Boxplots showing the distribution of decoder performance at the optimal timescale for each of the three stimulus conditions. In each plot, the central mark indicates the median, the edges of the box indicate the 25th and 75th percentiles, and the error bars extend to the most extreme values. The results of paired Wilcoxon tests comparing the medians of the distributions are indicated. Chance level performance was 12.5%. n.s., Not significant., ***p < 0.001

For this neuron, the decoder performance was unaffected by the change in intensity but was severely degraded by the change in ITD. As shown in Figure 4.2.5, right, across the sample of neurons, the change in ITD caused a significant decrease in decoder performance (at the timescale for which performance was maximal for each neuron) relative to the base condition, whereas the change in intensity had no significant effect (paired Wilcoxon tests, p < 0.001 for ΔITD and p = 0.27 for ΔSPL). Thus, the change in

the precision of spike timing across trials caused by a change in ITD had a strong effect on the ability of IC neurons to encode low-frequency sounds, whereas the change in the overall timing of spike events caused by a change in intensity did not (at least for the restricted range of intensities corresponding to the positive slope of the RLF; see below for the results for intensities corresponding to the negative slope of the RLF).

4.2.3 Changes in intensity, but not ITD, evoke a change in spectrotemporal filtering

Previous studies have demonstrated that changes in intensity evoke a shift in spectral and/or temporal filtering properties that may help preserve the flow of auditory information in the face of changes in the SNR of incoming sounds (Lesica & Grothe, 2008b; Nagel & Doupe, 2006; Rees & Møller, 1987). For example, as intensity is decreased, temporal filtering shifts toward low frequencies where the SNR in natural sounds is likely to be the largest (Lesica & Grothe, 2008b; Singh & Theunissen, 2003).To investigate whether such shifts could account for the differences in the effects of changes in intensity and ITD on the coding of low-frequency sounds illustrated above, we measured the SNR as a function of response frequency for the neurons in our sample. Because the sound tokens were identical for all three stimulus conditions and were uncorrelated (i.e., had equal power at all frequencies), the SNR at each response frequency is a direct reflection of the net effect of the spectrotemporal filtering properties of the system (note that we use the term spectrotemporal filtering because the neurons in our sample have low preferred frequencies and response power at a given frequency can reflect filtering of both envelope and fine structure). The mean SNR as a function of response frequency for the base, ΔSPL, and ΔITD conditions for our sample of neurons are shown in Figure 4.2.6, normalized such that the area under each curve is the same to compensate for the overall differences in SNR described above. As expected, the change in intensity caused a clear shift toward low response frequencies relative to the base condition. In contrast, the SNRs as functions of response frequency for the base and ΔITD conditions were nearly identical, indicating that the change in ITD did not evoke a shift in spectrotemporal filtering properties. Thus, the system appears to shift its spectrotemporal filtering properties to preserve the ability of single neurons to encode low-frequency sounds in response to changes in intensity, but not in response to changes in ITD.
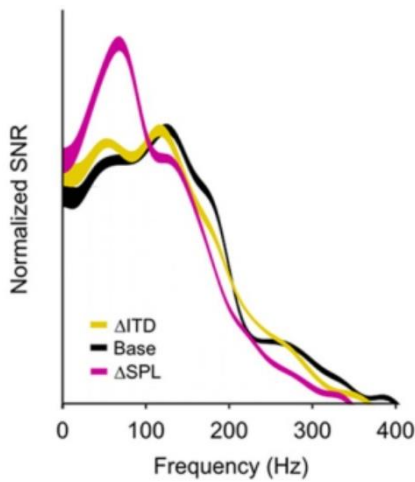
Figure 4.2.6. The SNR as a function of response frequency under the three stimulus conditions, averaged across all cells in the sample and normalized such that the area under each curve is the same. The thickness of the lines indicates the SEM. The change in intensity caused a clear shift toward low response frequencies relative to the base condition. In contrast, the SNRs as functions of response frequency for the base and ΔITD conditions were nearly identical, indicating that the change in ITD did not evoke a shift in spectrotemporal filtering properties. Thus, the system appears to shift its spectrotemporal filtering properties to preserve the ability of single neurons to encode low-frequency sounds in response to changes in intensity, but not in response to changes in ITD.

4.2.4. Coding is robust to changes in ITD at the population level

The results described above demonstrate that changes in intensity and ITD have different effects on the coding of low-frequency sounds in the responses of single neurons. However, these changes also have different effects on the overall spike rates of the entire population. For example, an increase in intensity will cause, on average, an increase in spike rate for the whole population (except, perhaps, at very high intensities). In contrast, because most binaural neurons with low preferred frequencies (including all in this study) respond most strongly to sounds located on the side contralateral to the brain hemisphere that they are in (corresponding to positive ITDs in this study), a change in the ITD of a sound will cause, on average, an increase in spike rate for neurons in one hemisphere and a decrease in spike rate for neurons in the other hemisphere. To determine how changes in intensity and ITD influenced the coding of sound content at the population level, we decoded the responses of many different random subpopulations of neurons using the same metric as described above.
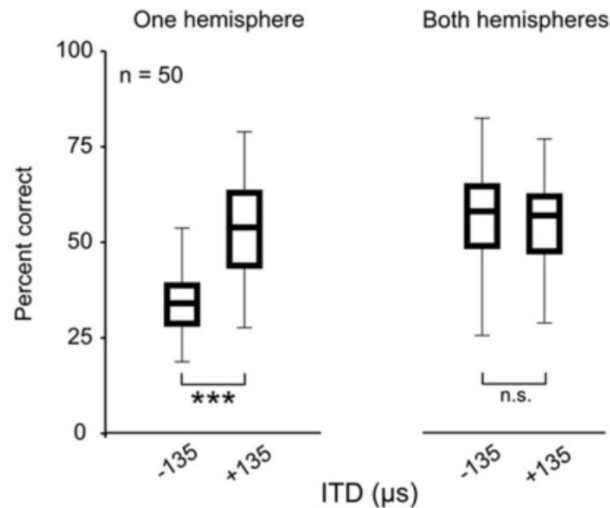
*Figure 4.2.7. Boxplots showing the distribution of decoder performance for 50 randomly chosen populations of 10 cells with either all cells from the same hemisphere or half of the cells from each hemisphere. Only the distributions for responses to sounds at 83 dB SPL are shown, but the distributions for other intensities were similar. All neurons were in fact recorded in the same hemisphere, but responses to sounds at -135and +135μs ITD were switched for half of the neurons to simulate responses from both hemispheres. n.s., Not significant. ***p < 0.001.*

As shown in Figure 4.2.7, when all of the cells in the population were taken from a single hemisphere, the change in ITD from + 135 to - 135 μs (corresponding to a change in location from the contralateral side to the ipsilateral side) caused a decrease in decoder performance similar to that observed in single cells (Wilcoxon test, p < 0.001, n = 50 different random subpopulations of 10 neurons). However, when half of the population was drawn from each hemisphere, the performance of the decoder was independent of ITD (Wilcoxon test, p = 0.96). These results suggest that opposing effects of a change in ITD in the two hemispheres offset each other; a change in ITD that degrades the coding of low-frequency sounds in one hemisphere enhances it in the other, such that there is no net change across the entire population.

4.2.5. The effects of intensity and ITD on coding are similar with and without background noise

Changes in listening conditions, such as the addition of background noise, have been shown to have strong effects on the processing of sound content in the IC (Kvale & Schreiner, 2004; Lesica & Grothe, 2008b; Rees & Palmer, 1988). To determine whether

the observed effects of changes in intensity and ITD on coding described above were also evident in the presence of background noise, we recorded responses of the same neurons to the same sound tokens in the presence of broadband background noise (SNR = 0 dB). The responses of a typical neuron to one sound with and without background noise presented at 63 dB SPL with 0 µs ITD are shown in Figure 4.2.8, left. Again, to study the effects of changes in intensity and ITD beyond those that result from changes in overall spike rate, we analyzed only those neurons for which we found a decrease in intensity and a negative change in ITD that caused approximately the same decrease in spike rate relative to an arbitrary base condition (n = 14). For this subset of neurons, the reductions in spike rate for the ΔSPL and ΔITD conditions relative to the base condition were not significantly different (paired Wilcoxon test, p = 0.54; median reduction was 31% for ΔSPL and 32% for ΔITD).
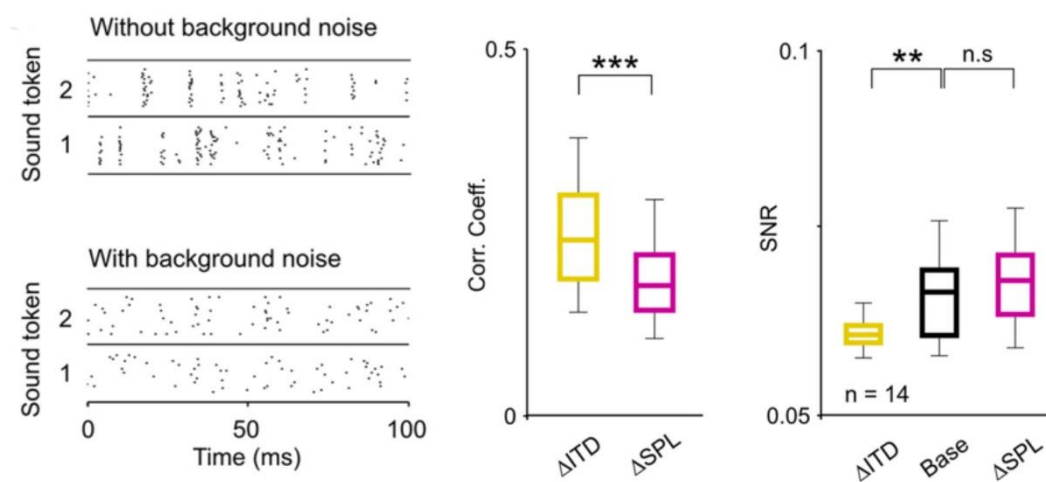
Figure 4.2.8. Left: A raster plot showing the spike trains recorded from a typical neuron in response to 20 repeated presentations of one sound token presented at 63dB SPL with 0 µs ITD with and without background noise at a signal-to-noise ratio of 0dB. Middle and right: Boxplots showing the distribution of CCs (middle) and response SNRs (right) between the PSTHs for the base condition and each of the two change conditions for a sample of 14 neurons. In each plot, the central mark indicates the median, the edges of the box indicate the 25th and 75th percentiles, and the error bars extend to the most extreme values. The results of paired Wilcoxon tests comparing the medians of the distributions are indicated. n.s., Not significant, **p < 0.01, ***p < 0.001 The results are similar as for the case without background noise (see Figures 4.2.3. and 4.2.4.).

The effects of changes in intensity and ITD on the timing of spike events with background noise were similar to those without; the CCs between responses for the ΔITD and base conditions were significantly larger than those between responses for the ΔSPL and base conditions (paired Wilcoxon test, $p < 0.001$) (Figure 4.2.8, middle) and the change in ITD caused a significant decrease in SNR relative to the base condition, whereas the change in intensity had no significant effect (paired Wilcoxon tests, $p = 0.002$ for ΔITD and $p = 0.24$ for ΔSPL) (Figure 4.2.8, right). The effects of changes in intensity and ITD on coding with background noise were also similar to those without; a change in ITD resulted in a significant decrease in decoder performance for single neurons relative to the base condition, whereas the change in intensity had no significant effect (paired Wilcoxon tests, $p = 0.02$ for ΔITD and $p = 0.61$ for ΔSPL) (Figure 4.2.9, left) and the effects of a change in ITD were canceled out at the population level when the population contained neurons from both hemispheres (Wilcoxon tests, $p < 0.001$ for one hemisphere, $p = 0.58$ for both hemispheres, n = 50 different random subpopulations of 10 cells) (Figure 4.2.9, right). The similarity of the results in Figures 4.2.1-4.2.9. suggest that, at least at a qualitative level, the effects of changes in ITD and intensity on the coding of low frequency sounds are independent of background noise level.
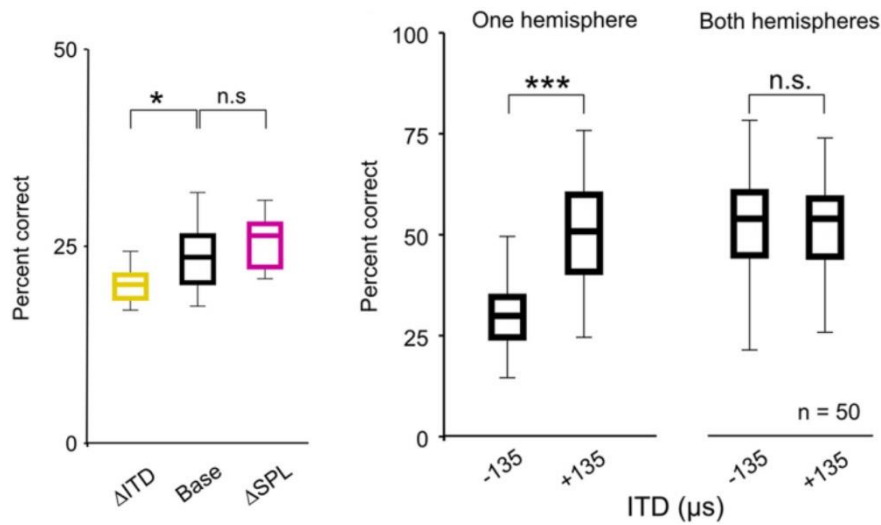
Figure 4.2.9. Left: Boxplots showing the distribution of decoder performance at the optimal timescale for each of the three stimulus conditions. In each plot, the central mark indicates the median, the edges of the box indicate the 25th and 75th percentiles, and the error bars extend to the most extreme values. The results of paired Wilcoxon tests comparing the medians of the distributions are indicated. Right: Boxplots showing decoder performance for randomly chosen populations with either all cells from the same hemisphere or half of the cells from each hemisphere. Chance level performance was 12.5%. n.s., Not significant, *p < 0.05, ***p < 0.001 The results are similar as for the case without background noise (see Figures 4.2.5. and 4.2.7.).

4.2.6. Changes in intensity along the positive and negative slope of the RLF have different effects on coding

Many neurons in the auditory system have RLFs that are nonmonotonic, i.e., spike rate increases with increasing intensity for soft sounds, but decreases with increasing intensity for loud sounds. To determine whether the observed effects of changes in intensity on coding differ depending on whether the changes are along the positive or negative slope of the RLF, we performed the same set of analyses on responses of those neurons (n = 13) for which we found an increase in intensity and a negative change in ITD that caused approximately the same decrease in spike rate relative to an arbitrary base condition (Figure 4.2.10, schematic diagram). For this subset of neurons, the reductions in spike rate for the ΔSPL and ΔITD conditions relative to the base condition were not significantly different (paired Wilcoxon test, p = 0.12; median reduction was 20 % for ΔSPL and 21 % for ΔITD).
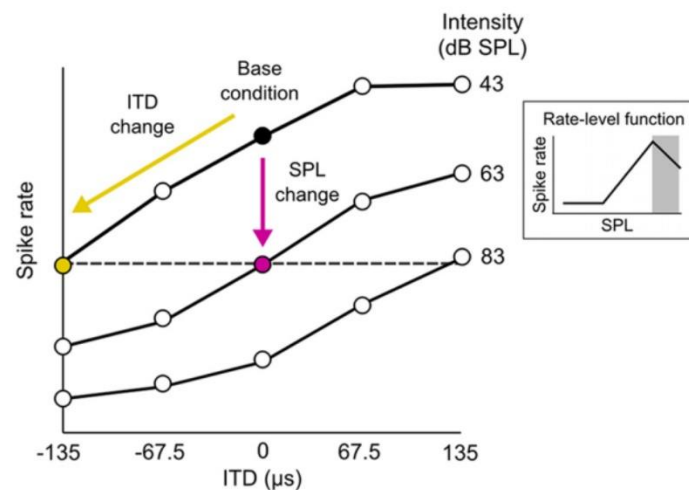
Figure 4.2.10. A schematic diagram depicting the relationship between the three stimulus conditions. Only those cells for which the ΔSPL (an increase in intensity) condition could be defined by a change in intensity along the negative slope of the RLF were analyzed.

As with changes in intensity along the positive slope of the RLF (Figure 4.2.3, right), changes in intensity along the negative slope of the RLF had a much stronger effect than changes in ITD on the overall timing of spike events; the CCs between responses for the ΔITD and base conditions were significantly larger than those between responses for the ΔSPL and base conditions (paired Wilcoxon test, p < 0.001) (Figure 4.2.11, left). This result was consistent with the shift in spectrotemporal filtering reflected in the frequency

content of responses for the ΔSPL condition (because the change in intensity is positive, the shift for the SPL condition is toward higher frequencies) (Figure 4.2.11, right).

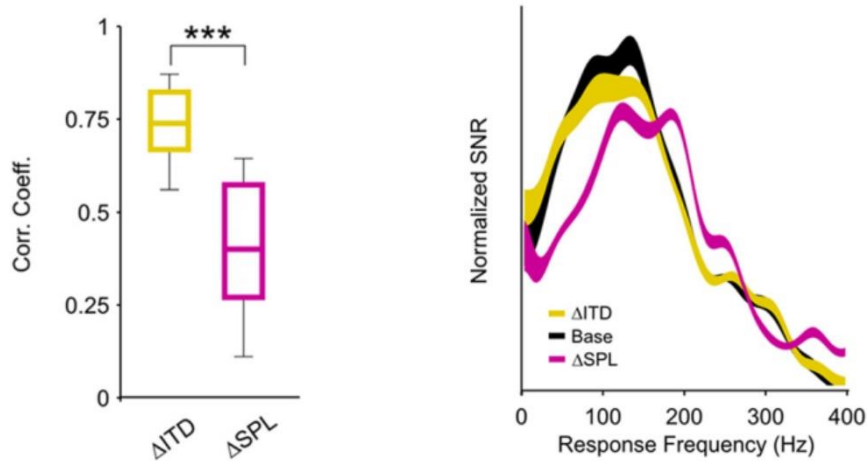*Figure 4.2.11. Left: Boxplots showing the distribution of CCs between the PSTHs for the base condition and each of the two change conditions for a sample of 13 neurons. Right: The SNR as a function of response frequency under the three stimulus conditions. ***p < 0.001. The changes in intensity along the negative slope of the RLF caused a decrease in SNR similar to that caused by a change in ITD.*

However, unlike changes in intensity along the positive slope of the RLF, which had no effect on SNR (Figure 4.2.4.), changes in intensity along the negative slope of the RLF caused a decrease in SNR similar to that caused by a change in ITD (paired Wilcoxon tests, $p = 0.006$ for ΔITD and $p = 0.05$ for ΔSPL) (Figure 4.2.12, left). As a result, the effects of changes in intensity along the negative slope of the RLF on coding were similar to those caused by a change in ITD. Both changes resulted in a significant decrease in decoder performance for single neurons relative to the base condition (paired Wilcoxon tests, $p < 0.001$ for ΔITD and $p = 0.01$ for ΔSPL) (Figure 4.2.12, right). Thus, the ability of single neurons to encode low-frequency sounds appears to be robust to changes in intensity along the positive slope of the RLF, but not to changes in intensity along the negative slope of the RLF.
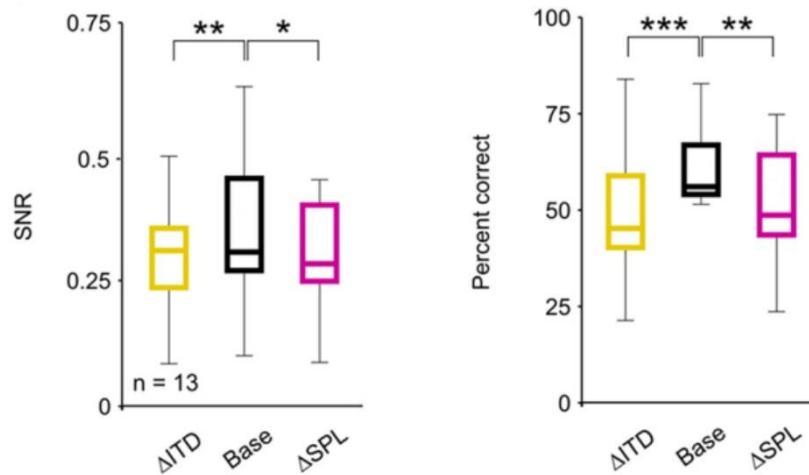
Figure 4.2.12. Left: Boxplots showing the distribution of response SNRs for each of the three stimulus conditions. Right: Boxplots showing the distribution of decoder performance at the optimal timescale for each of the three stimulus conditions. Chance level performance was 12.5%. In each plot, the central mark indicates the median, the edges of the box indicate the 25th and 75th percentiles, and the error bars extend to the most extreme values. The results of paired Wilcoxon tests comparing the medians of the distributions are indicated. *p < 0.05, **p < 0.01, ***p < 0.001. These results show that the ability of single neurons to encode low-frequency sounds appears to be robust to changes in intensity along the positive slope of the RLF, but not to changes in intensity along the negative slope of the RLF.

5. Discussion and conclusions

5.1. Decoding with distance metrics

We have demonstrated that when decoding population spike trains and calcium signals using distance metrics, optimization of the influence of each cell on the overall result can provide an increase in performance relative to simple weighting schemes. The results demonstrate that for populations of cells in the auditory and visual systems with a variety of signal and noise correlations, the benefit of genetic optimization can be relatively large (up to 10%). The results demonstrate that the optimal weights for population decoding cannot be derived simply from the performance of each cell as an individual, suggesting that there may be a relationship between the optimal weights for decoder performance and the correlations between cells in the population. One interesting avenue for further research would be to characterize this relationship, i.e., to explicitly describe the impact of correlations on the optimal weighting scheme when decoding population responses. This relationship also suggests the potential of optimization as a tool for measuring the contribution of individual cells to the population code. For example, as the stimulus and/or correlations in the population change, the corresponding changes in the optimal weights for different cells or groups of cells could be used to assay the change in the distribution of information across the population. A study by Ince and colleagues (Ince et al., 2013) addressed this question by analyzing neural population responses recorded from the auditory cortex of awake macaques. In that paper, the authors used the information carried by neuronal subpopulations for optimization and showed that a small fraction of temporally precise cells carries most of the information. This finding assumed that a few neurons carry large amounts of information. It was also found that neurons having high firing rates and short encoding times carried the high information. These results support the emerging picture that a small subset of all neurons recorded during an experiment contribute significantly to the information coding of a larger neural population (Panzeri et al., 2015). Studies on the auditory system summarized by (Panzeri et al., 2015) also show that the neurons contributing the most to the information content of the population are also the ones that fire sparsely (Centanni et al., 2014; Garcia-Lazaro et al., 2013; Ince et al., 2013). Conclusively, finding the optimal weights for neurons in a population is crucial for understanding the information coding capabilities of neural populations. Our

method provides a solution for this problem, however, with the advent of machine learning methods, faster and more optimal decoding of larger neural populations became possible (Glaser et al., 2020).

5.2. Revealing the effects of interaural time and intensity difference on the coding of low-frequency sounds

We have demonstrated that even when changes in intensity and ITD have similar effects on the spike rate of a single neuron, they can have different effects on the neuron's ability to encode low-frequency sounds. I found that a change in intensity along the positive slope of the RLF evoked a change in spectrotemporal filtering properties that changed the overall timing of spike events, but preserved the precision of spike timing across trials such that decoding of sound token identity from the responses of single neurons was not affected. In contrast, a change in ITD did not evoke a change in spectrotemporal filtering properties and, thus, had little impact on the overall timing of spike events, but had strong effects on the precision of spike timing across trials and, consequently, on decoding. However, because the two brain hemispheres are mirror images of each other in terms of ITD sensitivity, changes in ITD had no net effect on coding across the entire population. These effects were robust to the addition of background noise at both the single neuron and population level.

We also found that the effects of changes in intensity along the negative slope of the RLF were different from those of changes along the positive slope. Changes in intensity along the negative slope of the RLF caused changes in both the overall timing of spike events and the precision of spike timing across trials, and had effects on decoding that were similar to those caused by a change in ITD.

Our results show that, at least at moderate intensities, the auditory system can simultaneously employ fundamentally different strategies to maintain the flow of information in the face of changes in intensity and ITD. Because a change in intensity will have a similar effect on all neurons in the population, mechanisms that adjust the response properties of single neurons are necessary to preserve the flow of information at the population level. However, because a change in ITD will enhance coding in single

neurons in one hemisphere and degrade it in the other, its effects will balance out at the population level and no mechanisms that adjust the response properties of single neurons are necessary. The location and nature of the integration of the information about ITD from the two hemispheres remains a source of speculation (Porter & Groh, 2006). It seems clear that this integration does not take place in primary auditory cortex (Eggermont & Mossop, 1998; King & Campbell, 2005; Stecker et al., 2005), but there is evidence suggesting that it may take place in higher cortical areas (Miller & Recanzone, 2009; Stecker et al., 2003). However, the ITD coding strategy in the primary auditory cortex is already different compared to the IC. Belliveau and colleagues (Belliveau et al., 2014) showed that in the IC, most of the neurons responded maximally to contralateral ITDs close to the physiological range of the gerbil. This result corresponds well to my results presented in this thesis. However, Belliveau and colleagues have also found in the same study (Belliveau et al., 2014) that there is no such preference for ITD laterality, the preferred ITDs of A1 neurons were distributed evenly along the physiological range and there was an even number of neurons preferring contralateral and ipsilateral ITDs. These results suggest a fundamentally different coding strategy of the IC and A1 in gerbils. One possible explanation for this observation is that A1 may not be required for the localization of single sound sources in gerbils.

The intensity-dependent changes in spectrotemporal filtering properties observed here, as well as the associated changes in the overall timing of response events, are similar to those that have been observed throughout the auditory system (Lesica & Grothe, 2008a; Møller, 1977; Nagel & Doupe, 2006). These effects are due, at least in part, to the nonlinear properties of the basilar membrane, but central mechanisms such as inhibition within the IC may also play a role (Caspary et al., 2002). The origin of the ITD-dependent changes in the precision of spike timing across trials is less clear. Since a change in ITD does not actually affect the responses in the auditory nerves, but only the timing between them, the ITD-dependent changes that we observe must arise centrally after binaural convergence. One possible source of the observed effects is the coincidence detection mechanism in the MSO (which, presumably, provides the primary inputs to the IC cells studied here), the reliability of which has long been known to vary with overall spike rate (Goldberg & Brown, 1969).

The results of this study describe the effects of intensity and ITD on the coding of low-frequency sounds. It remains to be seen whether or not similar effects are evident for high-frequency sounds. The mechanisms that optimize response properties in single neurons in response to changes in intensity operate across a wide range of frequencies (Lesica & Grothe, 2008b; Nagel & Doupe, 2006; Rees & Møller, 1987), so the effects of changes in intensity on coding are likely to be similar for low- and high-frequency sounds. For high-frequency sounds, there are two important spatial cues: interaural level differences (ILDs) and spectral notches. Like ITDs, ILDs are computed centrally after binaural integration and have opposing effects on different subsets of the population, so no mechanism to compensate for changes in ILD at the single neuron level may be necessary (note the critical distinction between the changes in spectrotemporal filtering properties at issue here and mechanisms that adjust dynamic response range, which appear to operate in response to changes in both ITDs and ILDs (Dahmen et al., 2010; McAlpine et al., 2000; Spitzer & Semple, 1991)). Spectral notches, which are imposed by the pinnae, only affect a small subset of cells for a sound at any given location, so, again, no mechanism to adjust coding strategy may be necessary. However, because spectral notches are already present in the ear, it may be possible for the system to use some of the same machinery that compensates for changes in intensity with little additional overhead.

6. Summary

Most of the animals use three strategies for sound source localization (Palmer & Kuwada, 2005). Interaural time differences are used for low frequencies, while interaural level differences are used for higher frequencies where ITDs are not available. In addition, spectral content of the sound stimulus can also help in sound source localization. An interesting question about the sound source localization strategies of animals is how the coding of ITD and ILD changes in the brainstem for single neurons and for neuronal populations by changing time difference or level difference. This question can be formulated this way: do the changes of the two clues have the same effect on neuronal coding or do they change different aspects of the neuronal response? The importance of this question is that it can elucidate the strategies the central nervous system uses to keep sound information encoding robust in the face of changing sound features, for example when the position of the animal related to the sound source is changing.

We showed that changes in sound intensity evoked changes in spectrotemporal filtering that influenced the overall timing of spike events but preserved their precision across trials such that the decoding of single neuron responses was not affected. In contrast, changes in interaural time difference did not trigger changes in spectrotemporal filtering but did have strong effects on the precision of spike events and, consequently, on decoder performance. However, changes in ITD had opposing effects in the two brain hemispheres and, thus, canceled out at the population level. These results were similar with and without the addition of background noise (Horvath & Lesica, 2011).

For the study about the effects of changes in interaural time and level differences on single neuronal coding, decoding methods were used (Victor & Purpura, 1996). When more single neuronal responses are recorded from the same population, population response can be characterized by the population coding performance. The other published paper included in this thesis is about population decoder optimization. I contributed to the development of a genetic optimization method for decoding neuronal population responses. It was shown that genetic optimization is able to provide a superior distance metrics decoder performance to neuronal weight optimization using randomly assigned weights or weights obtained decoding each neuron individually. (Hofer et al., 2010)

7. Összefoglalás

Az állatvilágban háromféle stratégia figyelhető meg a hallott hang forrásának lokalizációjára. Az alacsony frekvenciás hangok esetében a fülközi időkülönbség (interaural time difference, ITD) hordozza a térbeli információt, míg a magasabb frekvenciák esetén, ahol az ITD már nem áll rendelkezésre, a fülközi hangerő különbséget (interaural level difference, ILD) felhasználva detektálható a hangforrás térbeli elhelyezkedése. Ezeken túl pedig a hanginger spektrális összetétele is hozzájárul a hangforrás lokalizációjához. Az ITD-t és az ILD-t vizsgálva felmerül az érdekes kérdés, hogy a két információforrás hatása azonos-e az idegi kódolásra, vagy pedig mindkettő az idegrendszeri válasz más aspektusait befolyásolja. A kérdés fontosságát az adja, hogy a válasz fény deríthet arra, milyen módon éri el az idegrendszer a hanginformáció robusztus kódolását, miközben a hanginger tulajdonságai folyamatosan változnak.

A disszertációban bemutatott tanulmányban kimutattuk, hogy a hangerő változása megváltoztatta a spektrotemporális szűrést a colliculus inferiorban. Ez befolyással volt az idegsejt tüzelések időzítésére, de az ismétlések között a tüzelések pontossága nem, ezen keresztül pedig az egyes idegsejtek válaszának dekódolása nem változott. Ezzel szemben, bár az ITD változások nem változtatták meg a spektrotemporális szűrést, az idegsejt tüzelések pontosságát jelentősen befolyásolták, ezen keresztül pedig a dekódolás pontosságára is hatással voltak. Az ITD változások a két agyféltekében ellentétes irányú változásokat indukáltak, így populációs szinten kioltották egymást. Az eredményeink ehhez hasonlóak voltak akkor is, ha a hangingerhez háttérzajt is kevertünk (Horvath & Lesica, 2011).

A fenti eredmények megállapításához dekódolási stratégiákat alkalmaztunk (Victor & Purpura, 1996). A disszertációban ismertetett másik publikáció egy sejtpopulációs aktivitás dekódolásának optimalizálásra használt genetikus algoritmus fejlesztését és használatát mutatja be. Megmutattuk, hogy a genetikai algoritmus alapú populációs dekódolás jobb dekóder teljesítményt eredményez a véletlenszerű súlyozáshoz vagy az egyes idegsejt egyenkénti válaszának dekódolásához képest (Hofer et al., 2010)

8. References

Aljadeff, J., Lansdell, B. J., Fairhall, A. L., & Kleinfeld, D. (2016). Analysis of Neuronal Spike

Trains, Deconstructed. *Neuron*, *91*(2), 221–259.

https://doi.org/10.1016/j.neuron.2016.05.039

Aronov, D., Reich, D. S., Mechler, F., & Victor, J. D. (2003). Neural Coding of Spatial Phase in

V1 of the Macaque Monkey. *Journal of Neurophysiology*, *89*(6), 3304–3327.

https://doi.org/10.1152/jn.00826.2002

Bal, R., Green, G. G. R., Rees, A., & Sanders, D. J. (2002). Firing patterns of inferior colliculus

neurons–histology and mechanism to change firing patterns in rat brain slices.

Neuroscience Letters, *317*(1), 42–46. https://doi.org/10.1016/S0304-3940(01)02425-9

Belliveau, L. A. C., Lyamzin, D. R., & Lesica, N. A. (2014). The Neural Representation of

Interaural Time Differences in Gerbils Is Transformed from Midbrain to Cortex.

Journal of Neuroscience, *34*(50), 16796–16808.

https://doi.org/10.1523/JNEUROSCI.2432-14.2014

Billimoria, C. P., Kraus, B. J., Narayan, R., Maddox, R. K., & Sen, K. (2008). Invariance and

Sensitivity to Intensity in Neural Discrimination of Natural Sounds. *Journal of

Neuroscience*, *28*(25), 6304–6308. https://doi.org/10.1523/JNEUROSCI.0961-08.2008

Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT

Press.

Boring Edwin G. (1942). *Sensation And Perception In The History Of Experimental

Psychology*. Appleton-Century Company.

http://archive.org/details/in.ernet.dli.2015.52372

Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Nature

Neuroscience*, *2*(11), 947–957. https://doi.org/10.1038/14731

Brown, E. N., Kass, R. E., & Mitra, P. P. (2004). Multiple neural spike train data analysis:

State-of-the-art and future challenges. *Nature Neuroscience*, *7*(5), 456–461.

https://doi.org/10.1038/nn1228

Brughera, A., Dunai, L., & Hartmann, W. M. (2013). Human interaural time difference thresholds for sine tones: The high-frequency limit. *The Journal of the Acoustical Society of America*, *133*(5), 2839–2855. https://doi.org/10.1121/1.4795778

Burger, R. M., & Pollak, G. D. (2001). Reversible Inactivation of the Dorsal Nucleus of the Lateral Lemniscus Reveals Its Role in the Processing of Multiple Sound Sources in the Inferior Colliculus of Bats. *The Journal of Neuroscience*, *21*(13), 4830–4843. https://doi.org/10.1523/JNEUROSCI.21-13-04830.2001

Cant, N. B. (2005). Projections from the Cochlear Nuclear Complex to the Inferior Colliculus. In J. A. Winer & C. E. Schreiner (Eds.), *The Inferior Colliculus* (pp. 115–131). Springer. https://doi.org/10.1007/0-387-27083-3_3

Cant, N. B., & Benson, C. G. (2006). Organization of the inferior colliculus of the gerbil (Meriones unguiculatus): Differences in distribution of projections from the cochlear nuclei and the superior olivary complex. *J Comp Neurol*, *495*(5), 511–528. https://doi.org/10.1002/cne.20888

Cariani, P. (2011). Jeffress model. *Scholarpedia*, *6*(7), 2920. https://doi.org/10.4249/scholarpedia.2920

Carr, C. E., & Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuroscience*, *10*(10), 3227–3246. https://doi.org/10.1523/JNEUROSCI.10-10-03227.1990

Caspary, D. M., Palombi, P. S., & Hughes, L. F. (2002). GABAergic inputs shape responses to amplitude modulated stimuli in the inferior colliculus. *Hearing Research*, *168*(1–2), 163–173. https://doi.org/10.1016/S0378-5955(02)00363-5

Centanni, T. M., Sloan, A. M., Reed, A. C., Engineer, C. T., Rennaker, R. L., & Kilgard, M. P. (2014). Detection and identification of speech sounds using cortical activity patterns. *Neuroscience*, *258*, 292–306. https://doi.org/10.1016/j.neuroscience.2013.11.030

Covey, E. (1993). The Monaural Nuclei of the Lateral Lemniscus: Parallel Pathways from Cochlear Nucleus to Midbrain. In M. A. Merchán, J. M. Juiz, D. A. Godfrey, & E.

Mugnaini (Eds.), *The Mammalian Cochlear Nuclei: Organization and Function* (pp. 321–334). Springer US. https://doi.org/10.1007/978-1-4615-2932-3_25

Dahmen, J. C., Keating, P., Nodal, F. R., Schulz, A. L., & King, A. J. (2010). Adaptation to Stimulus Statistics in the Perception and Neural Representation of Auditory Space. *Neuron*, *66*(6), 937–948. https://doi.org/10.1016/j.neuron.2010.05.018

Dean, I., Harper, N. S., & McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience*, *8*(12), Article 12. https://doi.org/10.1038/nn1541

Eggermont, J. J., & Mossop, J. E. (1998). Azimuth Coding in Primary Auditory Cortex of the Cat. I. Spike Synchrony Versus Spike Count Representations. *Journal of Neurophysiology*, *80*(4), 2133–2150. https://doi.org/10.1152/jn.1998.80.4.2133

Egorova, M., Ehret, G., Vartanian, I., & Esser, K.-H. (2001). Frequency response areas of neurons in the mouse inferior colliculus. I. Threshold and tuning characteristics. *Experimental Brain Research*, *140*(2), 145–161. https://doi.org/10.1007/s002210100786

Ehret, G., & Schreiner, C. E. (2005). Spectral and Intensity Coding in the Auditory Midbrain. In J. A. Winer & C. E. Schreiner (Eds.), *The Inferior Colliculus* (pp. 312–345). Springer New York. https://doi.org/10.1007/0-387-27083-3_11

Faingold, C. L., Anderson, C. A. B., & Randall, M. E. (1993). Stimulation or blockade of the dorsal nucleus of the lateral lemniscus alters binaural and tonic inhibition in contralateral inferior colliculus neurons. *Hearing Research*, *69*(1), 98–106. https://doi.org/10.1016/0378-5955(93)90097-K

Fraley, C., & Raftery, A. E. (1999). MCLUST: Software for Model-Based Cluster Analysis. *Journal of Classification*, *16*(2), 297–306. https://doi.org/10.1007/s003579900058

Fraley, C., & Raftery, A. E. (2002). Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association*, *97*(458), 611–631. https://doi.org/10.1198/016214502760047131

Garcia-Lazaro, J. A., Belliveau, L. A. C., & Lesica, N. A. (2013). Independent Population Coding of Speech with Sub-Millisecond Precision. *Journal of Neuroscience*, *33*(49), 19362–19372. https://doi.org/10.1523/JNEUROSCI.3711-13.2013

Glaser, J. I., Benjamin, A. S., Chowdhury, R. H., Perich, M. G., Miller, L. E., & Kording, K. P. (2020). Machine Learning for Neural Decoding. *eNeuro*, *7*(4). https://doi.org/10.1523/ENEURO.0506-19.2020

Goldberg, J. M., & Brown, P. B. (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization. *Journal of Neurophysiology*, *32*(4), 613–636. https://doi.org/10.1152/jn.1969.32.4.613

Groh, J. M., Kelly, K. A., & Underhill, A. M. (2003). A Monotonic Code for Sound Azimuth in Primate Inferior Colliculus. *Journal of Cognitive Neuroscience*, *15*(8), 1217–1231. https://doi.org/10.1162/089892903322598166

Hancock, K. E., & Delgutte, B. (2004). A Physiologically Based Model of Interaural Time Difference Discrimination. *Journal of Neuroscience*, *24*(32), 7110–7117. https://doi.org/10.1523/JNEUROSCI.0762-04.2004

Hartmann, W. M., & Macaulay, E. J. (2014). Anatomical limits on interaural time differences: An ecological perspective. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00034

Heffner, R. S., Heffner, H. E., Kearns, D., Vogel, J., & Koay, G. (1994). Sound localization in chinchillas. I: Left/right discriminations. *Hearing Research*, *80*(2), 247–257. https://doi.org/10.1016/0378-5955(94)90116-3

Hofer, S. B., Mrsic-Flogel, T. D., Horvath, D., Grothe, B., & Lesica, N. A. (2010). Optimization of population decoding with distance metrics. *Neural Networks*, *23*(6), 728–732. https://doi.org/10.1016/j.neunet.2010.04.007

Horvath, D., & Lesica, N. A. (2011). The Effects of Interaural Time Difference and Intensity on the Coding of Low-Frequency Sounds in the Mammalian Midbrain. *Journal of Neuroscience*, *31*(10), 3821–3827. https://doi.org/10.1523/JNEUROSCI.4806-10.2011

Houghton, C., & Sen, K. (2008). A new multi-neuron spike-train metric. *Neural Computation*, *20*(6), 1495–1511. https://doi.org/10.1162/neco.2007.10-06-350

Ince, R. A. A., Panzeri, S., & Kayser, C. (2013). Neural Codes Formed by Small and Temporally Precise Populations in Auditory Cortex. *Journal of Neuroscience*, *33*(46), 18277–18287. https://doi.org/10.1523/JNEUROSCI.2631-13.2013

Jeffress, L. (1948). A place theory of sound localization. *Journal of Comparative and Physiological Psychology*, *41*, 35–39.

Kennedy, J. F., Eberhart, R. C., & Shi, Y. (2001). *Swarm intelligence*. Morgan Kaufmann Publishers.

Keplinger, S., Beiderbeck, B., Michalakis, S., Biel, M., Grothe, B., & Kunz, L. (2018). Optogenetic Control of Neural Circuits in the Mongolian Gerbil. *Frontiers in Cellular Neuroscience*, *12*. https://doi.org/10.3389/fncel.2018.00111

King, A. J., & Campbell, R. A. A. (2005). Cortical Processing of Sound-Source Location. *ACTA ACUSTICA UNITED WITH ACUSTICA*, *91*, 11.

Kvale, M. N., & Schreiner, C. E. (2004). Short-Term Adaptation of Auditory Receptive Fields to Dynamic Stimuli. *Journal of Neurophysiology*, *91*(2), 604–612. https://doi.org/10.1152/jn.00484.2003

Langner, G., & Schreiner, C. E. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *Journal of Neurophysiology*, *60*(6), 1799–1822. https://doi.org/10.1152/jn.1988.60.6.1799

Le Beau, F. E., Rees, A., & Malmierca, M. S. (1996). Contribution of GABA- and glycine-mediated inhibition to the monaural temporal response properties of neurons in the inferior colliculus. *Journal of Neurophysiology*, *75*(2), 902–919. https://doi.org/10.1152/jn.1996.75.2.902

Leakey, D. M., Sayers, B. McA., & Cherry, C. (1958). Binaural Fusion of Low- and High-Frequency Sounds. *The Journal of the Acoustical Society of America*, *30*(3), 222–222. https://doi.org/10.1121/1.1909549

Lesica, N. A., & Grothe, B. (2008a). Dynamic Spectrotemporal Feature Selectivity in the Auditory Midbrain. *Journal of Neuroscience*, *28*(21), 5412–5421. https://doi.org/10.1523/JNEUROSCI.0073-08.2008

Lesica, N. A., & Grothe, B. (2008b). Efficient Temporal Processing of Naturalistic Sounds. *PLoS ONE*, *3*(2), e1655. https://doi.org/10.1371/journal.pone.0001655

Lesica, N. A., Lingner, A., & Grothe, B. (2010). Population coding of interaural time differences in gerbils and barn owls. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(35), 11696–11702. https://doi.org/10.1523/JNEUROSCI.0846-10.2010

Liberman, M. C., & Brown, M. C. (1986). Physiology and anatomy of single olivocochlear neurons in the cat. *Hear Res*, *24*(1), 17–36.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *The Journal of the Acoustical Society of America*, *106*(4 Pt 1), 1633–1654.

Lyzwa, D., & Wörgötter, F. (2016). Neural and Response Correlations to Complex Natural Sounds in the Auditory Midbrain. *Frontiers in Neural Circuits*, *10*. https://doi.org/10.3389/fncir.2016.00089

Maki, K., & Furukawa, S. (2005). Acoustical cues for sound localization by the Mongolian gerbil, Meriones unguiculatus. *The Journal of the Acoustical Society of America*, *118*(2), 872–886. https://doi.org/10.1121/1.1944647

McAlpine, D., & Grothe, B. (2003). Sound localization and delay lines – do mammals fit the model? *Trends in Neurosciences*, *26*(7), 347–350. https://doi.org/10.1016/S0166-2236(03)00140-1

McAlpine, D., Jiang, D., & Palmer, A. R. (2001). A neural code for low-frequency sound localization in mammals. *Nature Neuroscience*, *4*(4), 396–401. https://doi.org/10.1038/86049

McAlpine, D., Jiang, D., Shackleton, T. M., & Palmer, A. R. (2000). Responses of Neurons in the Inferior Colliculus to Dynamic Interaural Phase Cues: Evidence for a Mechanism of Binaural Adaptation. *Journal of Neurophysiology*, *83*(3), 1356–1365. https://doi.org/10.1152/jn.2000.83.3.1356

Meffin, H., & Grothe, B. (2009). Selective filtering to spurious localization cues in the mammalian auditory brainstem. *The Journal of the Acoustical Society of America*, *126*(5), 2437–2454. https://doi.org/10.1121/1.3238239

Merchán, M. A., Malmierca, M. S., Bajo, V. M., & Bjaalie, J. G. (1997). The Nuclei of the Lateral Lemniscus. In J. Syka (Ed.), *Acoustical Signal Processing in the Central Auditory System* (pp. 211–226). Springer US. https://doi.org/10.1007/978-1-4419-8712-9_20

Miller, L. M., & Recanzone, G. H. (2009). Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. *Proceedings of the National Academy of Sciences*, *106*(14), 5931–5935. https://doi.org/10.1073/pnas.0901023106

Møller, A. R. (1977). Frequency selectivity of single auditory-nerve fibers in response to broadband noise stimuli. *The Journal of the Acoustical Society of America*, *62*(1), 135–142. https://doi.org/10.1121/1.381495

Morest, D. K., & Oliver, D. L. (1984). The neuronal architecture of the inferior colliculus in the cat: Defining the functional anatomy of the auditory midbrain. *Journal of Comparative Neurology*, *222*(2), 209–236. https://doi.org/10.1002/cne.902220206

Mrsic-Flogel, T. D., Hofer, S. B., Ohki, K., Reid, R. C., Bonhoeffer, T., & Hübener, M. (2007). Homeostatic Regulation of Eye-Specific Responses in Visual Cortex during Ocular Dominance Plasticity. *Neuron*, *54*(6), 961–972. https://doi.org/10.1016/j.neuron.2007.05.028

Nagel, K. I., & Doupe, A. J. (2006). Temporal Processing and Adaptation in the Songbird
Auditory Forebrain. *Neuron*, *51*(6), 845–859.
https://doi.org/10.1016/j.neuron.2006.08.030

Oliver, D. L. (1984). Neuron types in the central nucleus of the inferior colliculus that project to
the medial geniculate body. *Neuroscience*, *11*(2), 409–424.
https://doi.org/10.1016/0306-4522(84)90033-2

Oliver, D. L. (2005). Neuronal Organization in the Inferior Colliculus. In J. A. Winer & C. E.
Schreiner (Eds.), *The Inferior Colliculus* (pp. 69–114). Springer New York.
https://doi.org/10.1007/0-387-27083-3_2

Osen, K. K. (1969). Cytoarchitecture of the cochlear nuclei in the cat. *Journal of Comparative
Neurology*, *136*(4), 453–483. https://doi.org/doi:10.1002/cne.901360407

Palmer, A. R., & Kuwada, S. (2005). Binaural and Spatial Coding in the Inferior Colliculus. In
J. A. Winer & C. E. Schreiner (Eds.), *The Inferior Colliculus* (pp. 377–410). Springer
New York. https://doi.org/10.1007/0-387-27083-3_13

Panzeri, S., Macke, J. H., Gross, J., & Kayser, C. (2015). Neural population coding: Combining
insights from microscopic and mass signals. *Trends in Cognitive Sciences*, *19*(3), 162–
172. https://doi.org/10.1016/j.tics.2015.01.002

Pecka, M., Zahn, T. P., Saunier-Rebori, B., Siveke, I., Felmy, F., Wiegrebe, L., Klug, A.,
Pollak, G. D., & Grothe, B. (2007). Inhibiting the inhibition: A neuronal network for
sound localization in reverberant environments. *The Journal of Neuroscience: The
Official Journal of the Society for Neuroscience*, *27*(7), 1782–1790.
https://doi.org/10.1523/JNEUROSCI.5335-06.2007

Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967). Neuronal spike trains and stochastic point
processes. I. The single spike train. *Biophysical Journal*, *7*(4), 391–418.
https://doi.org/10.1016/S0006-3495(67)86596-2

Peruzzi, D., Sivaramakrishnan, S., & Oliver, D. L. (2000). Identification of cell types in brain slices of the inferior colliculus. *Neuroscience*, *101*(2), 403–416. https://doi.org/10.1016/S0306-4522(00)00382-1

Pickles, J. O. (2012). *An introduction to the physiology of hearing* (Vol. 4). Emerald Group Publishing Limited.

Pickles, J. O. (2015). Auditory pathways: Anatomy and physiology. In *Handb Clin Neurol* (Vol. 129, pp. 3–25).

Porter, K. K., & Groh, J. M. (2006). Chapter 18 The "other" transformation required for visual–auditory integration: Representational format. In *Progress in Brain Research* (Vol. 155, pp. 313–323). Elsevier. https://doi.org/10.1016/S0079-6123(06)55018-6

Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nature Reviews Neuroscience*, *1*(2), 125–132. https://doi.org/10.1038/35039062

Quian Quiroga, R., & Panzeri, S. (2009). Extracting information from neuronal populations: Information theory and decoding approaches. *Nat Rev Neurosci*, *10*(3), 173–185. https://doi.org/10.1038/nrn2578

Rees, A., & Langner, G. (2005). Temporal Coding in the Auditory Midbrain. In J. A. Winer & C. E. Schreiner (Eds.), *The Inferior Colliculus* (pp. 346–376). Springer-Verlag. https://doi.org/10.1007/0-387-27083-3_12

Rees, A., & Møller, A. R. (1987). Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hearing Research*, *27*(2), 129–143. https://doi.org/10.1016/0378-5955(87)90014-1

Rees, A., & Palmer, A. R. (1988). Rate-intensity functions and their modification by broadband noise for neurons in the guinea pig inferior colliculus. *The Journal of the Acoustical Society of America*, *83*(4), 1488–1498. https://doi.org/10.1121/1.395904

Rolewicz, S. (1987). *Functional Analysis and Control Theory*. Springer Netherlands. https://doi.org/10.1007/978-94-015-7758-8

Rose, J. E., Gross, N. B., Geisler, C. D., & Hind, J. E. (1966). Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source. *Journal of Neurophysiology*, *29*(2), 288–314. https://doi.org/10.1152/jn.1966.29.2.288

Sadagopan, S., & Wang, X. (2008). Level Invariant Representation of Sounds by Populations of Neurons in Primary Auditory Cortex. *Journal of Neuroscience*, *28*(13), 3415–3426. https://doi.org/10.1523/JNEUROSCI.2743-07.2008

Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K., & Redish, A. D. (2005). Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience*, *131*(1), 1–11. https://doi.org/10.1016/j.neuroscience.2004.09.066

Schofield, B. R. (2005). Superior Olivary Complex and Lateral Lemniscal Connections of the Auditory Midbrain. In J. A. Winer & C. E. Schreiner (Eds.), *The Inferior Colliculus* (pp. 132–154). Springer New York. https://doi.org/10.1007/0-387-27083-3_4

Schreiner, C. E., & Langner, G. (1997). Laminar fine structure of frequency organization in auditory midbrain. *Nature*, *388*(6640), 383–386. https://doi.org/10.1038/41106

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*(6), 3394–3411. https://doi.org/10.1121/1.1624067

Sivaramakrishnan, S., & Oliver, D. L. (2001). Distinct K Currents Result in Physiologically Distinct Cell Types in the Inferior Colliculus of the Rat. *The Journal of Neuroscience*, *21*(8), 2861–2877. https://doi.org/10.1523/JNEUROSCI.21-08-02861.2001

Spitzer, M., & Semple, M. (1991). Interaural phase coding in auditory midbrain: Influence of dynamic stimulus features. *Science*, *254*(5032), 721–724. https://doi.org/10.1126/science.1948053

Stecker, G. C., Harrington, I. A., & Middlebrooks, J. C. (2005). Location Coding by Opponent Neural Populations in the Auditory Cortex. *PLOS Biology*, *3*(3), e78. https://doi.org/10.1371/journal.pbio.0030078

Stecker, G. C., Mickey, B. J., Macpherson, E. A., & Middlebrooks, J. C. (2003). Spatial

    Sensitivity in Field PAF of Cat Auditory Cortex. *Journal of Neurophysiology*, *89*(6),

    2889–2903. https://doi.org/10.1152/jn.00980.2002

Syka, J., Popelář, J., Kvašňák, E., & Astl, J. (2000). Response properties of neurons in the

    central nucleus and external and dorsal cortices of the inferior colliculus in guinea pig.

    Experimental Brain Research, *133*(2), 254–266.

    https://doi.org/10.1007/s002210000426

van Rossum, M. C. W. (2001). A Novel Spike Distance. *Neural Computation*, *13*(4), 751–763.

    https://doi.org/10.1162/089976601300014321

Victor, J. D. (2005). Spike train metrics. *Current Opinion in Neurobiology*, *15*(5), 585–592.

    https://doi.org/10.1016/j.conb.2005.08.002

Victor, J. D., & Purpura, K. P. (1996). Nature and precision of temporal coding in visual cortex:

    A metric-space analysis. *Journal of Neurophysiology*, *76*(2), 1310–1326.

    https://doi.org/10.1152/jn.1996.76.2.1310

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound

    localization. *Journal of Experimental Psychology*, *27*(4), 339–368.

    https://doi.org/10.1037/h0054629

Warr, W. B. (1992). Organization of Olivocochlear Efferent Systems in Mammals. In D. B.

    Webster, A. N. Popper, & R. R. Fay (Eds.), *The Mammalian Auditory Pathway:*

    Neuroanatomy (pp. 410–448). Springer New York.

Wen, B., Wang, G. I., Dean, I., & Delgutte, B. (2009). Dynamic Range Adaptation to Sound

    Level Statistics in the Auditory Nerve. *Journal of Neuroscience*, *29*(44), 13797–13808.

    https://doi.org/10.1523/JNEUROSCI.5610-08.2009

Winer, J. A., & Schreiner, C. E. (2005). The Central Auditory System: A Functional Analysis.

    In J. A. Winer & C. E. Schreiner (Eds.), *The Inferior Colliculus* (pp. 1–68). Springer

    New York. https://doi.org/10.1007/0-387-27083-3_1

Yang, L., & Pollak, G. D. (1994). The roles of GABAergic and glycinergic inhibition on

binaural processing in the dorsal nucleus of the lateral lemniscus of the mustache bat.

Journal of Neurophysiology, *71*(6), 1999–2013.

https://doi.org/10.1152/jn.1994.71.6.1999

Yin, T. C., & Kuwada, S. (1983). Binaural interaction in low-frequency neurons in inferior

colliculus of the cat. III. Effects of changing frequency. *Journal of Neurophysiology*,

50(4), 1020–1042. https://doi.org/10.1152/jn.1983.50.4.1020

Zurek, P. M. (1987). The Precedence Effect. In W. A. Yost & G. Gourevitch (Eds.), *Directional

Hearing* (pp. 85–105). Springer US. https://doi.org/10.1007/978-1-4612-4738-8_4

Zwislocki, J., & Feldman, R. S. (1956). Just Noticeable Differences in Dichotic Phase. *The

Journal of the Acoustical Society of America*, *28*(5), 860–864.

https://doi.org/10.1121/1.1908495

9. List of publications

Publications related to the thesis:

Hofer, S. B., Mrsic-Flogel, T. D., Horvath, D., Grothe, B., & Lesica, N. A. (2010). Optimization of population decoding with distance metrics. Neural Networks, 23(6), 728–732. https://doi.org/10.1016/j.neunet.2010.04.007

Horvath, D., & Lesica, N. A. (2011). The Effects of Interaural Time Difference and Intensity on the Coding of Low-Frequency Sounds in the Mammalian Midbrain. Journal of Neuroscience, 31(10), 3821–3827. https://doi.org/10.1523/JNEUROSCI.4806-10.2011

Publications not related to the thesis:

Dombovari B, Fiath R, Kerekes B, Toth E, Wittner L, Horvath D, Seidl K, Herwik S, Torfs T, Paul O, Ruther P, Neves H, Ulbert I. In vivo validation of the electronic depth control probes. Biomedizinische Technik 59 : 4 pp. 283-289. , 7 p. (2014)

Fiath R, Kerekes BP, Wittner L, Toth K, Beregszaszi P, Horvath D, Ulbert I. Laminar analysis of the slow wave activity in the somatosensory cortex of anesthetized rats. European Journal of Neuroscience 44 : 3 pp. 1935-1951. , 17 p. (2016)

Fiath R, Beregszaszi P, Horvath D, Wittner L, Aarts AA, Ruther P, Neves HP, Bokor H, Acsady L, Ulbert I. Large-scale recording of thalamocortical circuits: in vivo electrophysiology with the two-dimensional electronic depth control silicon probe. Journal of Neurophysiology 116 : 5 pp. 2312-2330. , 19 p. (2016)

Fiáth R, Raducanu BC, Musa S, Andrei A, Lopez CM, van Hoof C, Ruther P, Aarts A, Horváth D, Ulbert I. A silicon-based neural probe with densely-packed low-impedance titanium nitride microelectrodes for ultrahigh-resolution in vivo recordings. Biosensors & Bioelectronics 106 pp. 86-92. , 7 p. (2018)

Fiáth R, Hofer KT, Csikós V, Horváth D, Nánási T, Tóth K, Pothof F, Böhler C, Asplund M, Ruther P, Ulbert I. Long-term recording performance and biocompatibility of chronically implanted cylindrically-shaped, polymer-based neural interfaces. Biomedizinische Technik 63 : 3 pp. 301-315. , 15 p. (2018)

10. Acknowledgements

I would like to thank my supervisor, István Ulbert for his support during the several years I have been lucky to spend in his lab. I would also like thank the late György Karmos for always trusting and encouraging me. I am grateful to Nicholas Lesica for letting me perform the experiments described in my thesis and for all his advice, help and kindness. I would also like to thank Daniel Hillier for pushing me towards finishing this thesis and for his critical insights to my work. And, finally, I am deeply grateful to my wife, Zsófia Boronkai-Horváth, and my family for their patience and support during the years that led to the completion of this thesis.