Physiological Studies of Binaural Tone-in-Noise Detection in the Inferior Colliculus: the Role of Stimulus Envelope and Neural Fluctuations

by

Langchen Fan

Submitted in Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

Supervised by Professor Laurel H. Carney

Department of Biomedical Engineering

Arts, Sciences and Engineering

Edmund A. Hajim School of Engineering and Applied Sciences

University of Rochester

Rochester, NY

2020

Table of Contents

Biographical Sketch

Langchen Fan graduated from Peking University (China) with the degree Bachelor of Science in Psychology in 2014. Her thesis for the Bachelor degree was focused on psychophysical hearing studies with human listeners. She began her doctoral studies in the Department of Biomedical Engineering at University at Rochester in 2014. Her thesis work was focused on neural mechanisms of detecting a tone in noise, supervised by Dr. Laurel H. Carney. She received the degree Master of Science in Biomedical Engineering in 2016. She visited the Hearing System Group at Technology University of Denmark for a collaborative project from June to August in 2018.

The following publications were a result of work conducted during doctoral study:

Journal Publications:

Fan, L., Henry, K.S., Carney, L.H. (2018) “Challenging one model with many stimuli: simulating responses in the inferior colliculus,” *Acta Acustica united with Acustica*, *104*(5), 895-899.

Conference Publications:

Fan, L., Henry, K. S., & Carney, L. H. (2020). Responses to tones masked by gaussian or low-noise noise in the inferior colliculus of awake rabbits. Abstract, *Association for Reseearch in Otolaryngology.*

Fan, L., Henry, K. S., & Carney, L. H. (2019). Physiological studies of binaural tone-in-noise detection in the inferior colliculus of awake rabbit. *The Journal of the Acoustical Society of America*, *145*(3), 1719-1720.

Fan, L., Mesiano, P.A., Dau, T., Carney, L.H. (2019) “Physiological model simulations of comodulation masking release for listeners with normal hearing and hearing Loss,” Abstract, *Association for Research in Otolaryngology*.

Mesiano, P.A., Zaar, J., Kowalewski, B., **Fan, L.,** Carney, L.H., Dau, T. (2019) “Characterizing the role of hearing loss in comodulation masking release: behavioral measurements and computational model predictions”, Abstract, *Association for Research in Otolaryngology*.

Fan, L. & Carney, LH. (2017) “Neural responses in the inferior colliculus to diotic tone-in-noise stimuli support detection based on envelope and neural fluctuations,” Abstract, *Association for Research in Otolaryngology*.

Carney, L.H., **Fan, L.,** Galant, N.K., Li, Z., Maxwell, B., Teverovsky, D.E., Varner, T. (2016) “Visualizing population model responses of peripheral, brainstem and midbrain neurons to complex sounds,” Abstract, *Association for Research in Otolaryngology*.

Acknowledgements

I would like to thank my advisor, Dr. Laurel Carney, for the guidance and support that she provided me throughout my entire graduate study. Laurel is very knowledgeable in research and always gives me good suggestions when I have questions. I am especially grateful for the time that she put in to help me improve writing. I am also thankful for the opportunity that she provided me to visit the Dau Lab in Denmark. Her passion for work will always inspire me.

Thanks to Dr. Ken Henry, not only as a committee member, but also as a collaborator. Ken has provided me many insightful suggestions from which I benefited. Thanks to other committee members, Dr. Ralf Haefner, Dr. Marc Schieber, and Dr. Steve McAleavey, for providing me helpful feedback.

Thanks to our lab technician, Kris Abrams, for teaching me how to handle the rabbits and helping with the surgeries. Thanks to my previous and current lab members, Margaret Youngman and Meron Abate, for their help with the experiments. Thanks to our lab programmer, Doug Schwarz, for his help with Matlab.

Thanks to my fellow graduate students, Yingxuan Wang, Afagh Farhadi, Braden Maxwell, Paul Mitchell, and John Wilson, for helpful feedback on my analysis and writing. Thanks to my writing group members, Zheng Liu, Raul Rodriguez, Maddy Cappelloni, Paul Mitchell, and Shyanthony Synigal, for valuable feedback on my writing and the fun conversations.

Thanks to my parents, Dong Zhao and Guangliang Fan, for their unconditional and endless love, for their support, and for being there whenever I feel lost. Thanks to all my friends for the fun time together.

Abstract

Human detection thresholds in the tone-in-noise (TIN) paradigm are often explained by the power spectrum model, which evaluates the stimulus energy at the output of an auditory filter. However, human thresholds are minimally affected when the stimulus energy is made less reliable. Stimulus envelope has been proposed as a cue for the condition. The envelope of a narrowband maskers is encoded in the slow fluctuations of responses of auditory-nerve (AN) fibers, which is affected by cochlear compression and inner-hair-cell (IHC) saturation. Because, neurons in the inferior colliculus (IC) are sensitive to amplitude modulation (AM) depth and frequency, IC responses were hypothesized to explain detection of masked tones based on sensitivity to fluctuation, for both in-phase (diotic) and out-of-phase tones (dichotic) with identical noise maskers. Recordings were collected in the IC of awake rabbits. For the diotic condition, narrowband gaussian noise (GN) and low-noise noise (LNN) were used as fluctuation amplitudes in the stimulus envelope and in model AN responses decrease for GN maskers but increase for LNN upon addition of tones near threshold. IC neurons excited by AM were expected to have response rates positively correlated with fluctuation amplitudes, whereas neurons suppressed by AM were expected to have rates negatively correlated. As predicted, changes in rate upon addition of a tone were significantly correlated with sensitivity to AM. In the dichotic condition, previous modeling studies have proposed interaural-correlation (IAC) as a cue for detection, and sensitivity to interaural time difference (ITD) has been proposed to be the physiological basis. Here, IAC-, ITD-, and envelope-based hypothesis were all tested. Results showed that all three cues explained a proportion of dichotic TIN responses, and IAC-based cues topped. However, these cues may not be independent. Additionally, the rate-based thresholds of the most sensitive neurons for both GN and LNN

masker were similar to human detection threshold. However, the rate-based thresholds of the most sensitive neurons did not vary across frequencies for the dichotic condition, whereas human detection thresholds substantially decrease with increasing tone frequency. Therefore, threshold difference between the diotic and dichotic conditions were different between neural data and human psychophysical data.

Contributors and Funding Sources

This work was supervised by a dissertation committee consisting of Professor Laurel Carney (advisor) of Departments of Biomedical Engineering and Neuroscience, Professor Marc Schieber of Departments of Neurology, Center for Visual Science, Neuroscience, and Biomedical Engineering, Professor Stephen McAleavey of Departments of Biomedical Engineering and Electrical and Computer Engineering, Professor Kenneth Henry of Department of Otolaryngology, Neuroscience, and Biomedical Engineering, and Professor Ralf Haefner of the Department of Brain and Cognitive Sciences at the University of Rochester. All work conducted for the dissertation was completed by the student independently, under the supervision of Prof. Laurel H. Carney and supported by the National Institute of Health and National Institute on Deafness and Other Communication Disorders under grant number R01-010813.

List of Tables

List of Figures

Chapter 1. General Introduction

People with hearing loss have difficulty understanding conversations in noisy environments, and everyone experiences some degree of hearing loss with age. Yet the detection strategy used by listeners with normal hearing is still unclear. In the tone-in-noise (TIN) detection task, as the name indicates, listeners are presented with either a noise-alone waveform or a tone-plus-noise waveform, and they respond to indicate whether or not they hear the tone. With the simplest possible auditory target, a tone, studying TIN detection may improve our understanding of neural mechanisms of detection in noise, as well as general sound-encoding strategies of the auditory system.

In studies of TIN detection, researchers have varied parameters such as the tone frequency and noise bandwidth. Another variation is in the presentation of binaural stimuli; for example, comparisons are often made between detection thresholds for identical noise and tones presented to the two ears (N_0S_0, diotic) and identical noise but antiphasic tones (N_0S_π, dichotic). Behavioral monaural and binaural TIN detection have been studied for several decades (e.g., Fletcher, 1940; Hirsch, 1948), but the underlying neural mechanisms are still under debate. The goal of this thesis is to improve our understanding of the monaural and binaural neural cues used in TIN detection by recording single-neuron responses in the inferior colliculus of awake rabbits.

1.1 Diotic TIN detection

In the 1930s, knowledge of hearing mechanisms was focused on the peripheral auditory system: sounds are decomposed into frequencies along the basilar membrane of the cochlea (Helmholtz & Ellis, 1875; Fletcher, 1940), analogous to a bank of band-pass filters. According to the literature, to detect a tone in masking noise, the tone-to-noise

intensity ratio at the output of the filter tuned to the target frequency must exceed a criterion. Thus, higher tone levels are needed for detection as noise bandwidth increases (for a noise with fixed spectrum level), i.e. detection threshold increases with increasing bandwidth. According to this assumption, the threshold no longer increases once the masker bandwidth exceeds a certain bandwidth (i.e. the critical band), because the additional components outside the bandwidth of the filter tuned at the target frequency cannot mask the target tone. This concept of critical bands and the bank of auditory filters is the basis of the prevalent energy model, or power-spectrum model (Greenwood, 1961; Moore, 2007; Patterson, 1986). Later, researchers spent decades estimating the critical bandwidth using tone-detection thresholds, with or without modification of the noise masker (Greenwood, 1961; Irino & Patterson, 1997; Moore & Glasberg, 1983; Moore, Peters, & Glasberg, 1990; Patterson, 1976). The energy model has not only been used to explain human TIN detection thresholds, but also many other tasks, such as frequency discrimination (Wier, Jesteadt, & Green, 1977) and gap detection (Shailer & Moore, 1983). Additionally, the energy model has been used as the first level of processing in models for more complex auditory processing, such as pitch perception (Meddis & O'Mard, 1997), interaural correlation (Bernstein & Trahiotis, 1997), and speech segregation (Roman, Wang, & Brown, 2003).

However, the energy model fails in several conditions. First, in a roving-level paradigm, energy cues are made unreliable by randomly varying the stimulus level across stimulus intervals. Detection thresholds in the roving-level paradigm are much less affected than the energy model predicts (Kidd, Mason, Brantley, & Owen, 1989). Second, in an equal-energy paradigm, human thresholds cannot be explained by the energy cue

alone, whether the energy of the stimulus waveforms is equalized before or after adding a tone (Richards, 1992; Richards & Nekrich, 1993).

Physiological studies also do not support the power-spectrum model. Although incoming sounds are decomposed into frequencies on the basilar membrane, the width of the excitatory frequency range becomes larger with increasing sound level (Ruggero, 1992). At 60-70 dB SPL, at which listeners are usually tested in psychophysical studies, a much wider frequency range than the “critical band” (i.e. a few kilohertz) is excited on the basilar membrane (Ruggero, 1992); similarly, auditory-nerve (AN) fibers respond to a wider frequency range than critical bandwidth at 70 dB SPL (Geisler, Rhode, & Kennedy, 1974; Guinan & Gifford, 1988). Energy cues have been proposed to be encoded by the average rates of low-spontaneous-rate (LSR) AN fibers, i.e. increasing rates indicating higher sound levels (Costalupes, 1985; Delgutte, 1990). However, high-spontaneous-rate AN fibers are the major afferent input to the central system (Liberman, 1978), and usually saturate at approximately 30 dB SPL (Liberman, 1978; Schalk & Sachs, 1980; Smith, 1979). Neurons in the cochlear nucleus, which receives input from AN fibers and projects to higher levels, either have limited dynamic ranges (about 30 – 35 dB SPL, May & Sachs, 1992), or do not project to ascending pathways (review: Carney, 2018).

What could be the cue that listeners use for TIN detection if not energy? One possible answer is the envelope. In contrast to the temporal fine structure, the envelope is the relatively slow fluctuating outline of a sound waveform (Fig. 1.1 B and C). Narrowband gaussian noise has a fluctuating envelope, and adding a tone tends to flatten the envelope. Envelope cues have been shown to be accessible to human listeners (Hall, Haggard, & Fernandes, 1984; Kohlrausch et al., 1997; Richards, 2002; Richards & Nekrich, 1993). Previous studies have shown that an envelope-based model can better predict

human thresholds in a roving-level condition than an energy model can (Richards, 1992), and envelope-based models have been shown to explain other auditory tasks (Dau, Kollmeier, & Kohlrausch, 1997; Dau, Verhey, & Kohlrausch, 1999; Davidson, Gilkey, Colburn, & Carney, 2009; Mao & Carney, 2015; Mao, Vosoughi, & Carney, 2013). The envelope-based encoding strategy also has a strong physiological basis. The envelope of a sound waveform is embedded in the temporal responses of AN fibers (i.e. in the slow fluctuations of post-stimulus time histograms, PSTHs) (Dreyer & Delgutte, 2006; Joris & Yin, 1992; Shamma, 1985). Neurons in the IC are tuned to the modulation frequency of sinusoidally amplitude-modulated (SAM) sounds (e.g., Krishna & Semple, 2000; Langner & Schreiner, 1988; Nelson & Carney, 2007; review: Joris, Schreiner, & Rees, 2004), as described by the modulation transfer function (MTF); that is, IC neurons are sensitive to envelope frequency and modulation depth. However, SAM sounds are periodic stimuli, but most sounds encountered in real life or in psychophysical experiments are not (such as TIN stimuli). IC neurons have good timing precision in response to SAM sounds (e.g., Griffin, Bernstein, Ingham, & McAlpine, 2005; Krishna & Semple, 2000; Langner & Schreiner, 1988; Rees & Moller, 1983), as well as stimuli with steep envelopes (i.e. they phase-lock to the stimulus envelope) (Lee, Osman, Volgushev, Escabi, & Read, 2016; Zheng & Escabi, 2013); even in response to non-periodic stimuli, response timing is reliable and repeatable (Keller & Takahashi, 2000).

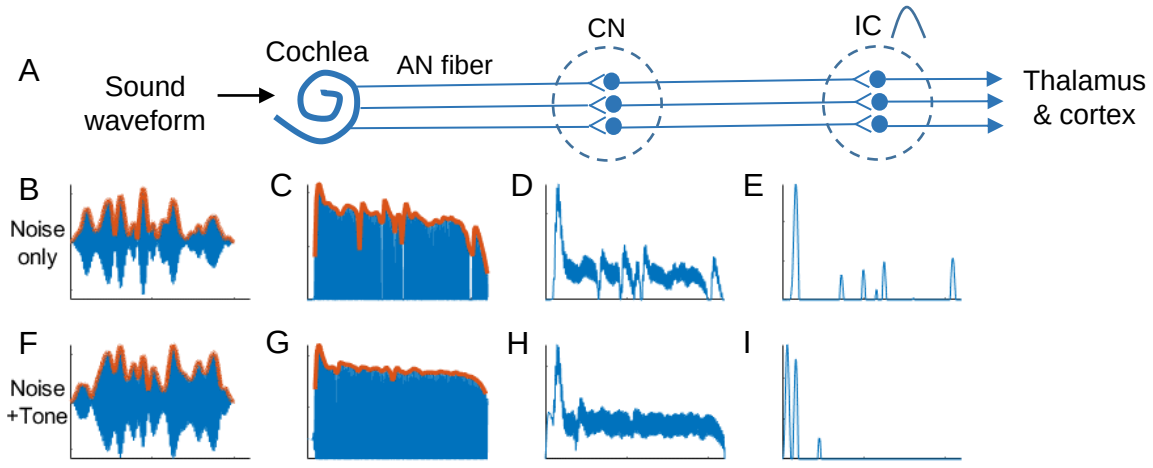

Figure 1.1: Simplified diagram of auditory pathway (A) and model responses to noise (B-E) and tone-plus-noise (F-I). *A)* Incoming sound waveform is first processed in the cochlea, then encoded by auditory-nerve (AN) fibers, the cochlear nucleus (CN), inferior colliculus (IC), and higher auditory levels. *B)* Narrowband gaussian noise. *C to E)* model responses (probability of firing vs. time) to the narrowband noise at the level of AN, CN, and IC (band-enhanced neuron). *F)* 0-dB SNR tone embedded in the narrowband gaussian noise. *G to I)* model responses to tone-plus-noise. Red curves in B, C, F, and G indicate stimulus envelope and AN neural fluctuation. Note that amplitude of AN fluctuation in C is larger than in G, in D is larger than in H, and rates in E is larger than I, upon addition of a tone.

1.2 Binaural TIN detection: in-phase vs. out-of-phase tones

Simply inverting the phase of the tone presented to one ear relative to the other (i.e. out-of-phase tones, S_π) in the presence of identical noise at the two ears (referred to as the N_0S_π condition) can result in a much lower detection threshold than for an in-phase tone (N_0S_0). The threshold difference between N_0S_0 and N_0S_π conditions is called the binaural masking level difference (BMLD, Hirsh, 1948). This phenomenon was first explained by Jeffress using a binaural model for spatial localization (Jeffress, Blodgett, Sandel, & Wood, 1956). In this model, each side of the brain is hypothesized to have a "delay net," in which neurons are sensitive to a specific delay between the inputs from the

two ears. Therefore, for an out-of-phase tone but identical noise at the two ears, representations of the tone and of the noise masker can be separated and the tone is easier to detect. Later, the equalization-cancellation (EC) model was proposed to explain BMLDs (Durlach, 1964a, 1964b). The EC model first determines whether the signals from the two ears are identical (equalization process); if not, the signals are subtracted one from another (cancellation process). Bernstein and Trahiotis later proposed a model based on normalized cross-correlation (Bernstein & Trahiotis, 1997, 2014, 2017) in which the waveforms presented at the two ears are processed before the step of cross-correlation, mainly by a gammatone filter with the “critical bandwidth” and an envelope compressor. All of these models have been shown to successfully explain some human detection thresholds. However, most of these models are based on signal processing of the stimulus waveforms; the physiological basis of the models was assumed to be the sensitivity of central auditory neurons to interaural differences in the stimuli.

When the sound source is not centered in front of the head, the ear closer to the source will receive the sound earlier and at a higher sound level than the opposite ear. The resulting interaural time (phase) and level differences (ITD/IPD, ILD) are proposed to be cues for sound localization. Neurons in the medial superior olive are sensitive to ITD or IPD (Goldberg & Brown, 1969; Yin & Chan, 1990), whereas neurons in the lateral superior olive are sensitive to ILD (Boudreau & Tsuchitani, 1968; Tollin, Koka, & Tsai, 2008; Tsuchitani, 1997; review: T. C. T. Yin, 2002). A peak ITD delay function (rate vs. ITDs) indicates that the neuron responds more strongly to small or zero ITDs than to large ITDs, whereas neurons with trough ITD delay functions respond in the opposite manner. Adding an out-of-phase tone with increasing level to identical (diotic) noise introduces ITDs to the stimulus, and at high tone levels, the 180°-phase difference of the tone

dominates the ITD. Therefore, as stimulus ITD increases from zero with increasing tone level, neurons with peak ITD delay functions are expected to have decreasing rate with increasing tone level, while neurons with trough ITD delay functions are expected to have increasing rate with increasing tone level. This hypothesis was examined in the inferior colliculus (IC) with a 500-Hz tone embedded in a gaussian noise, and neurons responded as expected (Jiang, McAlpine, & Palmer, 1997). However, if detection of an antiphasic tone were only based on ITD sensitivity, the BMLD would only exist at low frequencies (i.e. <1.5 kHz), as the frequency range of significant ITD cues is limited by the size of the head (Yin, 2002). However, listeners have BMLDs at frequencies well above 1.5 kHz (Goupell, 2012; van de Par & Kohlrausch, 1999). Previous studies have proposed that envelope-based cross-correlation is available at high frequencies (Bernstein & Trahiotis, 1996; Durlach, 1964b), but there have been no physiological studies to test this hypothesis.

1.3 Proposed envelope-encoding strategy

Based on the existing literature, we proposed an envelope-encoding strategy up to the level of IC to explain TIN detection and tested it with IC neural responses in this study. Figure 1.1 shows a flow chart that describes the processing stages (A) and model responses at each stage (B – I). Narrowband Gaussian noise maskers are used in this illustration. When sounds arrive at the cochlea, the basilar membrane vibrates so that the inner hair cells (IHCs) convert the mechanical signal into an electrical signal and stimulate the AN fibers. Note that most AN fibers have high spontaneous rates and saturate at medium-to-high sound levels (Liberman, 1978); that is, responses rates of these AN fibers do not change much with changes in stimulus energy (overall height in C vs. G). Therefore, the encoded envelopes in the slow fluctuation of AN responses, and in particular, the

amplitude of neural fluctuations, are similar across levels, thus providing a possible explanation for human performance in the roving-level paradigm. These high-spontaneous-rate fibers are proposed to form the basis of the envelope-encoding strategy, and descriptions in this study will be limited to this type of AN fiber. For the noise-alone presentation, the AN response fluctuates as the stimulus envelope fluctuates. When adding a tone, the stimulus envelope flattens, so the amplitude of AN neural fluctuations decreases. IHCs also play a critical role upon addition of a tone. IHCs have a nonlinear relationship between input sound level and response voltage (Russelll & Sellick, 1978). When adding a tone, the energy of the incoming sound increases and drives the IHC input/output transduction nonlinearity into saturation. This saturation flattens the neural fluctuations of AN responses. In summary, the amplitudes of AN neural fluctuation flatten upon addition of a tone, due to both the flattened envelope of the stimulus and the IHC nonlinearity. As Fig. 1.1C and G show, the red curve in C is more fluctuating than in G. The amplitude of neural fluctuations is enhanced in the cochlear nucleus (D vs. H) and later processed in the IC, reflected by the average rates of IC neurons. IC neurons are sensitive to envelope, and tuned to amplitude modulation, described by their MTF shape (Review: Joris et al., 2004). Band-enhanced MTF shape indicates that a neuron is excited by more modulated sounds, and thus is hypothesized to respond more strongly to larger fluctuations in general (e.g., narrowband gaussian noise). For example, in Fig. 1.1, the model IC neuron has a band-enhanced MTF; responses to noise (E) are stronger than to tone-plus-noise (I). A band-suppressed MTF shape, opposite to the band-enhanced MTF, indicates that a neuron responds more strongly to less modulated sounds, and thus is hypothesized to respond more strongly to lower amplitude fluctuations in general (e.g.,

tone-plus-noise). Therefore, changes in the average rates of IC neurons to tone-plus-noise stimuli are hypothesized to be affected by their MTF shape.

1.4 Overview of the thesis

Envelope-based cues have been shown to explain human N_0S_0 TIN detection thresholds when the energy model fails in diotic conditions, such as the roving-level paradigm (Richards, 1992), and to predict human detection performance better than classical ITD and ILD cues in the N_0S_π condition (Mao & Carney, 2014). Envelope cues also have a strong physiological basis, but they have not been tested with TIN stimuli. The goal of this thesis is to test the hypothesis that neurons in the IC respond to TIN stimuli according to their MTF shape, for both N_0S_0 and N_0S_π conditions. Single-neuron responses were recorded with a wide range of noise levels and tone levels for TIN stimuli.

For the N_0S_0 condition, two types of noise masker were used: narrowband gaussian noise (GN) and low-noise noise (LNN), which is narrowband but with flat envelope. These two maskers have different envelope statistics which help identify representations of envelope cues in IC neural responses to TIN stimuli. For N_0S_π conditions, prediction of the changes in IC neural responses as a function of SNR were compared between ITD-, interaural-correlation- and envelope-based hypotheses. The differences in neural thresholds between N_0S_0 and N_0S_π conditions (i.e. BMLDs) were calculated and compared with available human and rabbit behavioral data.

The next two chapters, both presented in the form of journal papers, describe the experimental design and results. Chapter 2 describes the IC neural responses to diotic tone-in-GN and tone-in-LNN stimuli, including both rate and timing analyses. Chapter 3 describes the neural responses to tone-in-GN stimuli for both N_0S_0 and N_0S_π conditions,

as well as neural BMLDs. Chapter 4 includes a general discussion, conclusion, and suggested future directions. The appendix presents a publication that describes simulation of neural responses to diotic tone-in-GN data with a monaural phenomenological envelope-based model that reflects the general hypothesis of this thesis.

Bibliography

Bernstein, L. R., & Trahiotis, C. (1996). On the use of the normalized correlation as an index of interaural envelope correlation. *J Acoust Soc Am, 100*(3), 1754-1763. doi:10.1121/1.416072

Bernstein, L. R., & Trahiotis, C. (1997). The effects of randomizing values of interaural disparities on binaural detection and on discrimination of interaural correlation. *J Acoust Soc Am, 102*(2 Pt 1), 1113-1120. doi:10.1121/1.419863

Bernstein, L. R., & Trahiotis, C. (2014). Accounting for binaural detection as a function of masker interaural correlation: effects of center frequency and bandwidth. *J Acoust Soc Am, 136*(6), 3211. doi:10.1121/1.4900830

Bernstein, L. R., & Trahiotis, C. (2017). An interaural-correlation-based approach that accounts for a wide variety of binaural detection data. *Journal of the Acoustical Society of America, 141*(2), 1150-1160. doi:10.1121/1.4976098

Boudreau, J. C., & Tsuchitani, C. (1968). Binaural interaction in the cat superior olive S segment. *Journal of Neurophysiology, 31*(3), 442-454. doi:10.1152/jn.1968.31.3.442

Carney, L. H. (2018). Supra-Threshold Hearing and Fluctuation Profiles: Implications for Sensorineural and Hidden Hearing Loss. *J Assoc Res Otolaryngol, 19*(4), 331-352. doi:10.1007/s10162-018-0669-5

Costalupes, J. A. (1985). Representation of Tones in Noise in the Responses of Auditory-Nerve Fibers in Cats .1. Comparison with Detection Thresholds. *Journal of Neuroscience, 5*(12), 3261-3269. Retrieved from <Go to ISI>://WOS:A1985AXC3000015

Dau, T., Kollmeier, B., & Kohlrausch, A. (1997). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *J Acoust Soc Am, 102*(5 Pt 1), 2892-2905. doi:10.1121/1.420344

Dau, T., Verhey, J., & Kohlrausch, A. (1999). Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers. *J Acoust Soc Am, 106*(5), 2752-2760. doi:10.1121/1.428103

Davidson, S. A., Gilkey, R. H., Colburn, H. S., & Carney, L. H. (2009). An evaluation of models for diotic and dichotic detection in reproducible noises. *Journal of the Acoustical Society of America, 126*(4), 1906-1925. doi:10.1121/1.3206583

Delgutte, B. (1990). Physiological mechanisms of psychophysical masking: observations from auditory-nerve fibers. *J Acoust Soc Am, 87*(2), 791-809. doi:10.1121/1.398891

Dreyer, A., & Delgutte, B. (2006). Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: Implications for sound localization. *Journal of Neurophysiology, 96*(5), 2327-2341. doi:10.1152/jn.00326.2006

Durlach, N. I. (1964a). Note on Binaural Masking-Level Differences as Function of Interaural Correlation of Masking Noise. *Journal of the Acoustical Society of America, 36*(9), 1613-&. doi:Doi 10.1121/1.1919254

Durlach, N. I. (1964b). Note on Binaural Masking-Level Differences at High Frequencies. *Journal of the Acoustical Society of America, 36*(3), 576-&. doi:Doi 10.1121/1.1919006

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics, 12*(1), 0047-0065. doi:DOI 10.1103/RevModPhys.12.47

Geisler, C. D., Rhode, W. S., & Kennedy, D. T. (1974). Responses to tonal stimuli of single auditory nerve fibers and their relationship to basilar membrane motion in the squirrel monkey. *Journal of Neurophysiology, 37*(6), 1156-1172. doi:10.1152/jn.1974.37.6.1156

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. doi:10.1152/jn.1969.32.4.613

Goupell, M. J. (2012). The role of envelope statistics in detecting changes in interaural correlation. *J Acoust Soc Am, 132*(3), 1561-1572. doi:10.1121/1.4740498

Greenwood, D. (1961). Auditory Masking and Critical Band. *Journal of the Acoustical Society of America, 33*(4), 484-&. doi:Doi 10.1121/1.1908699

Griffin, S. J., Bernstein, L. R., Ingham, N. J., & McAlpine, D. (2005). Neural sensitivity to interaural envelope delays in the inferior colliculus of the guinea pig. *Journal of Neurophysiology, 93*(6), 3463-3478. doi:10.1152/jn.00794.2004

Guinan, J. J., Jr., & Gifford, M. L. (1988). Effects of electrical stimulation of efferent olivocochlear neurons on cat auditory-nerve fibers. III. Tuning curves and thresholds at CF. *Hear Res, 37*(1), 29-45. doi:10.1016/0378-5955(88)90075-5

Hall, J. W., Haggard, M. P., & Fernandes, M. A. (1984). Detection in noise by spectro-temporal pattern analysis. *J Acoust Soc Am, 76*(1), 50-56. doi:10.1121/1.391005

Helmholtz, H. L. F., & Ellis, A. J. (1875). On the sensation of sound in general.

Hirsh, I. J. (1948). Binaural summation and interaural inhibition as a function of the level of masking noise. *Am J Psychol, 61*(2), 205-213. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18859554

Irino, T., & Patterson, R. D. (1997). A time-domain, level-dependent auditory filter: The gammachirp. *Journal of the Acoustical Society of America, 101*(1), 412-419. doi:Doi 10.1121/1.417975

Jeffress, L. A., Blodgett, H. C., Sandel, T. T., & Wood, C. L. (1956). Masking of Tonal Signals. *Journal of the Acoustical Society of America, 28*(3), 416-426. doi:Doi 10.1121/1.1908346

Jiang, D., McAlpine, D., & Palmer, A. R. (1997). Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions. *Journal of Neurophysiology, 77*(6), 3085-3106.

Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Reviews, 84*(2), 541-577. doi:10.1152/physrev.00029.2003

Joris, P. X., & Yin, T. C. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am, 91*(1), 215-232. doi:10.1121/1.402757

Keller, C. H., & Takahashi, T. T. (2000). Representation of temporal features of complex sounds by the discharge patterns of neurons in the owl's inferior colliculus. *Journal of Neurophysiology, 84*(5), 2638-2650. doi:10.1152/jn.2000.84.5.2638

Kidd, G., Jr., Mason, C. R., Brantley, M. A., & Owen, G. A. (1989). Roving-level tone-in-noise detection. *J Acoust Soc Am, 86*(4), 1310-1317.

Kohlrausch, A., Fassel, R., vanderHeijden, M., Kortekaas, R., vandePar, S., Oxenham, A. J., & Puschel, D. (1997). Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations. *Acustica, 83*(4), 659-669.

Krishna, B. S., & Semple, M. N. (2000). Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *Journal of Neurophysiology, 84*(1), 255-273. doi:10.1152/jn.2000.84.1.255

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. doi:10.1152/jn.1988.60.6.1799

Lee, C. M., Osman, A. F., Volgushev, M., Escabi, M. A., & Read, H. L. (2016). Neural spike-timing patterns vary with sound shape and periodicity in three auditory cortical fields. *Journal of Neurophysiology, 115*(4), 1886-1904. doi:10.1152/jn.00784.2015

Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. *J Acoust Soc Am, 63*(2), 442-455. doi:10.1121/1.381736

May, B. J., & Sachs, M. B. (1992). Dynamic range of neural rate responses in the ventral cochlear nucleus of awake cats. *Journal of neurophysiology*, *68*(5), 1589-1602.

Mao, J., & Carney, L. H. (2014). Binaural detection with narrowband and wideband reproducible noise maskers. IV. Models using interaural time, level, and envelope differences. *J Acoust Soc Am, 135*(2), 824-837. doi:10.1121/1.4861848

Mao, J., & Carney, L. H. (2015). Tone-in-noise detection using envelope cues: comparison of signal-processing-based and physiological models. *J Assoc Res Otolaryngol, 16*(1), 121-133. doi:10.1007/s10162-014-0489-1

Mao, J., Vosoughi, A., & Carney, L. H. (2013). Predictions of diotic tone-in-noise detection based on a nonlinear optimal combination of energy, envelope, and fine-structure cues. *J Acoust Soc Am, 134*(1), 396-406. doi:10.1121/1.4807815

Meddis, R., & O'Mard, L. (1997). A unitary model of pitch perception. *J Acoust Soc Am, 102*(3), 1811-1820. doi:10.1121/1.420088

Moore, B. C., & Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *J Acoust Soc Am, 74*(3), 750-753. doi:10.1121/1.389861

Moore, B. C., Peters, R. W., & Glasberg, B. R. (1990). Auditory filter shapes at low center frequencies. *J Acoust Soc Am, 88*(1), 132-140. doi:10.1121/1.399960

Moore, B. C. J. (2007). Basic auditory processes involved in the analysis of speech sounds. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1493), 947-963 %@ 0962-8436.

Nelson, P. C., & Carney, L. H. (2007). Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *Journal of Neurophysiology, 97*(1), 522-539. doi:10.1152/jn.00776.2006

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *J Acoust Soc Am, 59*(3), 640-654. doi:10.1121/1.380914

Patterson, R. D. (1986). Auditory filters and excitation patterns as representations of frequency resolution. *Frequency selectivity in hearing*.

Rees, A., & Moller, A. R. (1983). Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hear Res, 10*(3), 301-330. doi:10.1016/0378-5955(83)90095-3

Richards, V. M. (1992). The detectability of a tone added to narrow bands of equal-energy noise. *J Acoust Soc Am, 91*(6), 3424-3435. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1619118

Richards, V. M. (2002). Varying feedback to evaluate detection strategies: the detection of a tone added to noise. *J Assoc Res Otolaryngol, 3*(2), 209-221. doi:10.1007/s101620020034

Richards, V. M., & Nekrich, R. D. (1993). The incorporation of level and level-invariant cues for the detection of a tone added to noise. *J Acoust Soc Am, 94*(5), 2560-2574. doi:10.1121/1.407368

Roman, N., Wang, D., & Brown, G. J. (2003). Speech segregation based on sound localization. *J Acoust Soc Am, 114*(4 Pt 1), 2236-2252. doi:10.1121/1.1610463

Ruggero, M. A. (1992). Responses to sound of the basilar membrane of the mammalian cochlea. *Curr Opin Neurobiol, 2*(4), 449-456. doi:10.1016/0959-4388(92)90179-o

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., & Robles, L. (1997). Basilar-membrane responses to tones at the base of the chinchilla cochlea. *Journal of the Acoustical Society of America, 101*(4), 2151-2163. doi:Doi 10.1121/1.418265

Russell, I. J., & Sellick, P. M. (1978). Intracellular studies of hair cells in the mammalian cochlea. *The Journal of physiology*, *284*(1), 261-290.

Schalk, T. B., & Sachs, M. B. (1980). Nonlinearities in auditory-nerve fiber responses to bandlimited noise. *J Acoust Soc Am, 67*(3), 903-913. doi:10.1121/1.383970

Shailer, M. J., & Moore, B. C. (1983). Gap detection as a function of frequency, bandwidth, and level. *J Acoust Soc Am, 74*(2), 467-473. doi:10.1121/1.389812

Shamma, S. A. (1985). Speech processing in the auditory system. I: The representation of speech sounds in the responses of the auditory nerve. *J Acoust Soc Am, 78*(5), 1612-1621. doi:10.1121/1.392799

Smith, R. L. (1979). Adaptation, saturation, and physiological masking in single auditory-nerve fibers. *J Acoust Soc Am, 65*(1), 166-178. doi:10.1121/1.382260

Tollin, D. J., Koka, K., & Tsai, J. J. (2008). Interaural level difference discrimination thresholds for single neurons in the lateral superior olive. *Journal of Neuroscience, 28*(19), 4848-4860. doi:10.1523/JNEUROSCI.5421-07.2008

Tsuchitani, C. (1997). Input from the medial nucleus of trapezoid body to an interaural level detector. *Hear Res, 105*(1-2), 211-224. doi:10.1016/s0378-5955(96)00212-2

van de Par, S., & Kohlrausch, A. (1999). Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters. *J Acoust Soc Am, 106*(4 Pt 1), 1940-1947. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10530018

Wier, C. C., Jesteadt, W., & Green, D. M. (1977). Frequency Discrimination as a Function of Frequency and Sensation Level. *Journal of the Acoustical Society of America, 61*(1), 178-184. doi:Doi 10.1121/1.381251

Yin, T. C., & Chan, J. C. (1990). Interaural time sensitivity in medial superior olive of cat. *Journal of Neurophysiology, 64*(2), 465-488. doi:10.1152/jn.1990.64.2.465

Yin, T. C. T. (2002). Neural mechanisms of encoding binaural localization cues in the auditory brainstem. In *Integrative functions in the mammalian auditory pathway* (pp. 99-159): Springer.

Zheng, Y., & Escabi, M. A. (2013). Proportional spike-timing precision and firing reliability underlie efficient temporal processing of periodicity and envelope shape cues. *Journal of Neurophysiology, 110*(3), 587-606. doi:10.1152/jn.01080.2010

Chapter 2. Responses to Diotic Tone-in-Noise Stimuli in the Inferior Colliculus

2.1 Abstract

Human detection thresholds in diotic tone-in-noise (TIN) paradigms are often explained by the power-spectrum model, which evaluates the stimulus energy at the output of the auditory filter centered at the tone frequency. However, human thresholds are minimally affected when stimulus energy is made less reliable; therefore, the stimulus envelope has been proposed as a cue for TIN detection. The stimulus envelope is encoded in the slow fluctuations of responses of auditory-nerve (AN) fibers, and is affected by cochlear compression and inner-hair-cell (IHC) saturation. Here, IC responses were hypothesized to reflect changes in both the stimulus envelope and neural fluctuations upon addition of a tone to narrowband gaussian noise (GN) and low-noise noise (LNN). Neurons in the inferior colliculus (IC) are sensitive to amplitude-modulation depth and frequency, as described by modulation depth and transfer functions (MTFs). When adding a tone to GN, the fluctuation amplitudes in AN responses are reduced; in contrast, adding a near-threshold tone to LNN increases fluctuation amplitudes in the stimulus and model AN responses. Response rates of neurons with band-enhanced (BE) MTFs were expected to be positively correlated with fluctuation amplitudes, while response rates of band-suppressed (BS) neurons were expected to be negatively correlated. As predicted, of the neurons with measurable TIN-detection thresholds, most BE neurons had decreasing rate with increasing tone level for GN maskers and increasing rate for LNN maskers. Most BS neurons had increasing rate as tone level was increased for GN maskers and decreasing rate for LNN maskers. In addition, IC rate-based thresholds were comparable with

published human and rabbit behavioral data, and timing-based thresholds were more sensitive for the GN masker.

2.2 Introduction

Tone-in-noise (TIN) detection has been studied intensely over the past eighty years. The prevalent theory, the power-spectrum (or energy-based) model, is based on energy at the output of an auditory filter tuned at the frequency of the target tone (Fletcher, 1940; Greenwood, 1961; Moore, 2007; Patterson, 1974). However, energy-based models fail to explain detection thresholds for equal-energy noise waveforms (Richards, 1992), or the minimal effect on threshold when energy cues are made less reliable using a roving-level paradigm (randomly varying the sound level from interval to interval) (Kidd, Mason, Brantley, & Owen, 1989; Richards, Heller, & Green, 1991). These results suggest that human listeners can use cues other than energy for masked detection.

Adding a tone to a narrowband gaussian noise (GN) tends to flatten the stimulus envelope, in addition to increasing the energy (Richards, 1992). Envelope-based models can outperform the energy model in explaining differences in detection of tones across different reproducible noise waveforms (Davidson, Gilkey, Colburn, & Carney, 2009b; Mao, Vosoughi, & Carney, 2013). Neural responses provide a physiological basis for envelope-based models: auditory-nerve (AN) fibers synchronize to the envelope of complex sounds (review: Joris, Schreiner, & Rees, 2004). The stimulus envelope is reflected in the slow fluctuations of AN responses (i.e. in the envelopes of post stimulus time histograms, PSTHs), but this signal is distorted by peripheral nonlinearities. Particularly, the saturating transduction of inner hair cells (IHCs) results in capture, which flattens AN PSTHs in response to a tone near the fiber's characteristic frequency (CF) (Deng & Geisler, 1987;

Zilany & Bruce, 2007). Therefore, AN fluctuation amplitudes are hypothesized to be reduced upon addition of a tone to a noise masker (review: Carney, 2018). This change in fluctuation amplitude is preserved in the cochlear nucleus (Gai & Carney, 2006, 2008) and is reflected in the response rates of neurons in the inferior colliculus (IC), which are sensitive to modulation depth and to the frequency of sinusoidally amplitude-modulated (AM) sounds (Krishna & Semple, 2000; Nelson & Carney, 2007). In the current study, two types of masker, narrowband GN and low-noise noise (LNN), were chosen to further challenge energy- and envelope-based models; these maskers have similar energy, but different envelope statistics.

Sensitivity to the envelope in the IC has not previously been considered as a neural mechanism for TIN detection. Several studies in the IC have focused on changes in rate-level functions (RLFs) in response to tones at the CF with/without background noise, as a physiological representation of the energy model (Ramachandran, Davis, & May, 2000; Rees & Palmer, 1988; Rocchi & Ramachandran, 2018). Other physiological studies of binaural TIN detection include responses to diotic TIN stimuli for comparison with dichotic responses, but only for 500-Hz tones (e.g., (Jiang, McAlpine, & Palmer, 1997a, 1997b).

Sensitivities of IC neurons to envelope fluctuations are described by modulation transfer functions (MTFs, rate as a function of modulation frequency). In this study, neurons with different MTF types were hypothesized to respond differently upon addition of a tone to the two types of masker used (Fig. 2.1). For the narrowband GN stimulus, IC neurons with band-enhanced (BE) MTFs were hypothesized to respond with *higher* rates to fluctuating noise-alone stimuli, and with *lower* rates upon addition of a tone that reduces the fluctuation amplitudes in the AN responses. In contrast to GN, LNN has a flat envelope, and fluctuation amplitudes increase upon addition of a low-level tone (Kohlrausch et al.,

1997). Thus, BE neurons were hypothesized to have lower rates to LNN-alone stimuli and higher rates upon addition of a tone to LNN. On the contrary, neurons with band-suppressed (BS) MTFs were hypothesized to have the opposite changes in their responses. These predictions, based on stimulus envelope and nonlinear transformations of the peripheral system, are distinct from predictions based on energy cues.

Spike timing is an important aspect of neural responses. Neural timing in IC responses has been mainly examined with periodic signals such as pure tones or SAM stimuli (e.g. Krishna & Semple, 2000), so that temporal responses can be described by how well responses are phase-locked to the fine structure or envelope. However, most sounds in real-life and those used in psychophysical experiments are not periodic, and thus it is difficult to describe the temporal response patterns. Previous studies have made efforts to quantify temporal responses to non-periodic signals (e.g., Chase & Young, 2006; Gai & Carney, 2006, 2008; Joris, Louage, Cardoen, & van der Heijden, 2006; van Rossum, 2001). Correlation index (CI, Joris et al., 2006) was designed to quantify the consistency of spike timing across repetitions of the same stimulus; temporal correlation was used to calculate temporal thresholds (Gai & Carney, 2006). Other methods were ruled out for the following reasons: Spike distance (van Rossum, 2001) and fluctuation of PSTHs (Gai & Carney, 2006) are affected by average rates, an effect not easily removed by normalization. Mutual information is more suitable to identify relative importance of different cues (e.g., ITD vs. ILD in Chase & Young, 2006). Temporal reliability (Gai & Carney, 2006) is correlated with CI, but the CI calculation involved the timing of each spike, and thus CI value was considered more informative. Temporal thresholds were also calculated in the current study, as they have been shown to outperform average rate in predicting human modulation-detection thresholds (Carney, Zilany, Huang, Abrams, &

Idrobo, 2014; Henry, Neilans, Abrams, Idrobo, & Carney, 2016). Neural detection thresholds based on rate and temporal information were both compared with reported human thresholds (Goupell, 2012; Kohlrausch et al., 1997; van de Par & Kohlrausch, 1999) and rabbit behavioral data (Zheng et al., 2002).

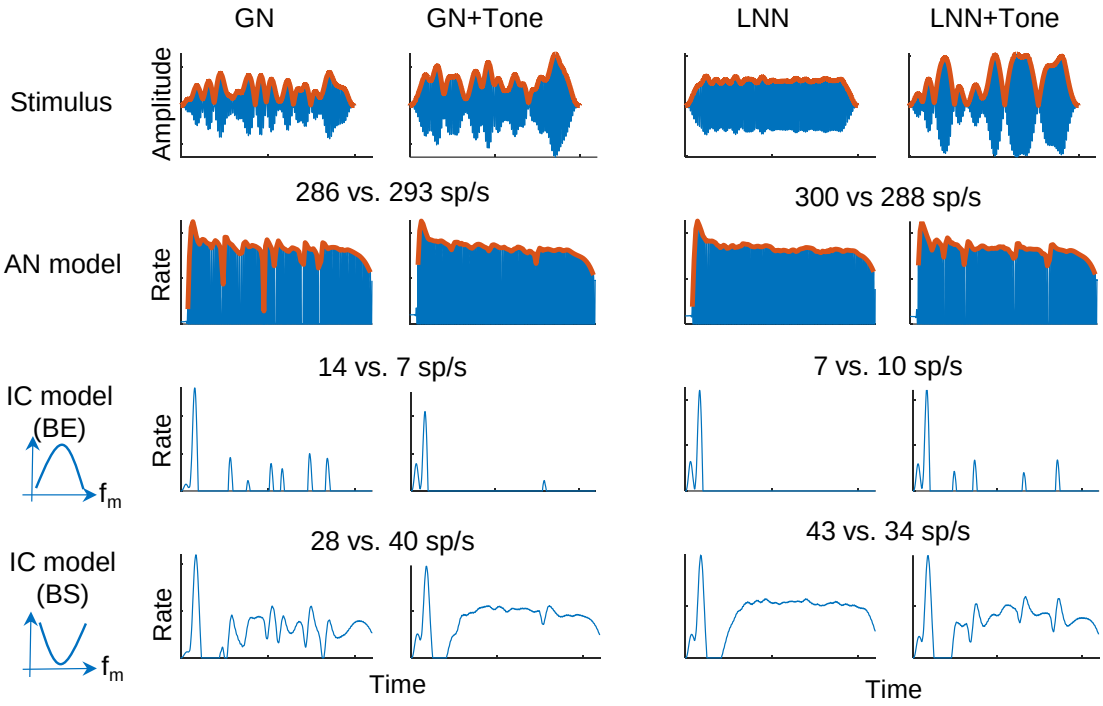

Figure 2.1: Example stimulus waveforms and responses of AN and IC models. Rows from top to bottom show: stimulus waveform (with envelope highlighted in red), high-spontaneous-rate AN model responses (with neural fluctuation highlighted in red), band-enhanced (BE) IC model responses, and band-suppressed (BS) IC model responses. Columns from left to right are for 1/3-octave narrowband gaussian noise (GN), GN with tone, low-noise noise (LNN), and LNN with tone. The tone frequency, center frequency of the noise maskers, and the model CFs were 1 kHz. Noise level is 65 dB SPL, and tone level is 0 dB SNR (slightly above human threshold). Average rates of model responses are shown in the format of “masker-alone rate vs. masker-plus-tone rate”. Model BE and BS neurons have decreasing or increasing rate upon addition of a tone, consistent with envelope-based hypothesis.

2.3 General Methods

Neurophysiological procedures were approved by the University Committee on Animal Resources. Recordings were from four female Dutch-belted rabbits with normal hearing, as monitored by distortion product otoacoustic emissions (Whitehead, Lonsburymartin, & Martin, 1992). Two rabbits were studied from age 17 to 55 months, and two rabbits were studied from age 13 to 23 months.

2.3.1 Procedures

Surgical procedures

Rabbits were anesthetized by intramuscular injection of ketamine (66 mg/kg) and xylazine (2 mg/kg) for both headbar placement and microdrive implantation surgeries. Additional injections were given as needed, based on pedal responses monitored throughout the surgery. A custom-designed, 3D-printed, hard plastic headbar with a chamber to hold the microdrive was mounted on the skull using stainless-steel screws and dental acrylic. After recovery, a craniotomy was made for insertion of the microdrive that held the tetrodes.

The microdrive (Five-drive, Neuralynx, Inc., USA) held four tetrodes. Each tetrode was twisted using four 18-μm platinum iridium wires, coated in epoxy (California Fine Wire Co., Grover Beach, CA, USA). Tetrodes were advanced slowly through the IC over the course of several weeks, before backing up and advancing again. Surgeries were performed to replace and reposition tetrodes as needed, usually every few months.

Recording procedure

All recordings were carried out in sound-attenuated booths (Acoustic Systems, Austin, Texas, USA). The rabbit sat in a custom-made chair, with the head fixed using the headbar, during daily 2-hour recording sessions. Custom ear molds (Hal-Hen Company, Inc., Garden City Park, New York, USA) made for each rabbit were positioned deep in the concha and included an Etymotic probe tube (Etymotic Research, Inc., Elk Grove Village, Illinois, USA) for calibration. The acoustic system included a MOTU audio interface (16A, Mark of the Unicorn, Cambridge, Massachusetts, USA), a Benchmark digital-to-analog converter (DAC3 HGC, Benchmark Media Systems, Inc., Syracuse, New York, USA), and Beyerdynamic DT990 (Beyerdynamic GmbH and Co., Heilbronn, Germany) or Etymotic ER2 earphones. A wideband search noise was used to locate auditory responses. Recordings were made with an RHD recording system (Intan Technologies, LLC., Los Angeles, California, USA). During the recording period, systematic increases in CF with increasing tetrode depth was used to determine that the tetrodes were located in the central nucleus of the IC (ICC); after euthanasia, electrode locations are verified histologically.

2.3.2 Stimuli

Speakers were calibrated with ER-7C or ER-10B+ probe-tube microphones (Etymotic Research, Inc., USA) at the beginning of each recording session. After calibration, several stimuli were presented to characterize all neurons. First, contralateral, ipsilateral and binaural wideband noise (0.1-19 kHz) were presented at several spectrum levels to determine the neuron's binaural sensitivity. Then interaural time difference (ITD) and interaural level difference (ILD) sensitivities were determined with wideband noise

(0.1-19 kHz). A response map, consisting of responses to pure tones between 0.25-20 kHz, presented at 10 to 70 dB SPL, was used to identify CF, the frequency at which the neuron responded at the lowest sound level. Modulation transfer functions were obtained with carriers of wideband noise (0.1-10 kHz) with a spectrum level of 30 dB SPL.

After characterizing the neurons, diotic TIN stimuli (identical at the two ears) were presented. Tone frequency was approximately equal to CF and log-centered within the 1/3 octave noise masker. Narrowband GN maskers were generated by filtering wideband noise with a 5000-order FIR band-pass filter. LNN maskers were generated using method #1 in Kohlrausch et al. (1997). Briefly, a 1/3-octave narrowband GN was normalized by its instantaneous envelope and then band-pass filtered to limit the noise bandwidth. This procedure was repeated ten times to obtain both a flat envelope in the time domain and a narrowband noise spectrum in the frequency domain. For both masker types, overall noise level varied from 35 to 75 dB SPL with a step size of 10 dB (35 dB SPL was omitted when it was below neural threshold). Signal-to-noise ratio for TIN stimuli varied from -12 to 8 dB with a step size of 4 dB. Typically, responses to 30 repetitions of each stimulus condition were recorded, with stimuli presented in a random sequence. Responses were collected for either random or reproducible TIN stimuli; if responses to both noise types were recorded for a neuron, datasets with responses to random noise waveforms were used for rate-based analyses, and datasets with responses to many repetitions of reproducible noise waveforms were used for temporal analyses.

2.3.3 Spike sorting

Spikes (action potentials) were sorted offline. Voltage recordings from each of the four tetrode wires were filtered with a 4^{th}-order Butterworth bandpass filter (0.3-6 kHz).

Potential spike waveforms were identified based on voltage threshold crossings and biphasic spikes. The initial threshold was set at 4.2 times the estimated standard deviation (σ_{est}) of the first 50-sec of the recording session, which included responses to noise stimuli. Σ_{est} was calculated as follows: $\sigma_{est} = median\left\{\frac{|x|}{\sqrt{2}*[1-erf(0.5)]}\right\}$ (Quiroga, Nadasdy, & Ben-Shaul, 2004), where x is the recorded voltage, $erf()$ is the error function and $median()$ takes the median value. The threshold was later adjusted manually if better clustering was attainable. For each threshold-crossing event on any of the four wires of one tetrode, 1-ms duration waveforms of all wires from the same tetrode were saved for sorting. Because well-isolated action potentials were biphasic, the mean amplitude of isolated spike waveforms was close to zero. Therefore, waveforms were excluded if they had mean values outside a range selected to be wide enough to include typical biphasic waveforms but narrow enough to eliminate many non-spike (often monophasic) waveforms. If more than one threshold-crossing event occurred within 0.5 ms, only the first one was included.

The sorting algorithm is described in Schwarz et al. (2012) (code available upon request). Briefly, a subset of 10,000 events were clustered using k-means with Mahalanobis distance scaled by cluster size (scale factor α = 1.5, Schwarz et al., 2012). In general, if a unit had a uniform biphasic spike shape and the percentage of interspike intervals shorter than 1-ms was less than 2%, the unit was determined to be a well-isolated single unit. If there was more than one cluster in the recording, the cluster quality was evaluated by how-well separated (vs. entangled) the clusters were, based on the Mahalanobis distance. Clusters were only considered well-isolated with a criterion of 0.1 for the cluster entanglement factor (smaller value indicates better cluster separation).

2.4 Rate analysis and Results

2.4.1 Analysis

MTF shape classification

The following rules were used to classify MTFs. The classification scheme was designed to be as simple as possible, while yielding classifications that generally agreed with a qualitative description of each MTF shape. The neurons were classified into four MTF types: band-enhanced (BE, i.e. rates were enhanced, in comparison to the response to an unmodulated wideband noise, over a band of modulation frequencies), band-suppressed (BS, i.e. rates were suppressed, in comparison to the response to unmodulated noise, over a band of modulation frequencies), hybrid (rates were both enhanced and suppressed, at different modulation frequencies), and all-pass (AP, no enhancement or suppression). Enhancement or suppression was identified with the Mann-Whitney test; at least two neighboring modulation frequencies were required to elicit significantly higher or lower response rates as compared to the unmodulated noise. If all response rates were lower than 5 spikes/s, the MTF was classified as unresponsive.

Rate analysis

The spike rate in response to each stimulus presentation was computed after excluding a 20-ms window after the stimulus onset; the mean and standard deviation across stimulus repetitions were calculated for each condition. At each noise level and SNR, a rate-based receiver-operating-characteristic (ROC, Egan, 1975) was calculated using rates in response to noise-alone and tone-plus-noise presentations. The area under the ROC curve provided an estimate of the percent-correct performance for a given

combination of tone and noise levels. The tone level resulting in 70.7%-correct performance was considered threshold, regardless of the direction of change in rate upon addition of a tone. 70.7% was used as equivalent to a psychophysical threshold obtained with a two-down, one-up tracking procedure (Levitt, 1971). Thresholds were estimated by linear interpolation between the first tone level with performance better than 70.7% and the tone level that was one step lower. Rate-based thresholds were then compared with rabbit and human behavioral thresholds.

In order to assess the trends in the rate versus SNR responses across the population, the *direction of change in rate with increasing SNR* (or *rate-change direction*) was labeled for each neuron, as follows: at each noise level, if the average rate at the ROC-based threshold was higher than for the noise-alone condition, the rate-change direction was designated as an “increase.” If the rate at threshold was lower than for the noise-alone condition, the rate-change direction was designated as a “decrease.” Otherwise, the neuron was labeled as “no-threshold.” Note that the direction of rate change was only designated at the threshold SNR; non-monotonic changes in rate at supra-threshold levels were not included in the population evaluation.

For each type of noise masker, the proportion of neurons with increasing or decreasing rate versus SNR were examined with χ^2 tests between BE and BS groups of neurons at each noise level, and a significant difference between the two groups of neurons was expected for the envelope-based hypothesis.

2.4.2 Results

Responses to tones with both GN and LNN maskers were studied in 120 isolated single units and with only GN maskers in 43 additional units. All units were tested using a

tone frequency within 1/3-oct of the neuron's CF. There were 54 BE units (33.1%), 72 BS units (44.2%), 12 hybrid units (7.4%), and 25 AP units (15.3%). The distribution of CFs for each MTF type is shown in Fig. 2.2.

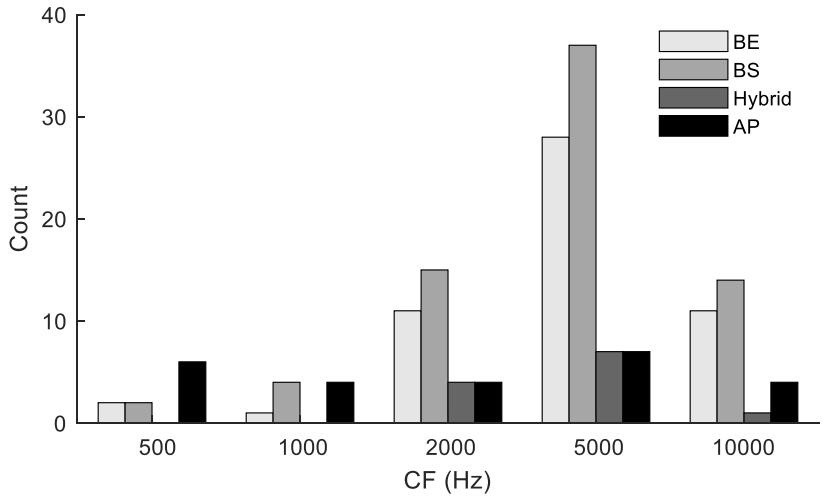

Figure 2.2: Distribution of MTFs across CF (in one-octave bins) for all units examined in this study. Gray shades from light to dark indicate units with band-enhanced (BE), band-suppressed (BS), hybrid, and all-pass (AP) MTF shapes. Most MTF types were represented across the range of CFs, except that hybrid MTFs were not observed at lower CFs. The largest numbers of units were in the frequency range for which rabbits have the most sensitive hearing (Heffner and Heffner, 2007).

Single-neuron responses to TIN

Figure 2.3 A,B shows MTFs of example BE (BE #1) and BS units (BS #1). Recall that BE or BS MTF shapes were identified as having rates in response to at least two consecutive modulation frequencies that were significantly different from the unmodulated condition (indicated by * in Fig. 2.3). IC neurons with BE MTFs (i.e. units with increased rate in response to fluctuating inputs) were expected to have decreased rate with increasing tone level, because stimulus envelopes and AN fluctuations flatten as the level

of a tone added to a GN masker increases. In contrast, neurons with BS MTFs (i.e. units with increased rate in response to unmodulated inputs) were expected to have rate that increased with increasing tone level. Figure 2.3C, D shows average rate in response to tone-in-GN as a function of SNR for different noise levels. Average rate for BE #1 in response to tone-in-GN decreased with increasing tone level; average rate for BS #1 increased with increasing tone level. Responses of these neurons support the hypothesis that the rate of BE neurons was positively correlated, and the rate of BS neurons was negatively correlated, to fluctuation amplitudes.

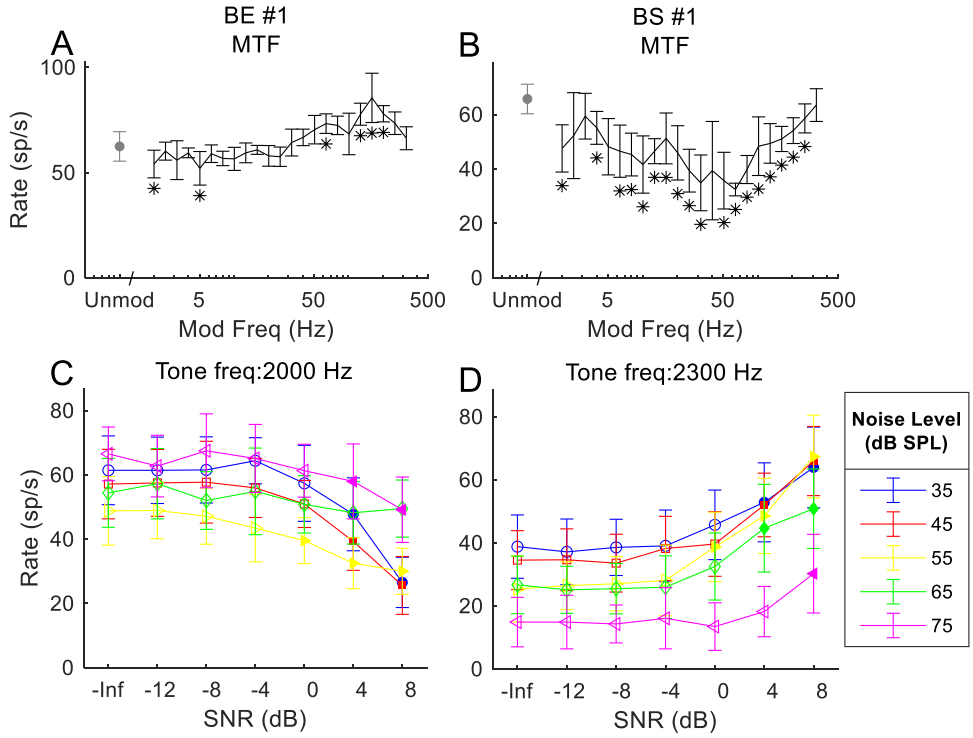

Figure 2.3: MTFs and responses to tone-in-GN for two example neurons. (A) (B) MTFs of BE #1 (band-enhanced, BE) and BS #1 (band-suppressed, BS). (C) (D) Average rate responses to tone-in-GN for BE #1 and BS #1 respectively. Errorbars indicate standard deviations. Colored lines with different symbols represent different overall noise levels. Filled symbols indicate supra-threshold SNRs based on ROC analyses of average rate. Tone frequency, near CF, is indicated in the title. BE #1 had decreasing rate as SNR increased. BS #1 had increasing rate.

Figure 2.4 shows responses of two additional example neurons to tone-in-GN (blue circles) and tone-in-LNN (red squares). MTFs are shown in the left-most column; the other columns show the two types of TIN responses as a function of SNR at five noise levels. For LNN stimuli, the flat envelopes are achieved by a specific configuration of phase across frequency (Pumplin, 1985). When a tone is added, the phase relationship is broken and the envelope is no longer flat (Kohlrausch et al., 1997) (Fig. 2.1). Therefore, for

neurons with BE MTFs, increasing rate was expected with increasing tone level for tone-in-LNN stimuli near threshold, and for neurons with BS MTFs, decreasing rate was expected. In both cases, the expected rate changes were opposite to those expected for tone-in-GN stimuli. BE #2 had decreased rate versus SNR at threshold for tone-in-GN stimuli at most noise levels and increased rate for LNN at all noise levels, as expected. BS #2 had increasing rate at threshold for tone-in-GN and decreasing rate for tone-in-LNN stimuli at all noise levels, as expected. Note a few details that further supported the hypothesis that responses are determined by amplitude fluctuations: First, in response to noise-alone stimuli at each noise level, BE #2 had higher rates for GN (fluctuating envelope) than LNN (flat envelope), whereas BS #2 had higher rates for LNN (flat envelope) than GN (fluctuating envelope). Second, the rate-based threshold was lower for LNN than for GN, consistent with changes in the normalized envelope slope that occur at lower SNRs for LNN than for GN (unpublished observations). Third, the rate versus SNR curves changed direction at high SNRs for tone-in-LNN stimuli (e.g. rate for BE #2 decreased and rate of BS #2 increased for high tone levels). This change was expected because the effect of an added tone at relatively high levels is to flatten the fluctuations in PSTHs of AN responses, due to IHC saturation. Thus, at the high tone levels, i.e. well above detection thresholds, average rate in response to tone-in-LNN stimuli changed in the same direction as for tone-in-GN stimuli.

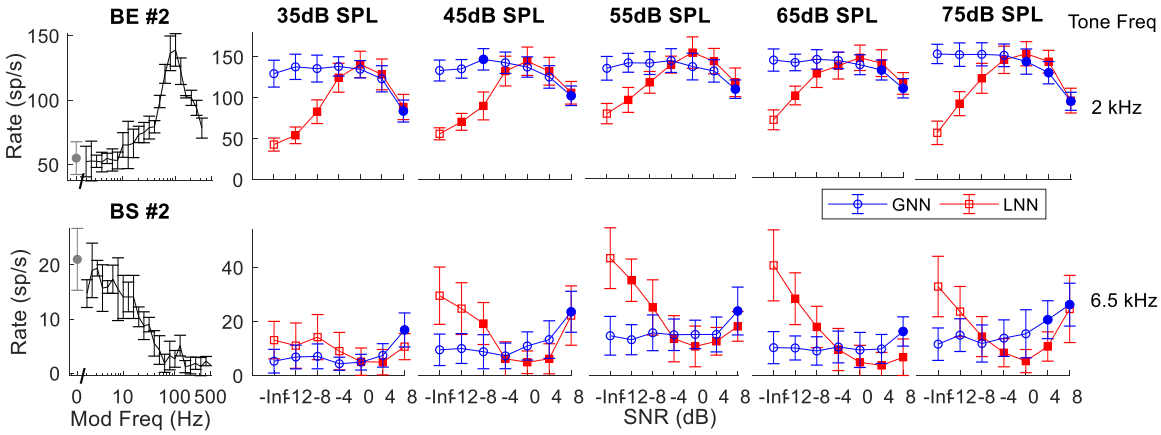

Figure 2.4: MTF and TIN responses of two example neurons. Left: MTFs. Right: Average rate in response to tone-in-GN (blue circles) and tone-in-LNN (red squares) at five overall noise levels. Filled symbols indicate supra-threshold SNRs based on ROC analyses. Tone frequency, near CF, is indicated at the right. Average rate for BE #2 decreased for tone-in-GN, except for a small increase for the 45-dB SPL noise; rate for tone-in-LNN increased. Average rate for BS #2 increased as SNR increased for the tone-in-GN and decreased for tone-in-LNN stimuli near threshold. LNN response curves change direction at high SNRs (see text).

Figure 2.5 shows average rate versus SNR across noise levels for five different example neurons. These neurons were chosen to illustrate complicated response patterns that were observed for many recorded units; each neuron had both expected and unexpected responses at different noise levels, as described below. In response to GN masker, expected responses versus SNR were observed for BE #3 at 65-75 dB SPL (decreasing rate), BS #3 at 55-75 dB SPL, and BS #4 at 35 and 75 dB SPL (increasing rate); unexpected responses versus SNR were observed for BE #3 at 35-55 dB SPL (no threshold or increasing rate), BS #3 at 35-45 dB SPL, and BS #4 at 45-65 dB SPL (decreasing rate). In response to LNN maskers, expected responses versus SNR were observed for BE #3 and BS #3 at almost all noise levels (increasing and decreasing rate,

respectively); unexpected responses vs SNR were observed for BE #3 only at 65 dB SPL (no threshold), BS #3 only at 35 dB SPL (increasing rate), and BS #4 at all noise levels. When considering responses to both GN and LNN maskers, expected responses vs SNR (opposite changes for GN and LNN) were observed for BE #3 only at 75 dB SPL, BS #3 at 55-75 dB SPL, BS #4 at 45-65 dB SPL. Because hybrid neurons had both enhancement and suppression, there was no specific prediction. Response changes among hybrid neurons varied. For example, the MTF of Hybrid #1 was similar to BE shape, and this neuron responded to GN and LNN maskers as expected for a BE neuron – decreasing rate versus SNR for GN maskers at all noise levels and increasing rate versus SNR for LNN maskers at 35-65 dB SPL. On the other hand, Hybrid #2 had an MTF similar to BS shape, but did not respond as expected for a BS neuron. Even though Hybrid #2 had increasing rate versus SNR at 35 and 45 dB SPL for GN masker, the similarity in responses to GN and LNN maskers indicated that the neuron possibly responded based on stimulus energy. At high noise levels, responses of Hybrid #2 were as expected for a BE neuron.

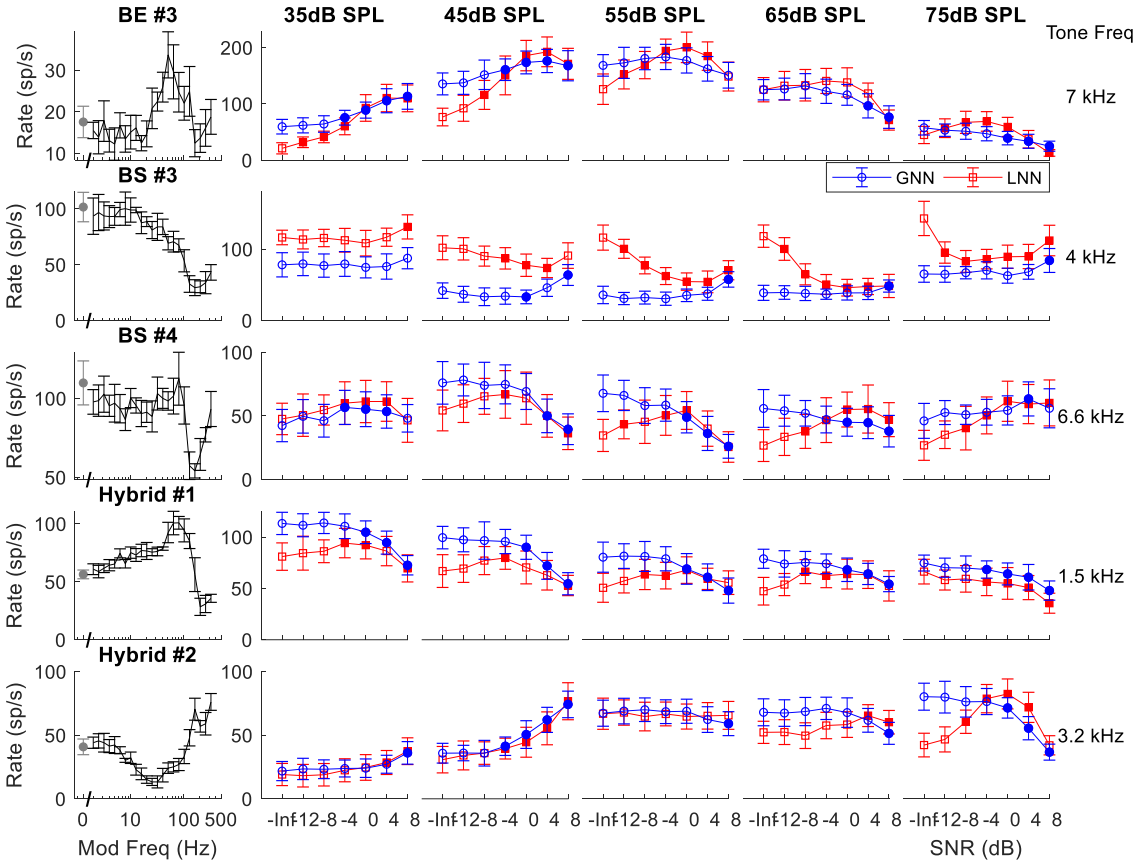

Figure 2.5: MTF and TIN responses of five example neurons. Left: MTFs. Right: tone-in-GN (circle) or tone-in-LNN (square) responses at five noise levels; filled symbols indicate suprathreshold responses based on ROC analyses. Same format as in Fig. 2.4.

Another complexity observed in the IC responses was non-monotonic rate in response to noise-alone narrowband stimuli (Figs. 2.4-2.5). Nonmonotonic rate in response to wideband noise has been observed previously (Ramachandran et al., 2000; Rees & Palmer, 1988). Responses to noise-alone stimuli at different levels for the nine example neurons are replotted in Fig. 2.6, with pure-tone RLFs extracted from response maps shown for comparison. Most neurons had non-monotonic noise RLFs for either GN or LNN, or both. The trends for noise RLFs were similar for GN and LNN in most neurons, but not necessarily similar to the pure tone RLF (e.g. BE #1, BS #3 and BS #4). Only

neurons BS #2, BE #3 and Hybrid #2 had noise RLFs that were similar to the tone RLF; in particular, the LNN RLFs of BE #2 and BS #2 were similar to the tone RLFs in both trends and rates. The differences between noise RLFs and tone RLF suggested that responses to narrowband noise could not simply be described as a sum of responses to tones across a range of frequencies near CF, even for neurons with “V” shaped response maps, which are primarily excited by a range of frequencies that broadens at higher tone levels (Ramachandran, Davis, & May, 1999). Therefore, a more complex interaction between inputs from different frequencies is suggested. The complex responses may be explained by inputs from the dorsal cochlear nucleus (Davis & Young, 2000; Young & Brownell, 1976). The range of discrepancies between noise and pure tone RLFs illustrated in Fig. 2.6 were typical for the population of neurons studied.

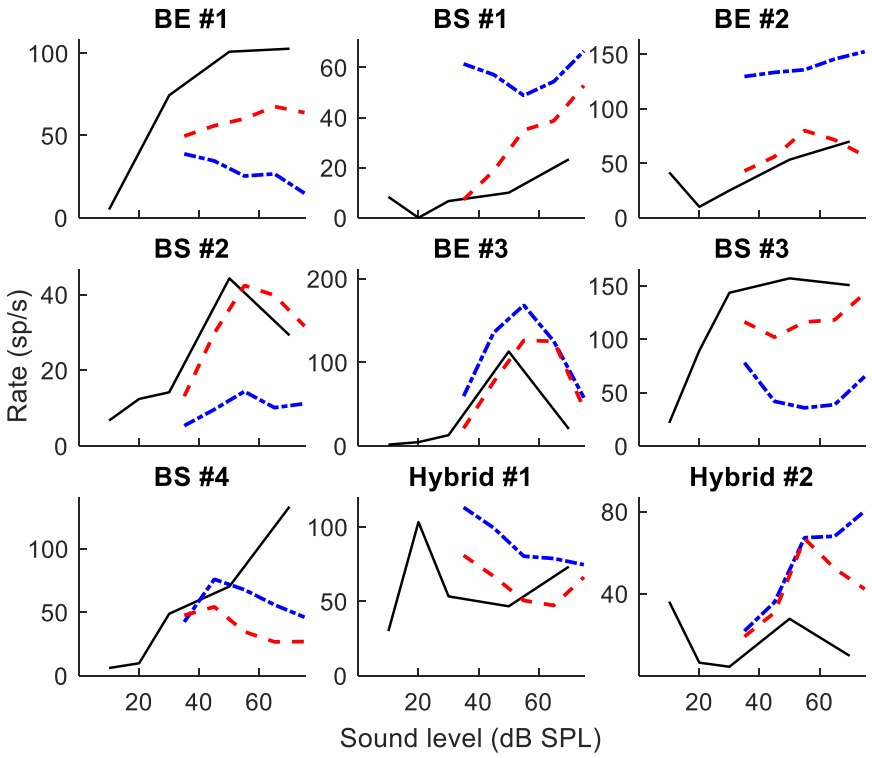

Figure 2.6: Rate-level functions (RLFs) of pure tone from response map (solid black), of GN responses (blue dot-dashed), and of LNN responses (red dashed). Titles show neuron numbers, from Figs. 2.3-2.5. Most neurons had non-monotonic RLFs, and noise RLFs often differed from tone RLFs (e.g. Neuron 5).

Trends in responses across the population

An analysis of responses across the population of neurons included tests for significant correlations between maximum changes in rate elicited by SAM stimuli (used for MTFs) and by addition of a tone to a masker (TIN responses). For SAM responses, the maximum difference between average rates in reponse to the unmodulated stimulus and across all modulation frequencies was determined for each neuron. BE neurons had a positive maximum rate difference (Fig. 2.7, triangles); BS neurons had a negative maximum rate difference (Fig. 2.7, squares); hybrid neurons could have either positive or

negative rate difference, depending upon whether enhancement or suppression was stronger (Fig. 2.7, diamonds); AP neurons were expected to have maximum rate differences near zero (Fig. 2.7, circles). Similarly, for TIN responses, the maximum difference in average rate across all SNRs tested and the average rate in response to the noise-alone stimulus was determined for each neuron, at each masker noise level.

For the GN masker, BE neurons were expected to have decreased rate with increasing SNR, and BS neurons were expected to have increased rate; therefore, the maximum rate differences for tone-in-GN responses were expected to be negatively correlated with the maximum rate differences in the MTFs. On the contrary, for the LNN masker, the maximum rate differences between tone-plus-noise and noise-alone stimuli were expected to be positively correlated with maximum rate differences in the MTFs. Results were consistent with this prediction at most noise levels: for the GN masker, the correlation coefficient was negative and significant at noise levels of 35, 55-75 dB SPL (Fig. 2.7, A), whereas for the LNN masker, the correlation coefficient was positive and significant at 55-75 dB SPL (Fig. 2.7, B). The correlation was stronger with increasing noise level for both GN and LNN condition, indicating a stronger influence of the envelope cue on responses at higher noise levels.

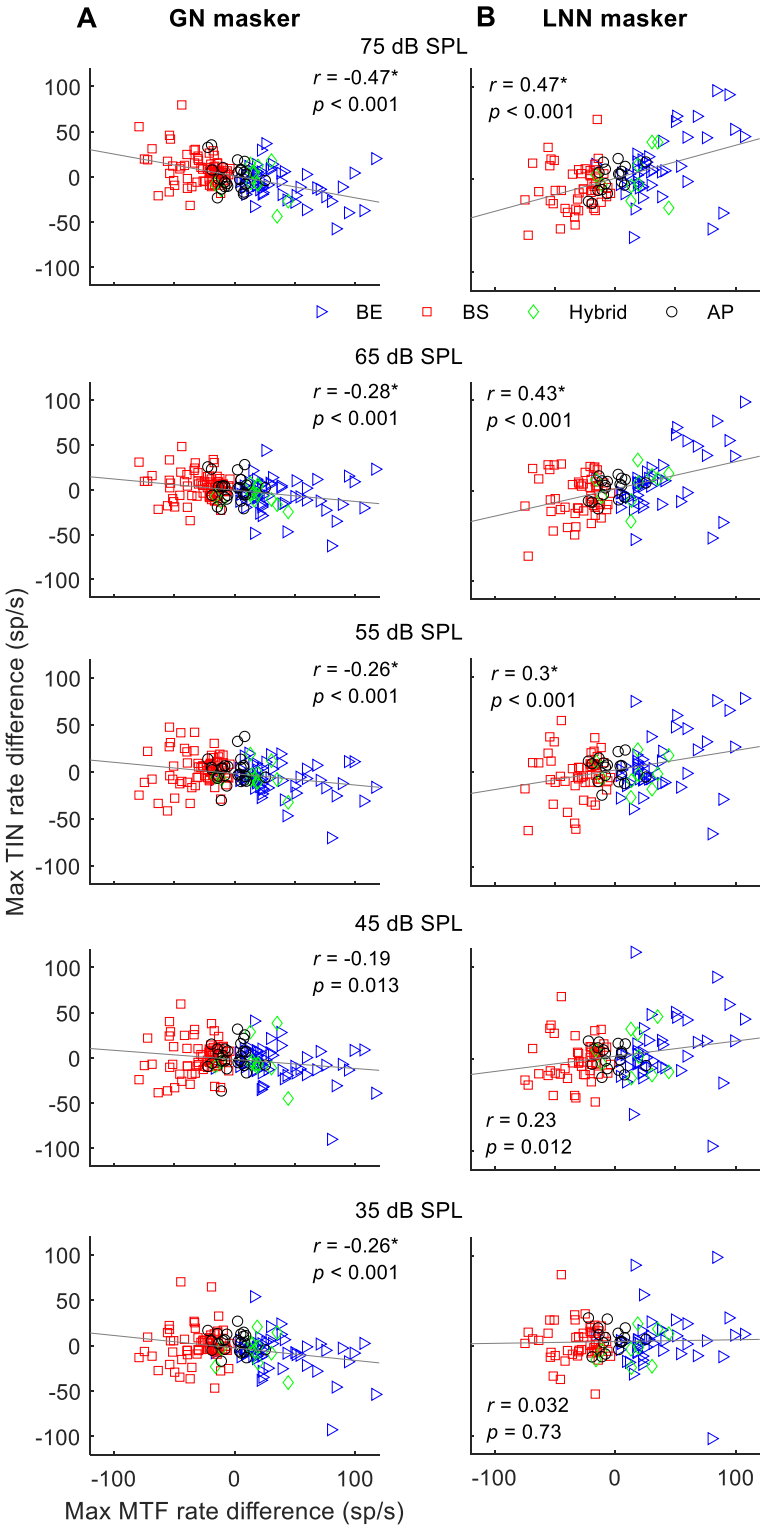
A
GN masker
B
LNN masker
75 dB SPL
r = -0.47*
p < 0.001
r = 0.47*
p < 0.001
BE
BS
Hybrid
AP
65 dB SPL
r = -0.28*
p < 0.001
r = 0.43*
p < 0.001
55 dB SPL
r = -0.26*
p < 0.001
r = 0.3*
p < 0.001
45 dB SPL
r = -0.19
p = 0.013
r = 0.23
p = 0.012
35 dB SPL
r = -0.26*
p < 0.001
r = 0.032
p = 0.73
Max TIN rate difference (sp/s)
Max MTF rate difference (sp/s)
100
50
0
-50
-100

Figure 2.7: Correlation between the maximum differences in response rate elicited by AM and addition of tone to a GN masker (A) or LNN masker (B) at each noise level. The maximum average rate difference elicited by AM was calculated between the peak or valley of the MTF and the rate in response to the unmodulated tone. The maximum rate difference elicited by addition of a tone was calculated between tone-plus-GN (A) or tone-plus-LNN (B) responses and the noise-alone response. Symbols indicate MTF types (see legend). In each plot, BE and BS neurons tend to be located on the right and left side, respectively, as expected. Correlation coefficient and the p value are shown in each plot; star indicates the correlation coefficient was significant after Bonferroni correction ($p < 0.01$). The maximum rate difference in the TIN responses was negatively correlated with that of the MTF for GN maskers, and positively correlated for LNN maskers, as expected.

Trends in the responses across the population of neurons were also evaluated by the proportion of neurons with the expected rate-change direction near threshold. The response of each neuron was labeled as increasing, decreasing, or as having no-threshold. Recall that in the responses of many neurons, especially to LNN maskers, the average rate versus SNR was non-monotonic. For these cases, the direction of rate change was only designated at the threshold SNR; non-monotonic changes in rate at supra-threshold levels were not included. For example, for the 75 dB SPL LNN masker, Hybrid #2 had an increasing rate-change direction, due to the significantly higher rate at -8 dB SNR as compared to the noise-alone condition, despite the fact that the rate decreased at higher SNRs (Fig. 6; also see rate versus SNR for LNN of BE #2 at all noise levels, Hybrid #2 at 75 dB SPL, and BS #4 at 55 dB SPL).

Figure 2.8 shows the percentage of units with measurable thresholds that had decreasing (blue) or increasing (red) rate changes, for BE or BS MTF types and for GN or LNN maskers. The proportion of BS neurons with the predicted rate-change direction increased with increasing masker level. The proportion of BE neurons with the expected

rate-change direction was more consistent across masker levels. A χ^2 test showed that for the GN masker, the ratio of units with decreasing or increasing rate was significantly different between BE and BS MTF types at all noise levels. For the LNN masker, the ratio was significantly different at noise levels of 65 and 75 dB SPL (Table 2.1).

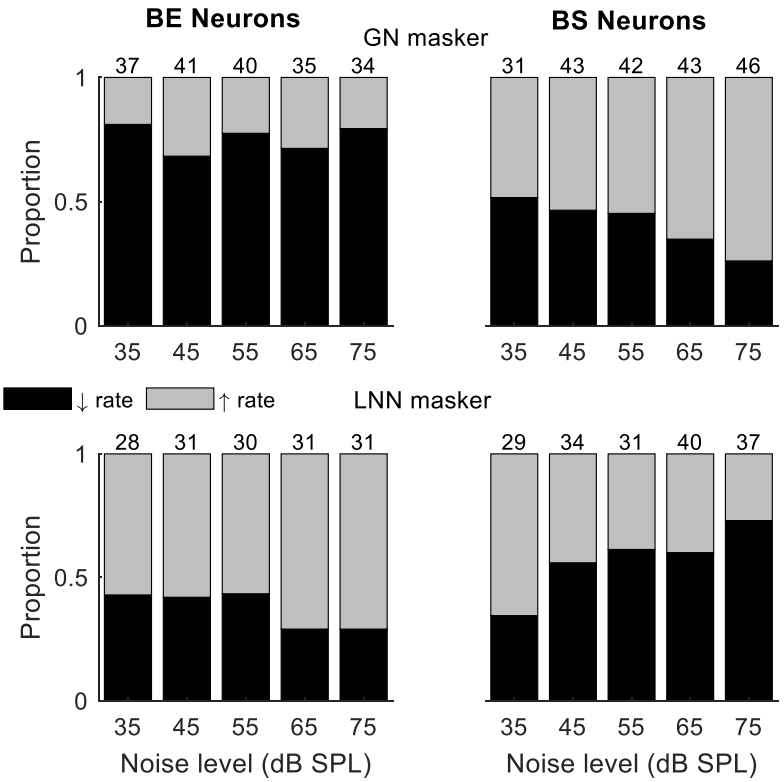

Figure 2.8 Proportion of neurons having decreasing (black) or increasing (gray) rate-changes among neurons with BE (left) and BS (right) MTF shapes in response to TIN with GN (upper plots) or LNN maskers (lower plots). In each plot, each bar shows results for one noise level. The percentage was calculated based on the total number of units (shown at the top of each bar) that had a rate-based threshold at that noise level.

Table 2.1. Statistics of χ^2 test to compare rate-change direction for BE and BS MTF groups.

	GN		LNN	
Noise levels (dB SPL)	$\chi^2(1,1)$	*p*	$\chi^2(1,1)$	*p*
35	6.69	0.0097	0.42	0.52
45	4.07	0.044	1.26	0.26
55	8.96	0.0028	1.97	0.16
65	10.31	0.0013	6.73	0.0095
75	22.25	<.0001	13.07	0.0003

The number of units with each MTF shape that had increasing or decreasing rate-change direction, or no threshold, at each noise level is shown in Table 2.2a for the GN masker and Table 2.2b for the LNN masker. As shown in Fig. 2.8, most BE and BS neurons with measurable rate-based thresholds responded with the expected rate-change direction (bold). With increasing noise levels, BS neurons not only had an increasing proportion of neurons with the expected change in rate versus SNR, but also an increasing proportion of neurons with measurable thresholds, while BE neurons were more consistent across noise levels.

There were smaller numbers of hybrid and AP neurons. There were also no specific predictions for rate-change direction for these neurons based on their MTF shapes, so only a qualitative description is provided here. Among neurons that had measurable thresholds, more hybrid neurons had decreasing rate in response to GN maskers, while approximately equal numbers had increasing or decreasing rate for LNN maskers. Many AP neurons did not have thresholds in response to GN maskers, but among those that had thresholds, more had increasing rate versus SNR at low noise levels for both GN and

LNN maskers, while approximately equal numbers of AP neurons had increasing or decreasing rate at medium-to-high noise levels.

Table 2.2a. Numbers of units with each rate-change direction at each GN level for all MTF shapes.

		Band-enhanced	Band-suppressed	Hybrid	All-pass
35 dB SPL	Decreasing↓	**30**	16	4	5
	Increasing↑	7	**15**	2	10
	No-threshold –	17	39	6	6
45 dB SPL	Decreasing↓	**28**	20	6	6
	Increasing↑	13	**23**	3	7
	No-threshold –	13	29	3	12
55 dB SPL	Decreasing↓	**31**	19	7	5
	Increasing↑	9	**23**	2	4
	No-threshold –	14	30	3	16
65 dB SPL	Decreasing↓	**25**	15	5	6
	Increasing↑	10	**28**	0	6
	No-threshold –	19	29	7	13
75 dB SPL	Decreasing↓	**27**	12	5	7
	Increasing↑	7	**34**	3	8
	No-threshold –	18	26	4	10

Note: Numbers of units with the expected rate-change direction for BE or BS MTF types are in bold font. One BE unit was not tested at 75 dB SPL; one BS and two AP units were not tested at 35 dB SPL.

Table 2.2b. The number of units with each rate-change directions at each LNN level for all MTF shapes.

		Band-enhanced	Band-suppressed	Hybrid	All-pass
35 dB SPL	Decreasing↓	12	**10**	4	3
	Increasing↑	**16**	19	3	8
	No-threshold –	15	22	2	6
45 dB SPL	Decreasing↓	13	**19**	2	4
	Increasing↑	**18**	15	4	8
	No-threshold –	12	17	3	5
55 dB SPL	Decreasing↓	13	**19**	3	4
	Increasing↑	**17**	12	3	5
	No-threshold –	13	20	3	8
65 dB SPL	Decreasing↓	9	**27**	4	5
	Increasing↑	**22**	10	3	5
	No-threshold –	10	14	2	7
75 dB SPL	Decreasing↓	9	**27**	4	5
	Increasing↑	**22**	10	3	5
	No-threshold –	10	14	2	7

Note: Numbers of units with the expected rate-change direction for BE or BS MTF types are in bold font. One BP unit was not tested at 75 dB SPL.

According to the envelope-based hypothesis, changes in rate upon addition of a tone to GN and LNN maskers were expected to be opposite. The number of units with different MTF shapes that had opposite or same rate-change directions for GN and LNN maskers were counted (Table 4). At lower noise levels, the number of units with opposite-direction rate changes was about the same or smaller than the number of units with the same-direction rate changes. At higher noise levels, more units had opposite-direction rate changes.

Table 2.3. The number of units with opposite or same rate change direction for GN and LNN maskers for each MTF type.

		Band-enhanced	Band-suppressed	Hybrid	All-pass	Total
35 dB SPL	opposite	**9**	**4**	**1**	**2**	**16**
	same	12	12	4	6	34
45 dB SPL	opposite	**10**	**7**	**3**	**3**	**26**
	same	17	17	2	4	40
55 dB SPL	opposite	**10**	**11**	**3**	**1**	**25**
	same	14	12	3	3	32
65 dB SPL	opposite	**9**	**16**	**3**	**2**	**22**
	same	12	13	1	4	29
75 dB SPL	opposite	**13**	**18**	**3**	**4**	**38**
	same	10	13	4	3	30

Note: Units without thresholds for both maskers not included. Units with expected responses are in bold.

The results in Table 4 were consistent with the correlation between maximum average rate differences in response to tone-in-GN and tone-in-LNN (Fig. 2.9). According to the envelope-based hypothesis, the correlation was expected to be negative. However, results showed that the rate differences for the two maskers were negatively correlated only at a noise level of 75 dB SPL. The significant positive correlation at low noise levels is consistent with energy-based encoding of tones in the presence of low-level noise. As noise level increases, the correlation coefficients in Fig. 2.9 transition from significant positive, to insignificant, to significant negative correlations, consistent with a transition from energy to envelope encoding mechanisms from low to high noise levels. The discrepancy between results based on separate evaluations of responses to GN and LNN

(Figs. 2.7 and 2.8) or direct comparisons of these responses (Fig. 2.9) will be discussed later.

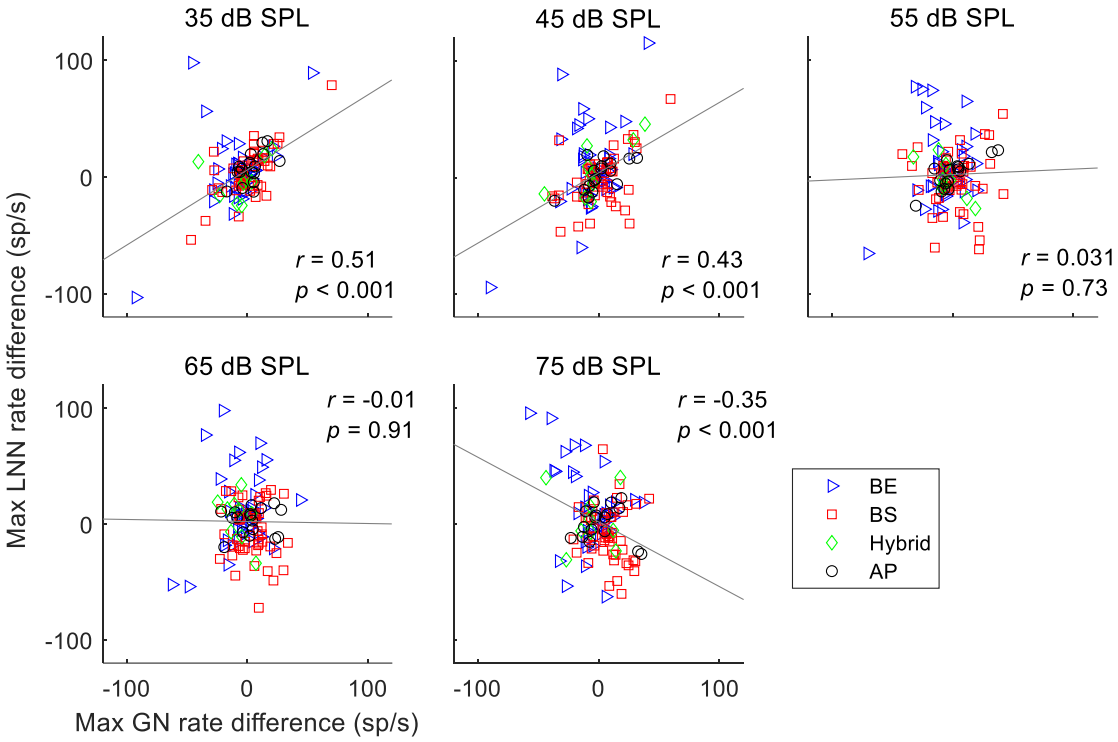

Figure 2.9: Correlation between maximum differences in average rate between TIN responses across all SNRs and the noise-alone responses for the two masker types. MTF types are indicated by different symbols (see legend). Correlation coefficients and p values are shown in each plot. The maximum rate differences between TIN responses of the two noise maskers were significantly positively correlated at noise levels of 35 to 45 dB SPL, not significantly correlated at 55 to 65 dB SPL, and significantly negatively correlated at 75 dB SPL.

Figure 2.10 shows GN (left) and LNN (right) TIN rate-based thresholds for all neurons, including increasing (red upward triangles) and decreasing (blue downward triangles) rate changes; gray circles at the top of the plot indicate neurons with no threshold. For the GN masker, 150 out of 163 units had measurable thresholds at one or

more noise levels; only 42 neurons had measurable thresholds at all five noise levels. For the LNN masker, 114 out of 120 units had threshold at one or more noise levels, and 36 neurons had thresholds at all five noise levels. There was not a clear trend for increasing/decreasing rate versus tone frequency (i.e. CF), except that at low frequencies and low levels, most of the neurons had increasing rate at threshold. As mentioned above, this trend could be due to a lack of saturation in AN responses, which would occur more often at lower frequencies due to higher audiometric thresholds. Human detection thresholds from several studies are plotted for comparison to neural thresholds. Note that sound levels and bandwidths used in the behavioral studies were slightly different from those in the current study. Listeners were tested with masker bandwidths of: 100, 250, 500 and 1000 Hz (compared with 116, 232, 463 and 924 Hz used here) for tone frequencies of 0.5, 1, 2 and 4 kHz respectively from van de Par & Kohlrausch (1999) at a noise level of 70 dB SPL; 100 Hz from Kohlrausch et al. (1997) at 60 dB SPL; 78, 240, 456 and 888 Hz (compared with 116, 463, 926 and 1852 Hz used here) for tone frequencies of 0.5, 2, 4 and 8 kHz, respectively, from Goupell (2012). Rabbits were tested with 200-Hz bandwidth maskers in Zheng et al. (2002). Human thresholds from Goupell (2012) are slightly lower than those of van de Par & Kohlrausch (1999), possibly due to narrower masker bandwidths. In general, human thresholds can be better explained by rate-based thresholds at high CFs than at low CFs, for both GN and LNN maskers, despite the differences in parameters. At high CFs, rate-based thresholds were lower than human thresholds.

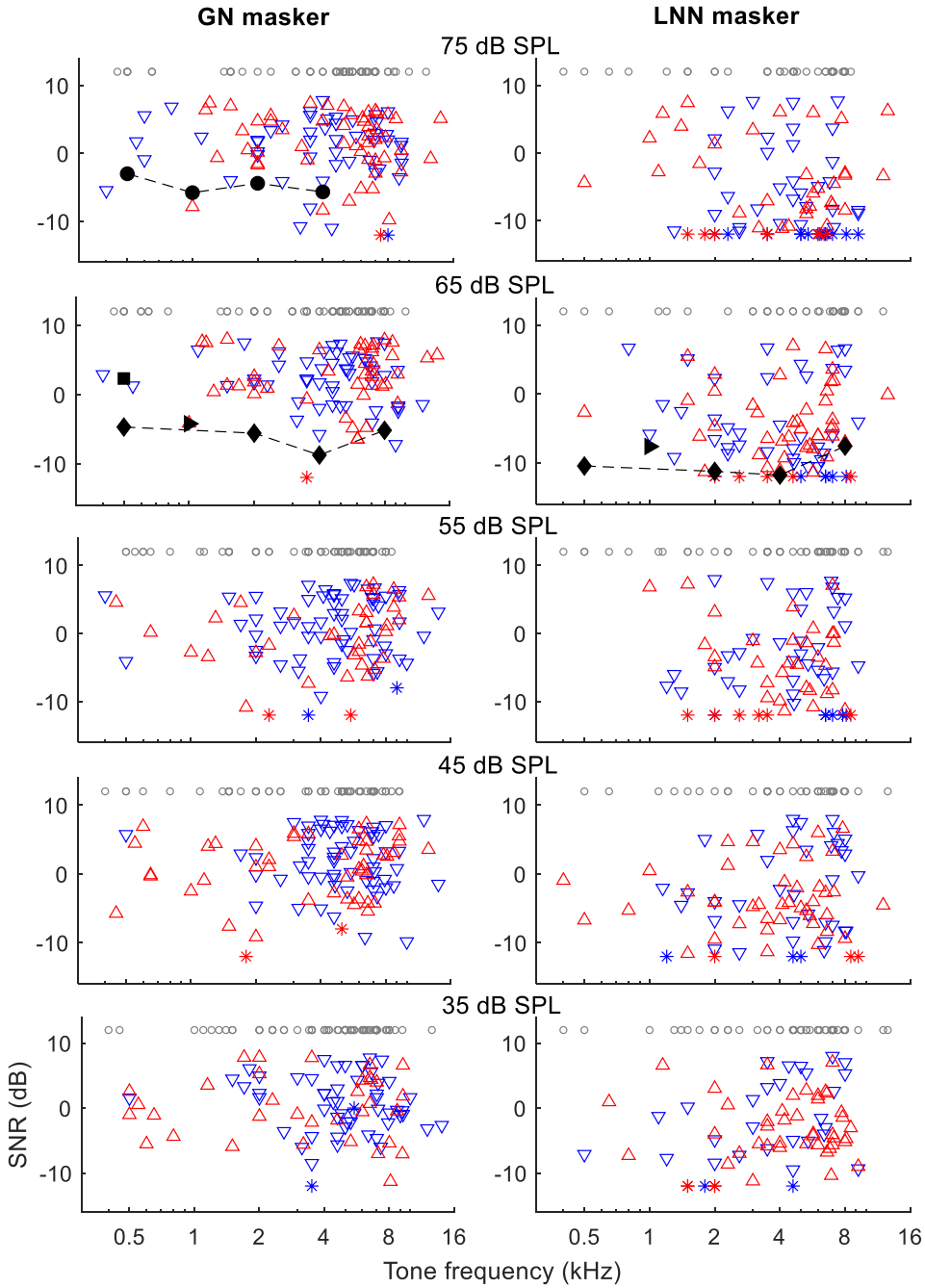
GN masker
LNN masker
75 dB SPL
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
SNR (dB)
10
0
-10
0.5
1
2
4
8
16
Tone frequency (kHz)
Neural Threshold
↑ rate at threshold
↓ rate at threshold
Threshold < lowest SNR
No threshold
Human behavioral data
Leong et al., 2020
van de Par & Kohlrausch, 1999
Kohlrausch et al., 1997
Goupell, 2012
Rabbit behavioral data
Zheng et al., 2002

Figure 2.10: Rate-based thresholds for all units at each noise level with GN (left) and LNN (right) maskers. Thresholds of the most sensitive neurons across frequency were similar to human behavioral data and rabbit behavioral data at 500 Hz (reference see legends).

Figure 2.11 shows threshold differences between rate in response to the two maskers (LNN – GN) for individual neurons at each noise level. Only units with measurable thresholds for both GN and LNN are shown here. If thresholds were lower than the lowest SNR tested, the lowest SNR was used for a conservative estimation (triangles), but neurons were excluded if the thresholds for both GN and LNN were lower than the lowest SNR. Two neurons did not have measurable thresholds for either masker type at any noise level, and 20 neurons had thresholds for both maskers at all noise levels. Negative values indicate that the threshold was lower for the LNN than GN masker, as expected based on human behavior (Kohlrausch et al., 1997). Recall that adding a tone to a GN masker reduces amplitudes of both stimulus envelope and neural fluctuations, whereas adding a near-threshold tone to an LNN masker increases amplitudes of both stimulus envelope and neural fluctuations. Therefore, opposite rate-change directions versus SNR were expected for GN and LNN maskers. Units with the expected opposite rate changes are shown in black (i.e. A, black stems indicate units with decreasing rate at threshold for GN maskers and increasing rate for LNN maskers; opposite in B). At lower noise levels, among neurons with expected opposite rate changes for GN and LNN maskers, more had decreasing rate for GN masker. The number of units with opposite rate-change directions for GN and LNN maskers increased with increasing noise levels. A Student t test showed that threshold differences for GN and LNN maskers were significantly greater for neurons with opposite-direction rate changes than for neurons with same-direction rate changes, $t(289) = 15.31, p < 0.0001$.

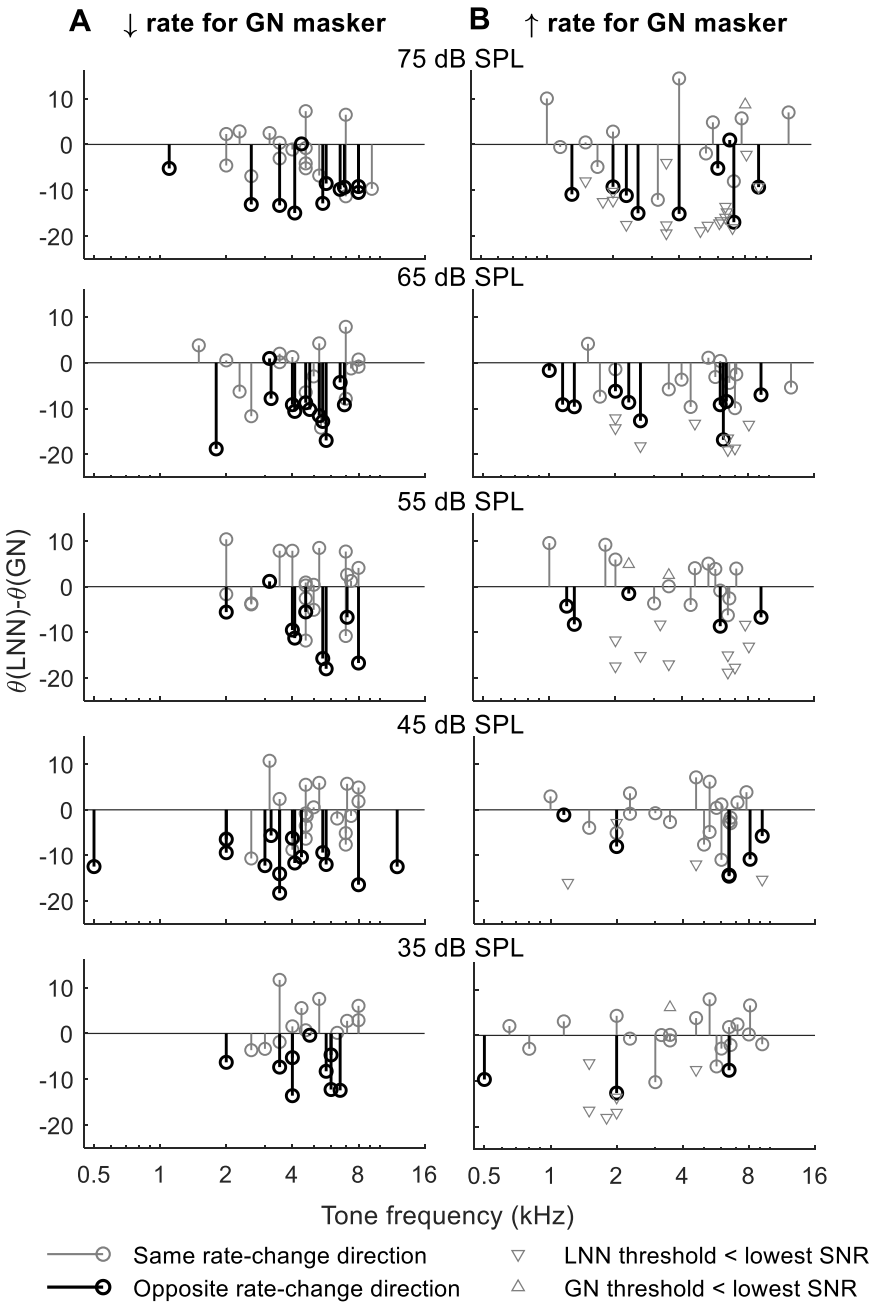

Figure 2.11: Threshold differences between GN and LNN maskers for units with increased (A) or decreased (B) rate at threshold for GN. Triangles indicate that the threshold for the GN or for LNN masker was lower than the lowest tested SNR, thus caused a conservative estimate. Neurons without a threshold or with thresholds for both GN and LNN maskers lower than the lowest tested SNR are excluded. Neurons that had opposite rate-change directions for GN and LNN maskers generally had larger threshold differences.

2.5 Temporal analysis and Results

2.5.1 Analysis

Correlation Index

Correlation index (CI) describes how reliable or repeatable spike times are in response to multiple repetitions of the same stimulus, similar to the concept of phase-locking in response to a periodic signal. Large CI values indicate more reliable spike times. The method to calculate correlation index was the same as in Joris et al. (2006), and is briefly described here. Each stimulus waveform was presented 30 times in the current study. For a spike train in response to one stimulus presentation, the total number of spikes from all other spike trains that happen within the coincidence window of each spike in the first spike train was counted. This procedure was repeated for each spike train and the number of spikes were summed (N_c). CI is a dimensionless number calculated by normalizing N_c using the following equation:

$$CI = \frac{N_c}{M(M-1)r^2\omega D},$$

where r is the average response rate (spikes/sec), ω is the length of the coincidence window (0.2 ms, similar results between 0.2 – 1 ms), D is the stimulus duration, and M is the number of stimulus repetitions. The confidence interval was calculated by the 95th percentile of the N_c distribution calculated with 1000 versions of scrambled responses (yielding a significance level of α=0.05). A CI was determined significant if it exceeded the confidence interval. CI was only calculated for conditions with average rates greater than 15 spikes/s to avoid large CI values due to normalization by small rates. Adjusted CI (CI_A), confidence interval subtracted from CI, was used for statistical tests.

Temporal Threshold

CI provided a quantitative description for reliability of spike times; however, there was no criterion to estimate a threshold based on CI. Therefore, neural temporal thresholds were estimated by evaluating changes in the PSTH upon addition of a tone. The process is demonstrated in Fig. 2.12. At each noise level, two PSTHs, with 1-ms binwidths, were calculated from responses to the 15 odd- and 15 even-numbered repetitions of the 30 total repetitions of the TIN stimulus at each SNR. Correlations were calculated between the two PSTHs for the noise-alone condition (r_0), and between each combination of noise PSTHs and tone-plus-noise PSTHs (r_1, r_2, r_3, r_4) at each tone level. The correlation between noise PSTH and tone-plus-noise PSTH was expected to decrease with increasing tone level, because adding higher-level tones would result in larger differences in the stimulus waveforms as well as in the peripheral responses. The lowest SNR above which r_1 , r_2 , r_3 and r_4 were all significantly lower than r_0 was considered the temporal threshold. The significance level was adjusted with Bonferroni correction for multiple comparisons (Goeman & Solari, 2014). Note that the temporal threshold estimated here was supra-threshold, and the true threshold would lie between this tone level and its lower neighbor.

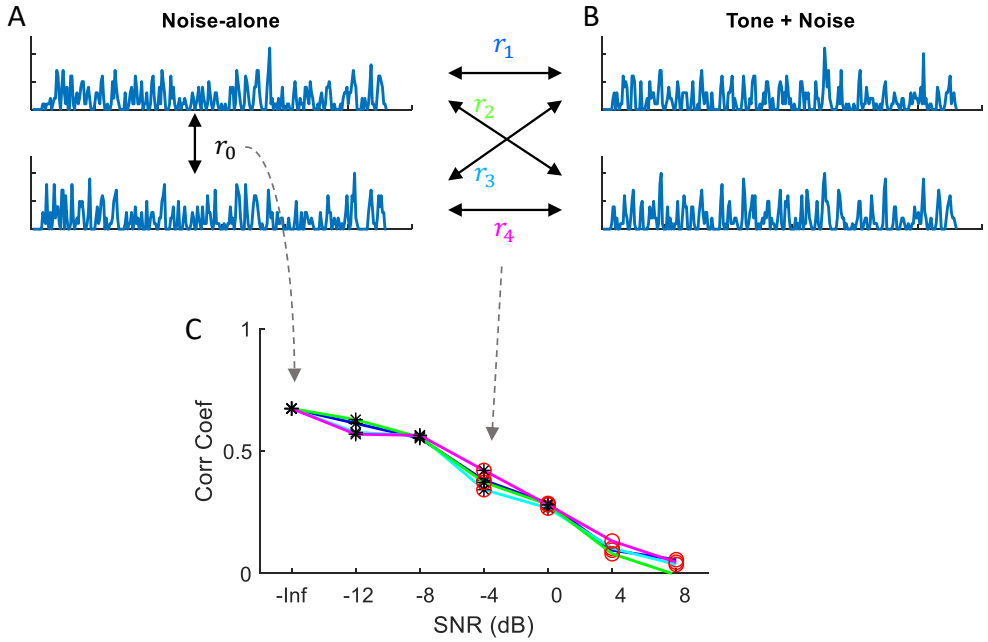

Figure 2.12: Illustration of temporal threshold calculation. A) PSTHs for odd- and even-numbered repetitions of noise-alone stimulus. B) PSTHs for odd- and even-numbered repetitions of noise with tone at SNR = -4 dB. Correlation coefficients between the noise PSTHs (r_0) and between PSTHs of noise and tone-plus-noise trials (r_1, r_2, r_3, r_4) were calculated to find temporal threshold. C) Correlation coefficients of r_0 and r_1, r_2, r_3, r_4 as a function of SNR. * indicates significance of the correlation coefficient; ○ indicates that all four correlations (r1-r4) were significantly lower than r_0. Estimated temporal threshold was the lowest SNR above which r_1, r_2, r_3 and r_4 were all significantly lower from r_0. In this example, the threshold was -4 dB.

2.5.2 Results

Temporal analysis was carried out for 113 units (out of 163 units) for which responses were recorded to many repetitions of reproducible GN maskers. Among the 113 units, responses of 74 units were also recorded for reproducible LNN maskers.

Reliability of spike times

Neurons in the IC respond reliably when the same stimulus is presented, as examined with SAM stimuli (e.g. Krishna & Semple, 2000; Joris et al., 2004). When spike times were reliable, PSTHs had many large peaks (bottom panel, Fig. 2.13A); otherwise, PSTHs had small fluctuations (bottom panel, Fig. 2.13B). Note that the reliability was not hypothesized to be related to the neuron's MTF shape. Figure 2.13 shows an example neuron for which the "peakiness" of PSTHs for GN and LNN masker varied with SNR. CI_A described the "peakiness" of the PSTH, as it counts the numbers of spikes that fall in the coincidence window. CI_A is plotted as a function of SNR together with average rate for GN (C) and LNN (D) (Fig. 2.13). In response to GN maskers, the CI_A tended to decrease with increasing SNR. For LNN maskers, the CI_A tended to first increase versus SNR, and then decrease for high SNRs. In this example neuron, reliability of spike times had similar trends as rate versus SNR.

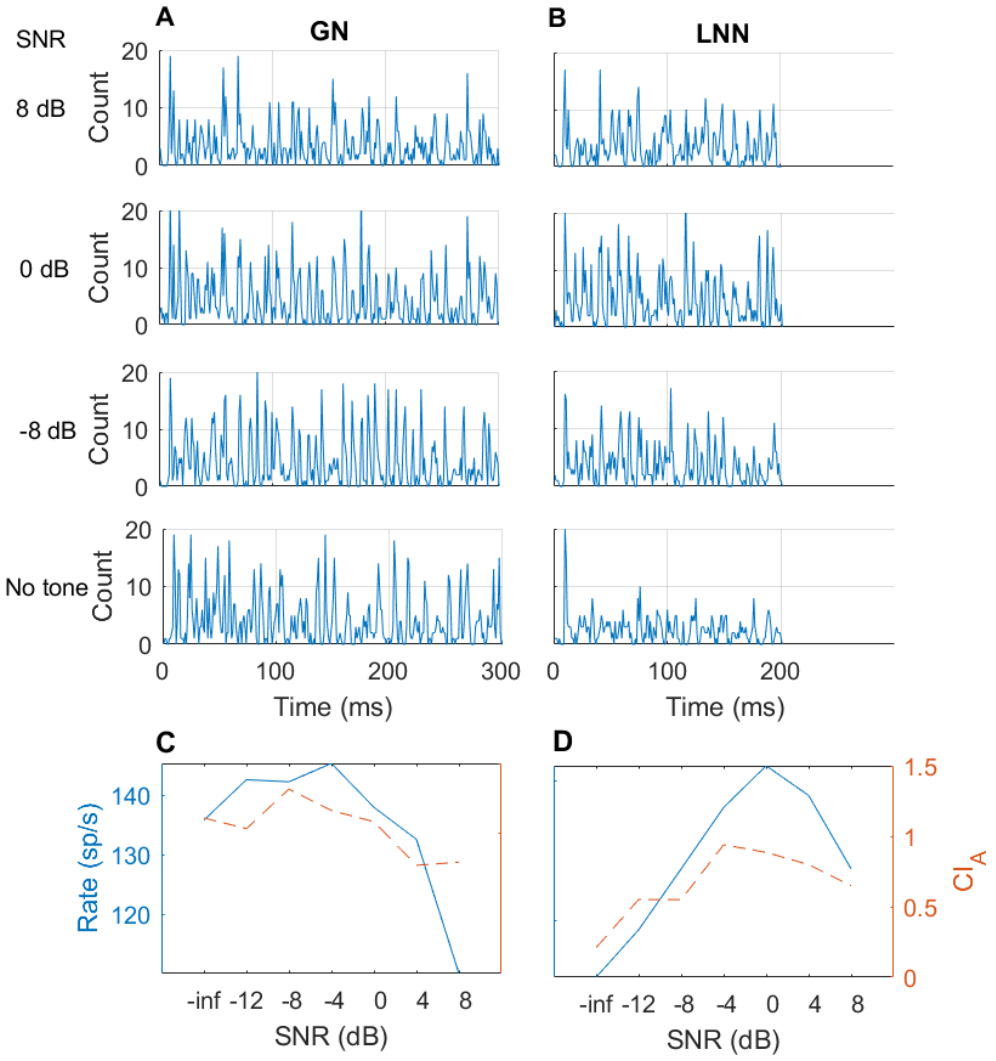

Figure 2.13: Example PSTHs used for temporal analyses (A, B), and changes in average rate and CI_A vs. SNR (C, D) for a noise level of 55 dB SPL. A and B) PSTHs at selected SNRs for GN (A) and LNN (B) maskers. Note that the duration of GN and LNN maskers were different, but PSTHs were plotted on the same scale to allow visual comparison between responses to GN and LNN maskers. C and D) Average rate (solid) and CI_A (dashed) as a function of tone level for GN (C) and LNN (D) maskers. Temporal reliability decreased for the GN masker as SNR increased, but reliability first increased and then decreased for LNN masker vs. SNR.

Figure 2.14 shows another IC neuron's PSTHs for selected SNRs, and CI_A as well as average rate as a function of SNR. The PSTHs are more "peaky" at low SNR than at high SNR, which can be best seen in Fig. 2.14C and D: the CI_A of this neuron decreased with increasing tone level for both GN and LNN, opposite to the change in rate. Recall that CI_A was normalized by rate, so the effect of rate was removed. If there was a certain feature in the stimulus that drove the IC neuron, the spike time of the neuron in Fig. 2.13 was more informative, as reduced response rate resulted in reduced temporal reliability (both CI_A and rate decreased with increasing SNR). The spike times of the neuron in Fig. 2.14 was less informative, as more spikes did not improve CI_A. On the other hand, CI_As for both neurons were greater than zero, higher than values for randomized spike times, so the spike times of both neurons were generally informative (i.e. reliable).

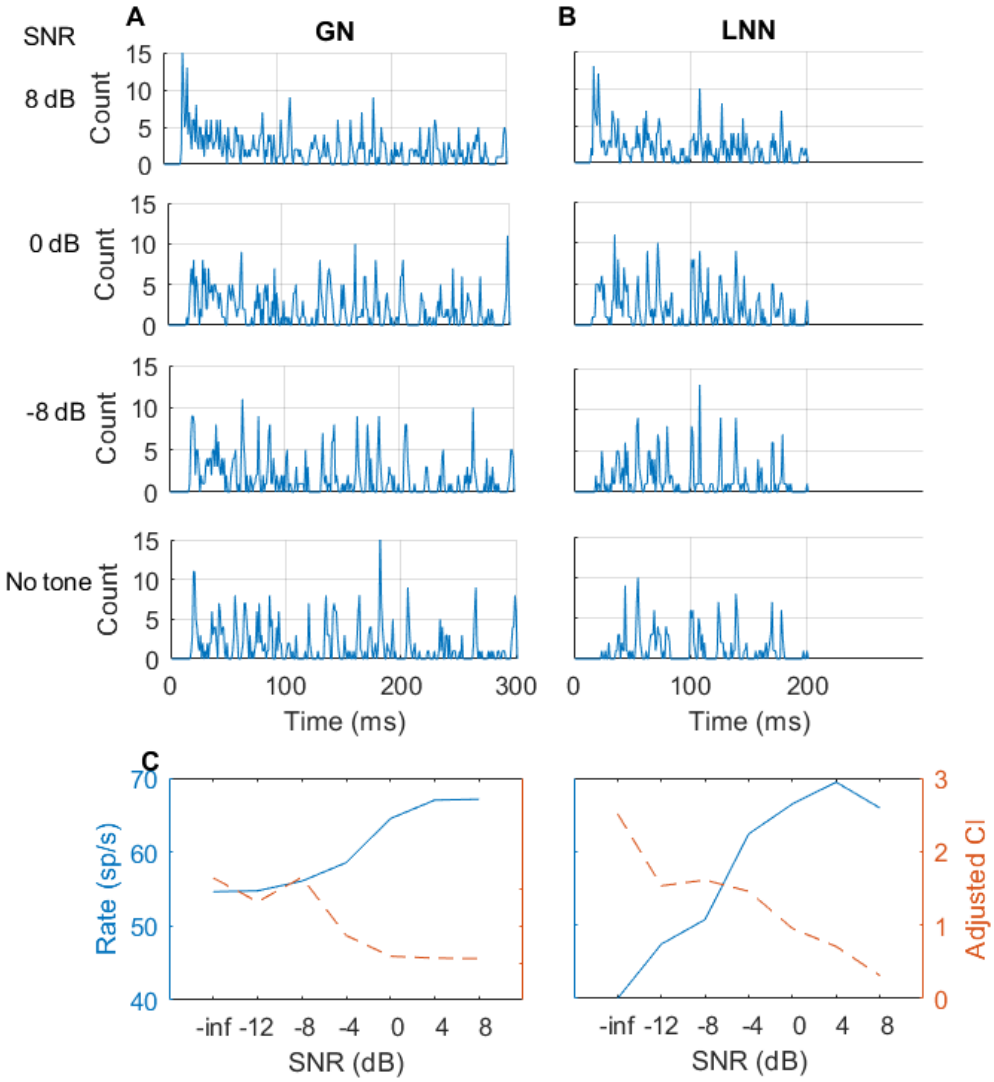

Figure 2.14: Example PSTHs (A, B) used for temporal analysis, and changes in average rate and CI_A vs. SNR (C, D) for another neuron in response to TIN stimuli with maskers at 35 dB SPL. Format is the same as in Fig. 2.13. Temporal reliability decreased for both GN and LNN maskers with increasing SNR.

To better evaluate the trends across the population in the CI_A versus SNR, a two-way ANOVA was conducted for both GN and LNN maskers. Fig. 2.15 shows the mean and standard deviation of CI_A as a function of SNR for GN (left) and LNN (right) maskers. For the GN masker, both main effects of noise level and SNR were significant:

$F(4,3077) = 2.47, p = 0.043$; $F(6,3077) = 16.65, p < 0.0001$. The interaction was not significant, $F(24, 3077) = 0.56, p = 0.96$. A post hoc test showed that CI_A for the noise level of 35 dB SPL was significantly different (higher) from that at 55 dB SPL; CI_A of 4- and 8-dB SNRs were significantly different from no-tone and SNRs of -12 to -4 dB, and CI_A for 0-dB SNR was significantly different from 8-dB SNR.

For the LNN masker, both main effects of noise level and SNR were significant: $F(4,1978) = 6.33, p < 0.0001; F(6,1978) = 11.23, p < 0.0001$. The interaction was not significant, $F(24, 1978) = 0.51, p = 0.98$. A post hoc test showed that CI_A for noise levels of 35 dB SPL was significantly different (higher) than for 55 to 75 dB SPL masker levels; CI_A of no-tone, SNRs of -12, -8, and 0 dB were significantly different (higher) than 8-dB SNR, while SNR of -4 dB was significantly different (higher) than both 4- and 8-dB SNRs. The statistics indicate that spike times became less reliable with increasing tone level for the GN masker, while they became slightly more reliable for the LNN masker at low SNRs.

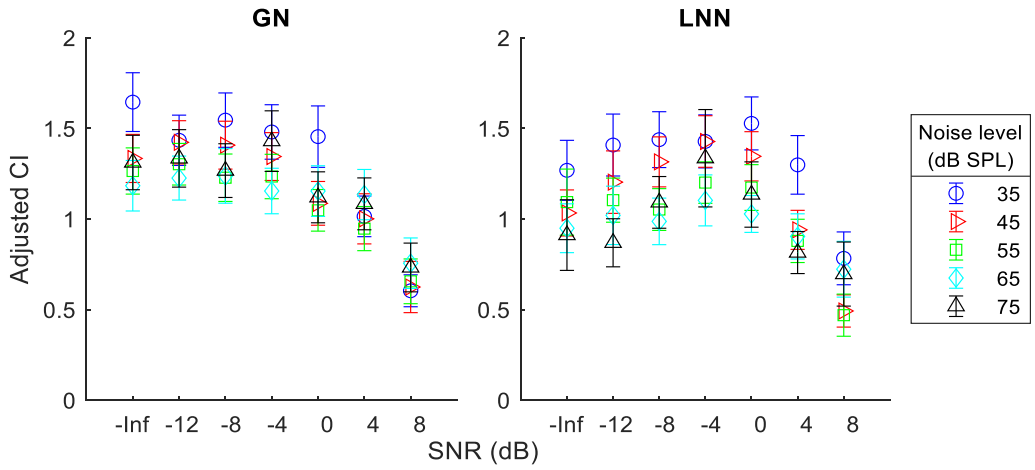

Figure 2.15: CI_A as a function of SNR at different noise levels for GN (left) and LNN (right) maskers. Noise levels are indicated in the legend. Errorbars show standard error.

Temporal thresholds

Temporal threshold was difficult to calculate for responses to noise stimuli. However, as shown in Figs. 2.13 and 2.14, temporal information was embedded in the shape of PSTHs. Here, temporal thresholds were calculated for each available unit based on changes in correlation coefficients between noise-alone PSTHs and noise-plus-tone PSTHs. Thresholds are plotted in Fig. 2.16 for both GN (left) and LNN (right) at each noise level, together with human and rabbit detection thresholds (as in Fig. 2.10). Tested SNRs had 4 dB steps, which limited the resolution of temporal threshold estimates. Also, due to the current method of estimating temporal thresholds, true thresholds for each unit should be the same or lower than plotted here. Since temporal thresholds were calculated based on changes in PSTHs, the relationship between CI_A and temporal thresholds was examined. For GN maskers, there was significant correlation between CI_As for the noise-alone condition and temporal thresholds, $r = -0.36, p < 0.0001$; that is, when the spike times were more reliable, it was easier to detect a change in the temporal pattern, and thus the temporal threshold was lower. However, there was no such trend for LNN maskers ($r = -0.015, p = 0.87$).

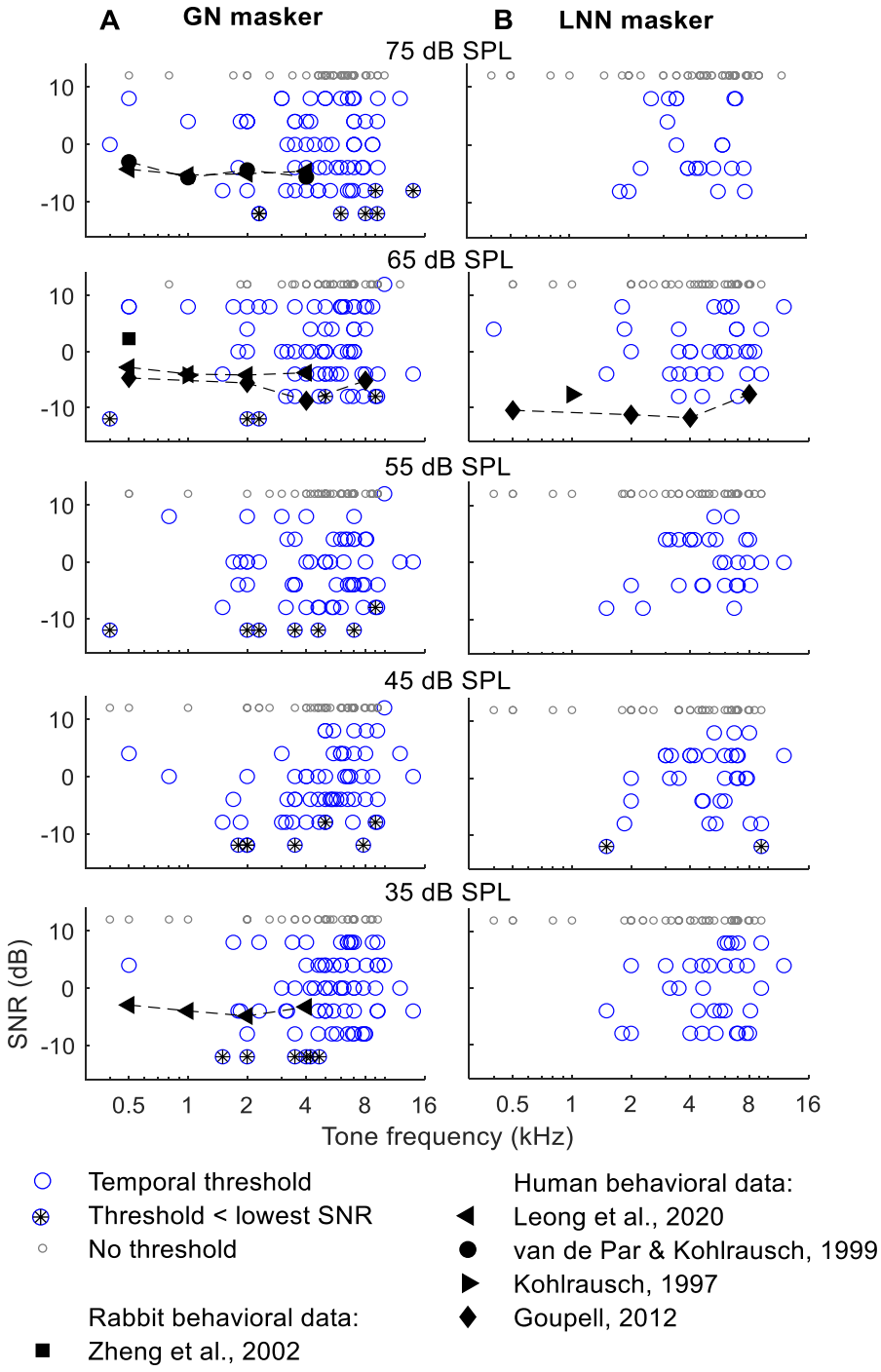

Figure 2.16: Temporal thresholds at each noise level for GN (left) and LNN (right) maskers. Human and rabbit behavioral data are the same as in Fig. 2.10.

Temporal thresholds have been shown to be lower than rate thresholds, and can better predict human performance than rate thresholds in a SAM detection task (Carney et al., 2014; Henry et al., 2016). To compare the rate and temporal thresholds in the current study, temporal thresholds were plotted against rate thresholds in Fig. 2.17 for both GN (left) and LNN (right) maskers at each noise level. Only neurons with measurable thresholds for both rate- and timing-based calculations were included here. In each panel, circles above the diagonal line indicate higher temporal than rate thresholds, and circles below the diagonal line indicate lower temporal than rate thresholds. The number of higher or lower temporal thresholds is shown in the top right corner of each panel. For the GN masker, at most noise levels there were more neurons with lower temporal than rate thresholds, consistent with previous studies. However, for the LNN masker, the majority of neurons had lower rate than temporal thresholds at all noise levels.

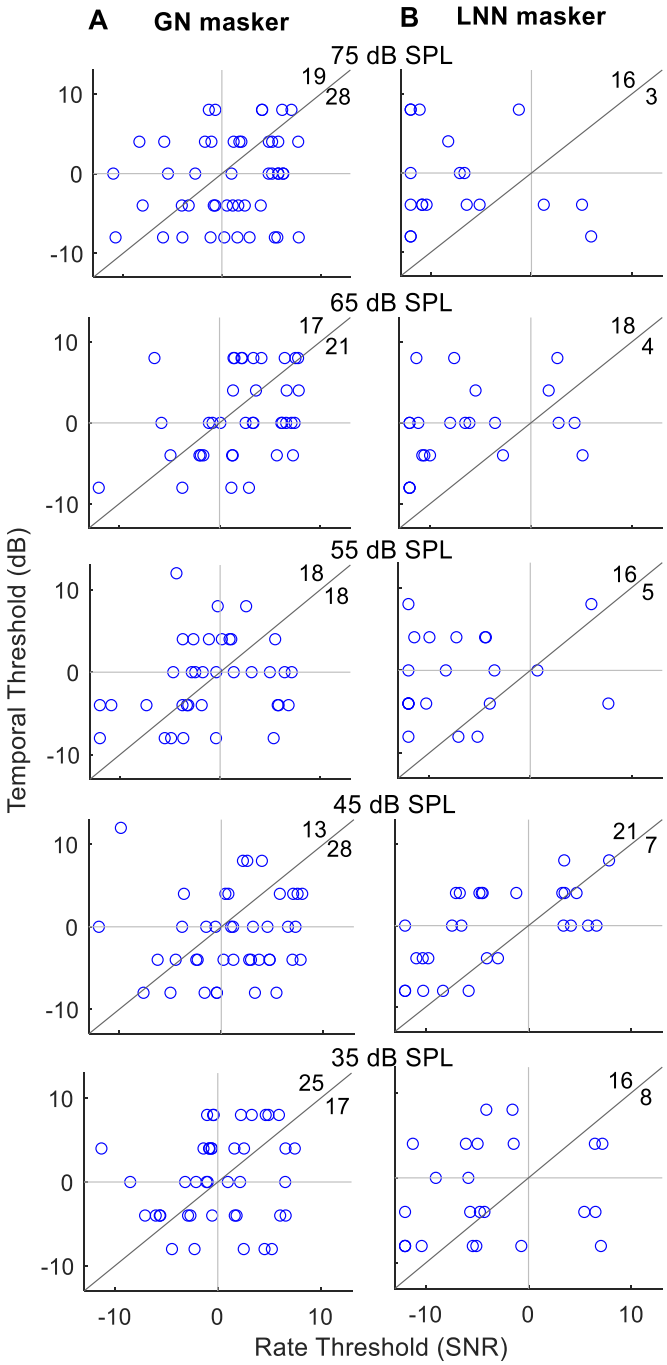

Figure 2.17: Temporal vs. rate thresholds for GN masker (A) and LNN masker (B) for neurons with measurable thresholds for both. Diagonal lines indicate identical temporal and rate thresholds. Neurons above the diagonal line had lower rate thresholds, whereas neurons below the diagonal line had lower temporal thresholds, and the numbers were shown above and below the diagonal line, respectively.

2.6 Discussion

Responses of single neurons in the IC to TIN stimuli were examined over a wide range of noise levels and SNRs for both GN and LNN maskers. Most IC neurons are tuned to amplitude modulations, as described by MTFs. Adding a tone flattens the envelope of a narrowband GN waveform (Richards, 1992), yet introduces fluctuations in the envelope of an LNN waveform for SNRs near behavioral detection thresholds (Kohlrausch et al., 1997). IC neurons were hypothesized to be sensitive to these changes in the envelope, as conveyed by slow fluctuations in the instantaneous rates of their neural inputs, upon addition of a tone to the two masker types. The nature and direction of the changes in IC responses with SNR was hypothesized to depend on the neuron's MTF type. As predicted, for neurons that had rate-based thresholds for GN maskers, most neurons with BE MTFs had decreasing rate and most neurons with BS MTFs had increasing rate at neural detection thresholds. Also as predicted, for LNN maskers, most neurons with BE MTFs had increasing rate, and most neurons with BS MTFs had decreasing rate at detection thresholds. Additionally, the maximum average rate changes across SNR of tone-in-GN responses were negatively correlated with maximum average rate differences in the MTF (between peaks or troughs in the MTF and unmodulated response rates). In contrast, maximum rate changes of tone-in-LNN responses were positively correlated to maximum changes in the MTF. These results support the hypothesis IC responses to narrowband TIN stimuli are largely explained by changes in the stimulus envelope, which are represented, and possibly enhanced, by slow fluctuations of AN responses.

Temporal analysis showed that most neurons had significant CIs. In response to GN maskers, adjusted CI decreased with increasing SNR, whereas in response to LNN

maskers, adjusted CI first increased then decreased with increasing SNR. Reliability of spike times was negatively correlated with temporal thresholds for GN conditions but had no correlation to threshold for LNN conditions. Among neurons with both measurable rate and temporal thresholds, most had lower temporal thresholds for GN conditions, but had lower rate thresholds for LNN conditions.

Neural encoding of envelope cues

Researchers have mainly focused on energy-based cues in TIN detection tasks (e.g., Fletcher, 1940). However, the energy model cannot explain human behavior for roving-level conditions (Kidd et al., 1989). Previous peripheral physiological studies have also shown that HSR AN fibers, the majority afferent group, had small changes in rate with increasing tone level at medium to high noise levels (Rose & Capranica, 1985), indicating "non-energy" coding for this group of AN fibers. At the level of the cochlear nucleus, neurons that project to more central nuclei have also been shown to be more sensitive to envelope cues than to energy (Gai & Carney, 2006). Many IC neurons also have non-monotonic rate versus sound level, in both the current study (Fig. 2.6) and in previous studies of responses to TIN (Jiang et al., 1997a, 1997b; Ramanchandran et al., 2000), a result that complicates energy-based rate coding. On the contrary, responses at the levels of AN, cochlear nucleus, and IC could support envelope-based neural encoding of the stimulus. Several studies have proposed that the envelope provides another detection cue for human listeners (Dau et al., 1996; Davidson et al., 2009a; Kohlrausch et al., 1997; Mao et al., 2013; Richards, 1992), and envelope information has been shown to be critical for speech recognition (e.g., Drullman, 1995). However, these studies have mainly focused on the stimulus envelope statistics and signal processing approaches. There has been a

lack of physiological studies focusing on changes in neural responses based on envelopes. Our study fills this gap, and the results show that most IC neurons with measurable TIN detection thresholds support the envelope-based hypothesis. In addition, this hypothesis may explain why fewer BS neurons responded as predicted at low sound levels. When overall sound level is low, HSR AN fiber average rates are not saturated; therefore, when adding a tone to narrowband GN, AN rates will be affected by energy (Carney, 2018) and having increasing rate (e.g., Neuron 8). As shown in Fig. 2.11 (bottom-right panel), at low masker levels there are many neurons that responded to both GN and LNN maskers with increasing rate for increasing tone level. Some neurons had increasing rate thresholds with increasing noise level, which can also be explained by the envelope-based hypothesis, as follows: At low noise levels, AN fibers did not have saturated average rate, and thus neural responses would fluctuate more than at high noise levels. Adding a tone would result in flattening of AN responses, changing the rate of IC cells more at low noise levels than at high noise levels.

Responses across the population for GN and LNN maskers: separate or comparative analysis of the two maskers.

When responses to GN and LNN maskers were analyzed separately for the population of IC neurons studied, the responses to TIN stimuli were consistent with sensitivity to amplitude modulation (Figs. 2.7 and 2.8), supporting the envelope-based hypothesis. According to this hypothesis, neural responses to tone-in-GN and tone-in-LNN should have opposite rate-change directions. Therefore, maximum rate differences for the GN and LNN condition were expected to be negatively correlated; however, significant negative correlation was only found at 75 dB SPL (Fig. 2.9). The correlation between

maximum rate differences of GN and LNN conditions suggested a transition from energy-based cues to envelope-based cues from low to high noise levels. The two conclusions seem to contradict each other but may be reconciled as follows: first, correlation coefficients for the GN or LNN condition alone increased with increasing noise level, suggesting stronger encoding of envelope-based cues at higher levels. Second, TIN responses of many BS neurons were not consistent with their MTF shape (Fig. 2.8), and these could largely bias the correlation between GN and LNN responses (Fig. 2.9, squares in the top left plot).

Complexity in the rate responses of IC neurons

Even though most BE and BS neurons with measurable thresholds supported the envelope-based hypothesis, the hypothesis could not explain responses of all neurons at all noise levels, indicating complexity in the IC responses. One complexity lies within the variation of MTFs, even within MTF types. For example, the best or worst modulation frequency, at which BE or BS neurons had the largest difference in rate as compared to an unmodulated sound, can vary over about 100 Hz. Our unpublished data also show that neurons that had similar MTF shapes for SAM with a tone carrier at CF may have different MTF shapes when the tone carrier was far off CF. This change in MTF shape with carrier frequency suggests differences in the neural inputs among neurons (e.g. inputs from lower auditory pathways or from other IC neurons). Differences of neural inputs can also be seen in Fig. 2.6 – the noise RLF did not always have the same trend as the tone RLF.

Another factor that may affect neural responses is the bandwidth of the noise waveform. When a frequency component in the stimulus is far from CF, the AN fiber at CF will have a higher threshold in response to that frequency; therefore, a higher sound level

is required to yield rate saturation similar to that for a CF frequency. The 1/3-oct bandwidth used in this study is not wide, but responses to different frequency components may vary slightly, especially for high-frequency stimuli that had wider bandwidths than low-frequency stimuli. A mismatch between tone frequency and CF may also affect neural responses to TIN stimuli. However, this mismatch could not be easily addressed, as IC neurons typically do not have as sharp a tip in the tuning curve as observed at lower levels of the auditory pathway.

Sound level may affect the categorization of MTF groups. There was a trend for more BS neurons to respond as expected with increasing noise level. As explained above, this trend could be due to unsaturated IHC and AN fibers. This trend could also be due to changes in MTF shape at lower sound levels (Krishna & Semple, 2000). Based on our limited data, changes in BE MTFs across sound level were mostly reflected in the amplitude of enhancement. On the contrary, BS MTF shape requires high rates at low modulation frequencies, but at low sound levels, the rate in response to low modulation frequencies may decrease, and the MTF may become hybrid or even weakly band-enhanced. Note that there was an increasing number of BS units that had measurable thresholds with increasing noise level, but the number was similar for BE units across noise level (Fig. 2.8).

Temporal reliability

Temporal information allows decoding spike times across a large population of neurons by synchrony or correlation. Encoding of auditory stimuli requires high resolution in time, thus temporal information may be more straightforward than rate information, especially for time-varying auditory stimuli – IC neurons can phase lock to the envelope

of sinusoidal amplitude-modulated sounds (Krishna & Semple, 2000), meaning that timing of spikes is faithful to the phase across cycles of amplitude modulation. IC neurons also have reliable spike times in response to non-periodic signals, such as noise (Carney and Yin, 1989; Keller & Takahashi, 2000). Thus, one may expect that envelope slopes or peaks drive the neuron, whether the stimulus is periodic or not. If so, in response to TIN stimuli, spike times in response to GN-alone stimuli would be most reliable, and reliability would decrease with increasing SNR, as the envelope progressively flattens. On the other hand, spike times in response to LNN-alone stimuli would be most unreliable, and reliability would first increase then decrease with increasing tone level, as the envelope becomes more fluctuating and then flatter. This expectation was observed in the CI_As across SNR, which had small but significant trends for both GN and LNN. However, temporal responses to LNN were shown to be more complicated. For LNN-alone, due to the flat envelope, IC PSTHs were expected to be flat or less fluctuating; instead, IC PSTHs were peaky (Fig. 2.13), and spike times were reliable (similar CI_A as for GN). One reason could be that peripheral filtering may introduce neural fluctuations in the AN responses, especially when the tone frequency and the neuron's CF were mismatched; such neural fluctuations would be expected to entrain the IC responses.

Temporal thresholds

Although CI values provide a quantitative description for reliability of spike timing, there is no criterion to determine a significant change in CI on a single-neuron basis. Additionally, CI can describe the reliability of spike times, but not changes in the overall pattern in the PSTH caused by addition of a tone. Here, the change in the correlation of PSTHs as SNR was increased was used to estimate temporal thresholds. Note that the

correlation between the PSTHs within TIN conditions was used to compute the reliability of spike timing, whereas the correlations between PSTHs for noise-alone and tone-plus-noise conditions reflected the general change in neural responses upon addition of a tone. For neurons with measurable thresholds based on both rate and temporal measures, the temporal thresholds tended to be lower than rate thresholds for GN maskers (Fig. 2.15), consistent with previous studies. However, for LNN maskers, temporal thresholds tended to be higher than rate thresholds (Fig. 2.15). Higher temporal thresholds indicated that temporal response patterns did not change significantly upon addition of a tone until the SNR was relatively high. Together with the result that CI_A slightly increased at low tone levels, temporal patterns in response to LNN, with or without low-level added tones, were driven by more than just envelopes.

Neural vs. behavioral threshold

Based on limited rabbit behavioral results (Zheng et al., 2002), rabbit rate-based neural thresholds matched rabbit behavioral data at 500 Hz (Fig. 2.10), and temporal thresholds were lower than behavioral data (Fig. 2.17). A few human datasets from different studies (Kohlrausch et al., 1997; van de Par & Kohlrausch, 1999; Goupell, 2012) were also compared to neural thresholds; human thresholds from different studies generally agree with each other despite the slight difference between stimuli tested. The neural thresholds of the most sensitive neurons across frequency can explain human detection thresholds, and may be slightly more sensitive than human thresholds at high frequencies. Qualitatively, the lowest neural thresholds across frequencies were similar for different noise levels. Note that human detection thresholds with LNN masker is about 5 dB lower than with GN masker (Kohlrausch et al., 1997), and this difference was

reflected in the rate-based neural thresholds of the most sensitive neurons across the population in the current study.

Temporal thresholds were only found lower than rate thresholds in most neurons for GN masker, whereas most neurons had lower rate thresholds for LNN masker. These results indicate that temporal responses did not always provide extra information to facilitate detection tasks; changes in IC average rates may be sufficient to explain behavioral thresholds.

Conclusion and future work

The role of envelope has not been well studied with TIN tasks in physiological studies, and neural fluctuations have not previously been studied. Our results showed that most neurons in the IC support the envelope-based hypothesis, which suggested that both stimulus envelopes and neural fluctuations are worthy of more attention, not just in TIN tasks, but also in other psychophysical paradigms. Future studies may systematically examine the role of best/worst frequency and bandwidth of MTFs in explaining TIN responses. On the other hand, neural responses in the IC are complicated, which may arise from convergence of various afferent inputs (Winer, Schreiner, & SpringerLink, 2005), and sensitivity to the envelope is not enough to explain all responses to other stimuli. A recent study also showed that responses of IC neurons were affected by the neuron location within the IC, possibly due to the unique local circuits in the IC (Ono, Bishop, & Oliver, 2017). Our study showed that even for neurons within the same MTF category, there may be substantial differences in the responses. One can imagine high-dimensional objects – two objects that look the same from one angle may look different in another. An IC neuron can have binaural sensitivity, interaural time and level sensitivity, SAM

sensitivity, RM type, direction selectivity, etc., representing many neural dimensions. Therefore, neurons in the IC, and perhaps also neurons at higher levels, should be characterized in more than one way in the future. The amount of recording required will increase, but may bring more insights in the neural mechanisms of the auditory system.

Bibliography

Carney, L. H. (2018). Supra-Threshold Hearing and Fluctuation Profiles: Implications for Sensorineural and Hidden Hearing Loss. *J Assoc Res Otolaryngol, 19*(4), 331-352. doi:10.1007/s10162-018-0669-5

Carney, L. H., & Yin, T. C. (1989). Responses of low-frequency cells in the inferior colliculus to interaural time differences of clicks: excitatory and inhibitory components. *Journal of Neurophysiology*, *62*(1), 144-161.

Carney, L. H., Zilany, M. S., Huang, N. J., Abrams, K. S., & Idrobo, F. (2014). Suboptimal use of neural information in a mammalian auditory system. *Journal of Neuroscience, 34*(4), 1306-1313. doi:10.1523/JNEUROSCI.3031-13.2014

Chase, S. M., & Young, E. D. (2006). Spike-timing codes enhance the representation of multiple simultaneous sound-localization cues in the inferior colliculus. *Journal of Neuroscience, 26*(15), 3889-3898. doi:10.1523/JNEUROSCI.4986-05.2006

Dau, T., Puschel, D., & Kohlrausch, A. (1996). A quantitative model of the "effective" signal processing in the auditory system. I. Model structure. *J Acoust Soc Am, 99*(6), 3615-3622. doi:10.1121/1.414959

Davidson, S. A., Gilkey, R. H., Colburn, H. S., & Carney, L. H. (2009a). An evaluation of models for diotic and dichotic detection in reproducible noises. *Journal of the Acoustical Society of America, 126*(4), 1906-1925. doi:10.1121/1.3206583

Davidson, S. A., Gilkey, R. H., Colburn, H. S., & Carney, L. H. (2009b). An evaluation of models for diotic and dichotic detection in reproducible noises. *J Acoust Soc Am, 126*(4), 1906-1925. doi:10.1121/1.3206583

Davis, K. A., & Young, E. D. (2000). Pharmacological evidence of inhibitory and disinhibitory neuronal circuits in dorsal cochlear nucleus. *Journal of Neurophysiology, 83*(2), 926-940. doi:10.1152/jn.2000.83.2.926

Deng, L., & Geisler, C. D. (1987). Responses of Auditory-Nerve Fibers to Multiple-Tone Complexes. *Journal of the Acoustical Society of America, 82*(6), 1989-2000. doi:Doi 10.1121/1.395643

Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. *J Acoust Soc Am, 97*(1), 585-592. doi:10.1121/1.413112

Egan, J. P. (1975). *Signal detection theory and ROC-analysis*. New York: Academic Press.

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics, 12*(1), 0047-0065. doi:DOI 10.1103/RevModPhys.12.47

Gai, Y., & Carney, L. H. (2006). Temporal measures and neural strategies for detection of tones in noise based on responses in anteroventral cochlear nucleus. *Journal of Neurophysiology, 96*(5), 2451-2464. doi:10.1152/jn.00471.2006

Gai, Y., & Carney, L. H. (2008). Statistical analyses of temporal information in auditory brainstem responses to tones in noise: Correlation index and spike-distance metric. *Jaro-Journal of the Association for Research in Otolaryngology, 9*(3), 373-387. doi:10.1007/s10162-008-0129-8

Goeman, J. J., & Solari, A. (2014). Multiple hypothesis testing in genomics. *Stat Med, 33*(11), 1946-1978. doi:10.1002/sim.6082

Goupell, M. J. (2012). The role of envelope statistics in detecting changes in interaural correlation. *J Acoust Soc Am, 132*(3), 1561-1572. doi:10.1121/1.4740498

Greenwood, D. (1961). Auditory Masking and Critical Band. *Journal of the Acoustical Society of America, 33*(4), 484-&. doi:Doi 10.1121/1.1908699

Heffner, H. E., & Heffner, R. S. (2007). Hearing ranges of laboratory animals. *J Am Assoc Lab Anim Sci, 46*(1), 20-22. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17203911

Henry, K. S., Neilans, E. G., Abrams, K. S., Idrobo, F., & Carney, L. H. (2016). Neural correlates of behavioral amplitude modulation sensitivity in the budgerigar midbrain. *Journal of Neurophysiology, 115*(4), 1905-1916. doi:10.1152/jn.01003.2015

Jiang, D., McAlpine, D., & Palmer, A. R. (1997a). Detectability index measures of binaural masking level difference across populations of inferior colliculus neurons. *Journal of Neuroscience, 17*(23), 9331-9339. Retrieved from <Go to ISI>://WOS:A1997YG47000039

Jiang, D., McAlpine, D., & Palmer, A. R. (1997b). Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions. *Journal of Neurophysiology, 77*(6), 3085-3106. Retrieved from <Go to ISI>://WOS:A1997XJ17700019

Joris, P. X., Louage, D. H., Cardoen, L., & van der Heijden, M. (2006). Correlation index: a new metric to quantify temporal coding. *Hear Res, 216-217*, 19-30. doi:10.1016/j.heares.2006.03.010

Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Reviews, 84*(2), 541-577. doi:10.1152/physrev.00029.2003

Kidd, G., Jr., Mason, C. R., Brantley, M. A., & Owen, G. A. (1989). Roving-level tone-in-noise detection. *J Acoust Soc Am, 86*(4), 1310-1317. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2808906

Kohlrausch, A., Fassel, R., vanderHeijden, M., Kortekaas, R., vandePar, S., Oxenham, A. J., & Puschel, D. (1997). Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations. *Acustica, 83*(4), 659-669. Retrieved from <Go to ISI>://WOS:A1997XU29500009

Krishna, B. S., & Semple, M. N. (2000). Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *Journal of Neurophysiology, 84*(1), 255-273. doi:10.1152/jn.2000.84.1.255

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J Acoust Soc Am, 49*(2), Suppl 2:467+. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/5541744

Mao, J., Vosoughi, A., & Carney, L. H. (2013). Predictions of diotic tone-in-noise detection based on a nonlinear optimal combination of energy, envelope, and fine-structure cues. *J Acoust Soc Am, 134*(1), 396-406. doi:10.1121/1.4807815

Moore, B. C. J. (2007). Basic auditory processes involved in the analysis of speech sounds. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1493), 947-963 %@ 0962-8436.

Nelson, P. C., & Carney, L. H. (2007). Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *Journal of Neurophysiology, 97*(1), 522-539. doi:10.1152/jn.00776.2006

Ono, M., Bishop, D. C., & Oliver, D. L. (2017). Identified GABAergic and Glutamatergic Neurons in the Mouse Inferior Colliculus Share Similar Response Properties. *Journal of Neuroscience, 37*(37), 8952-8964. doi:10.1523/JNEUROSCI.0745-17.2017

Patterson, R. D. (1974). Auditory filter shape. *J Acoust Soc Am, 55*(4), 802-809. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/4833074

Pumplin, J. (1985). Low-noise noise. *The Journal of the Acoustical Society of America, 78*(1), 100-104 %@ 0001-4966.

Quiroga, R. Q., Nadasdy, Z., & Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput, 16*(8), 1661-1687. doi:10.1162/089976604774201631

Ramachandran, R., Davis, K. A., & May, B. J. (1999). Single-unit responses in the inferior colliculus of decerebrate cats. I. Classification based on frequency response maps. *Journal of Neurophysiology, 82*(1), 152-163. doi:10.1152/jn.1999.82.1.152

Ramachandran, R., Davis, K. A., & May, B. J. (2000). Rate representation of tones in noise in the inferior colliculus of decerebrate cats. *Jaro, 1*(2), 144-160. doi:DOI 10.1007/s101620010029

Rees, A., & Palmer, A. R. (1988). Rate-Intensity Functions and Their Modification by Broad-Band Noise for Neurons in the Guinea-Pig Inferior Colliculus. *Journal of the Acoustical Society of America, 83*(4), 1488-1498. doi:Doi 10.1121/1.395904

Richards, V. M. (1992). The detectability of a tone added to narrow bands of equal-energy noise. *J Acoust Soc Am, 91*(6), 3424-3435. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1619118

Richards, V. M., Heller, L. M., & Green, D. M. (1991). The detection of a tone added to a narrow band of noise: the energy model revisited. *Q J Exp Psychol A, 43*(3), 481-501. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1775653

Rocchi, F., & Ramachandran, R. (2018). Neuronal adaptation to sound statistics in the inferior colliculus of behaving macaques does not reduce the effectiveness of the masking noise. *Journal of Neurophysiology, 120*(6), 2819-2833. doi:10.1152/jn.00875.2017

Rose, G. J., & Capranica, R. R. (1985). Sensitivity to amplitude modulated sounds in the anuran auditory nervous system. *Journal of Neurophysiology, 53*(2), 446-465. doi:10.1152/jn.1985.53.2.446

Schwarz, D. M., Zilany, M. S., Skevington, M., Huang, N. J., Flynn, B. C., & Carney, L. H. (2012). Semi-supervised spike sorting using pattern matching and a scaled

Mahalanobis distance metric. *J Neurosci Methods, 206*(2), 120-131. doi:10.1016/j.jneumeth.2012.02.013

van de Par, S., & Kohlrausch, A. (1999). Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters. *J Acoust Soc Am, 106*(4 Pt 1), 1940-1947. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10530018

van Rossum, M. C. (2001). A novel spike distance. *Neural Comput, 13*(4), 751-763. doi:10.1162/089976601300014321

Whitehead, M. L., Lonsburymartin, B. L., & Martin, G. K. (1992). Evidence for 2 Discrete Sources of 2f1-F2 Distortion-Product Otoacoustic Emission in Rabbit .1. Differential Dependence on Stimulus Parameters. *Journal of the Acoustical Society of America, 91*(3), 1587-1607. doi:Doi 10.1121/1.402440

Winer, J. A., Schreiner, C. E., & SpringerLink. (2005). *The Inferior Colliculus* [text](pp. XX, 705 p. 168 illus). doi:10.1007/b138578

Young, E. D., & Brownell, W. E. (1976). Responses to tones and noise of single cells in dorsal cochlear nucleus of unanesthetized cats. *Journal of Neurophysiology, 39*(2), 282-300. doi:10.1152/jn.1976.39.2.282

Zheng, L., Early, S. J., Mason, C. R., Idrobo, F., Harrison, J. M., & Carney, L. H. (2002). Binaural detection with narrowband and wideband reproducible noise maskers: II. Results for rabbit. *Journal of the Acoustical Society of America, 111*(1), 346-356. doi:Doi 10.1121/1.1423930

Zilany, M. S. A., & Bruce, I. C. (2007). Representation of the vowel (epsilon) in normal and impaired auditory nerve fibers: Model predictions of responses in cats. *Journal of the Acoustical Society of America, 122*(1), 402-417. doi:10.1121/1.2735117

Chapter 3. Responses to Dichotic Tone-in-Noise Stimuli in the Inferior Colliculus

3.1 Abstract

Human listeners are more sensitive to tones embedded in diotic noise when the tones are out-of-phase at the two ears (N_0S_π) than when they are in-phase (N_0S_0). The difference between the tone-detection thresholds for these two conditions is called the binaural masking level difference (BMLD), which reflects a benefit of binaural processing. Detection in the N_0S_0 condition has recently been shown to be better explained by envelope-based cues than by energy-based cues. Detection in the N_0S_π condition has been explained in modeling studies by changes in interaural correlation (IAC). Physiological studies of the BMLD have focused on neural responses to interaural time differences (ITDs). Here, responses to N_0S_π and N_0S_0 stimuli were recorded in the inferior colliculus of awake rabbits. ITD-based, IAC-based, and envelope-based hypotheses for binaural detection were examined here. Results confirmed that for the N_0S_0 condition changes in response rates upon addition of a tone could be explained by a neuron's sensitivity to envelope cues at most noise levels. For the N_0S_π condition, ITD-, IAC-, and envelope-based cues all contributed to neural responses, and these cues are not independent. Average rate-based thresholds were calculated for both N_0S_0 and N_0S_π conditions to estimate neural BMLDs. The neural BMLD at 500 Hz matched the behavioral BMLD in rabbit; the most sensitive neural BMLDs had a different trend across frequency than human BMLDs.

3.2 Introduction

Performance of human listeners benefits from binaural hearing. For example, in the tone-in-noise (TIN) detection task, the threshold for detection of out-of-phase tone in identical noise at the two ears (N_0S_π) is lower (i.e. better) than that for detection of an in-phase tone (N_0S_0) (e.g., Hirsh, 1948; Hawley, Litovsky, & Culling, 2004). Several binaural cues have been proposed to explain the improvement in detection, but these cues have not been examined together in physiological studies.

In N_0S_π stimuli, the difference between the tone-plus-noise waveforms at the two ears results in a decrease in the interaural correlation (IAC, e.g., Bernstein & Trahiotis, 1996). Several models for binaural detection have focused on detection of a decrease in IAC (e.g., Colburn, 1977; Bernstein & Trahiotis, 1997, 2017). Physiological studies of detection of tones in N_0S_π stimuli have focused on different types of sensitivity to interaural time differences (ITDs) (Jiang, McAlpine, & Palmer, 1997a, 1997b) due to the correlation between IAC and ITD: binaural stimuli that have larger ITDs have lower IACs. Responses to ITDs are described by the noise-delay function (NDF), average rate as a function of ITD, and by the ITD that elicits the strongest/weakest response, referred to as the best/worst ITD. For ITD-sensitive neurons, there are two typical shapes: peak-like, with a best ITD close to zero and decreasing rates for increasing ITDs, and trough-like, with a worst ITD close to zero and increasing rates for increasing ITDs. Because ITDs are introduced upon addition to diotic noise of an out-of-phase tone at the two ears (N_0S_π), neurons with peak-like NDFs have been hypothesized to have decreased average rates with increasing tone level, whereas neurons with trough-like NDF would have increased average rates.

Because many low-frequency IC neurons are sensitive to ITDs, this ITD-based hypothesis was tested in the inferior colliculus (IC) (Jiang et al., 1997a), but only for 500-Hz tones.

The difference in detection thresholds between the N_0S_0 and N_0S_π conditions is called the binaural masking level difference (BMLD). Human listeners can have substantial BMLDs (> 3 dB) up to at least 8 kHz (van de Par & Kohlrausch, 1999; Goupell, 2012), yet modeling and physiological studies have mainly focused on low frequencies, where the BMLD is typically larger (up to 20 dB, depending on bandwidth) (e.g., van de Par & Kohlrausch, 1999).

The current study aimed to expand the scope of physiological studies of N_0S_π detection in the inferior colliculus (IC) to a wider frequency range. Neural responses were also recorded over a wide range of noise levels to explore trends across sound level, which have not previously been reported in physiological studies, but human psychophysical studies showed that BMLDs are robust across a range of noise levels (Buss et al., 2003) and roving-level paradigm (i.e. stimulus level randomly varies from interval to interval) (Henning et al., 2005). One goal of this study was to examine changes in neural responses upon addition of a dichotic tone to test the ITD-based hypothesis. Note that there is not a clear prediction for high-frequency binaural detection based on ITD processing, because the neural ability to phase-lock to temporal fine structure decreases with increasing frequency.

IC neurons receive afferent inputs from almost all lower levels of the auditory system (Cant & Oliver, 2018), therefore IC neurons are sensitive to many different features of stimuli, including ITDs (e.g., Wagner, Takahashi, & Konishi, 1987) and envelope frequency and depth (e.g., Langner & Schreiner, 1988; Nelson & Carney, 2007; Zheng &

Escabi, 2013). An envelope-based hypothesis predicts changes in neural responses to N_0S_0 stimuli better than models based on classical energy cues (Chapter 2), and thus this hypothesis was further tested with detection under the N_0S_π condition. Briefly, adding a tone to narrowband noise flattens the stimulus envelope (Richards, 1992) and subsequently the neural fluctuations (review: Carney, 2018). The sensitivity of IC neurons to amplitude-modulation frequency is described by modulation transfer functions (MTFs, average rate as a function of modulation frequency); for example, band-enhanced (BE) and band-suppressed (BS) MTFs have average rates responses over some range of modulation frequencies that are stronger or weaker than the response to unmodulated stimuli. Thus, IC neurons were hypothesized to have average rate responses that were positively or negatively correlated with the neural fluctuation amplitudes that are set up in the auditory periphery and projected up to the midbrain. Therefore, according to an envelope- or neural-fluctuation-based hypothesis, neurons with BE MTFs were expected to have decreasing rate with increasing tone level for both N_0S_0 and N_0S_π conditions, whereas neurons with BS MTFs were expected to have increasing rate with tone level for both stimulus conditions.

In addition to the ITD- and envelope-based hypotheses, the role of IAC in neural responses to N_0S_π stimuli was evaluated. As mentioned earlier, neural sensitivity to ITD has been proposed to be the physiological basis for the IAC mode (Jiang et al., 1997), but recent studies have shown that predictions for tone detection differ for these two cues (van der Heijden & Joris, 2010; Culling, 2011). If interaural decorrelation explains the neural responses to N_0S_π stimuli, then the difference in average rate between IC responses to diotic noise and uncorrelated noise was expected to be correlated to the rate difference

between responses to the noise-alone condition and the noise-plus-dichotic-tone condition. This correlation was directly tested in this study.

The current study also computed rate-based IC neural thresholds for comparison with published detection thresholds for human listeners (Buss, Hall, & Grose, 2003; Goupell, 2012; van de Par & Kohlrausch, 1999) and rabbits (Zheng et al., 2002).

3.3 Materials and Methods

3.3.1 Procedures

Surgical and recording procedures were the same as described in Chapter 2. Briefly, rabbits were anesthetized with an intramuscular injection of ketamine (66 mg/kg) and xylazine (2 mg/kg) for both the headbar placement and microdrive (five-drive, Neuralynx, Inc., Bozeman, MTF, USA) implantation surgeries. The headbar was a custom-designed, 3D-printed hard plastic with a chamber that held the microdrive. The headbar was permanently mounted on the rabbit scull with stainless-steel screws and dental acrylic. A craniotomy was allowed for insertion of the microdrive after the rabbit recovered from the headbar surgery. One microdrive held four tetrodes and allowed for advancing and retracting tetrodes. Each tetrode consisted of four twisted 18-µm platinum iridium wires, coated in epoxy (California Fine Wire Co., Grover Beach, CA, USA). The microdrive was replaced as needed to search for new neurons.

During recording sessions, the rabbit was placed in a double-walled, sound-proof chamber, with head fixed using the headbar. Sound was delivered using Beyerdynamic DT990 (Beyerdynamic GmbH & Co., Heilbronn, Germany) or Etymotic ER2 earphones (Etymotic Research, Inc., Elk Grove Village, Illinois USA) with custom ear molds for each

rabbit. Ear molds were positioned deep in the concha and included an Etymotic probe tube for calibration. Search stimuli (wideband noise bursts) were presented to locate auditory responses. When the characteristic frequencies (CF) increased with increasing tetrode depth, the tetrodes were determined to be advanced into the central nucleus of the IC (ICC). Post-mortem histology is applied to verify the tetrode location. Recordings were sorted offline, as described in Chapter 2.

3.3.2 Stimuli

Speakers were calibrated with ER-7C or ER-10B+ microphones (Etymotic Research, Inc., USA) at the beginning of each recording session. The neurons were characterized in several ways before being presented with TIN stimuli. The neuron's binaural sensitivity was determined by responses to contralateral, ipsilateral, and binaural wideband noise (0.1-19 kHz) at several sound levels. Responses to pure tones between 0.25 and 20 kHz from 10 to 70 dB SPL were used to identify CF, the frequency at which the neuron responded at the lowest sound level. The NDF describes rate responses to noise stimuli as a function of ITD; NDFs were recorded with wideband noise (0.1-19 kHz), a duration of one sec, and spectrum level of 30 dB SPL. The ITDs ranged from -2000 to 2000 μs with a step of 200 μs. Responses to ILDs were recorded with the same noise bandwidth and duration as for the NDF. ILDs ranged from -15 to 15 dB with a step of 5 dB; the stimulus on the contralateral side had a fixed spectrum level of 30 dB SPL. Responses to sinusoidally-amplitude-modulated (SAM) noise were collected to identify the shape of the MTF. One-second-duration SAM noises were created with the following equation:

$$s = [1 + \sin(2\pi f_m t)]/2 * n$$

where n is the wideband noise with a spectrum level of 30 dB SPL, and f_m is the modulation frequency. Modulation frequencies used in the study were logarithmically spaced between 2 to 350 Hz. Responses to unmodulated noise were also recorded.

For TIN stimuli, the tone frequency and the center frequency of the 1/3-oct gaussian noise masker were chosen to be approximately equal to CF. Noise maskers were generated by filtering wideband noise with a 5000th-order FIR band-pass filter. Out-of-phase tone was added to the two ears before ramping. TIN stimuli had a duration of 0.3 sec and a cos^2 ramp duration of 0.01 sec (*tukeywin* in Matlab). Overall noise levels ranged from 35 to 75 dB SPL, with a step of 10 dB. Signal-to-noise ratio (SNR) ranged from -12 to 8 dB, with a step of 4 dB, plus the no-tone condition. Tone levels and noise levels were presented in random order, and the order was shuffled for each of the 30 repetitions of the stimulus set. If more than one dataset was recorded, the dataset with responses to multiple random noise waveforms was used.

To test the role of IAC, responses to diotic and dichotic noise were recorded. Diotic noise was binaurally identical noise (N_0), whereas dichotic noise was binaurally uncorrelated noise (N_u). Both N_0 and N_u condition used 1/3-octave gaussian narrowband noise, with a duration of 2 sec and sound level of 65 dB SPL.

3.4 Analysis

NDF shape classification

The shape of NDF, the best ITD (d_{BITD}), and the frequency of ITD tuning (f_{ITD}) were determined by fitting the NDF with a Gabor function (Lane & Delgutte, 2005), a sinusoid modulated by a gaussian function:

$$G_1 = Ae^{-(d_{ITD}-d_{BITD})^2/2\sigma^2}\cos[2\pi f_{ITD}(d_{ITD}-d_{BITD})] + B,$$

where A, B, and σ are parameters for the amplitude, DC offset, and the standard deviation of the gaussian function. If the neuron's CF was more than twice f_{ITD} (i.e. a high frequency neuron), which indicated that the neuron did not have fine-structure-based ITD sensitivity, f_{ITD} was set to be zero, and the NDF were refitted with the following gaussian function:

$$G_2 = Ae^{-(d_{ITD}-d_{BITD})^2/2\sigma^2} + B.$$

In either case, the function was half-wave rectified to remove negative rates and to better match NDF. The least-square fit was obtained with trust-region-reflective algorithm (Matlab's *lsqcurvefit*).

NDF was classified into peak-like, trough-like, and insensitive categories. In the following cases, the neuron was considered as having ITD sensitivity: 1) for NDFs fitted with G_1, if the absolute amplitude (A) was more than 5; 2) for NDFs fitted with G_2, if the prominence (A/B) was more than 0.25; 3) for NDFs fitted with G_2, for a fit $6 \cdot 10e^{-5} < \sigma < 0.001$. If the amplitude (A) was positive, the neuron was classified as having peak-like NDF; otherwise, the neuron was classified as having trough-like NDF. Other neurons were classified as insensitive. The classification of NDF generally agreed with the qualitative description. Example NDF shapes are shown in Fig. 3.1.

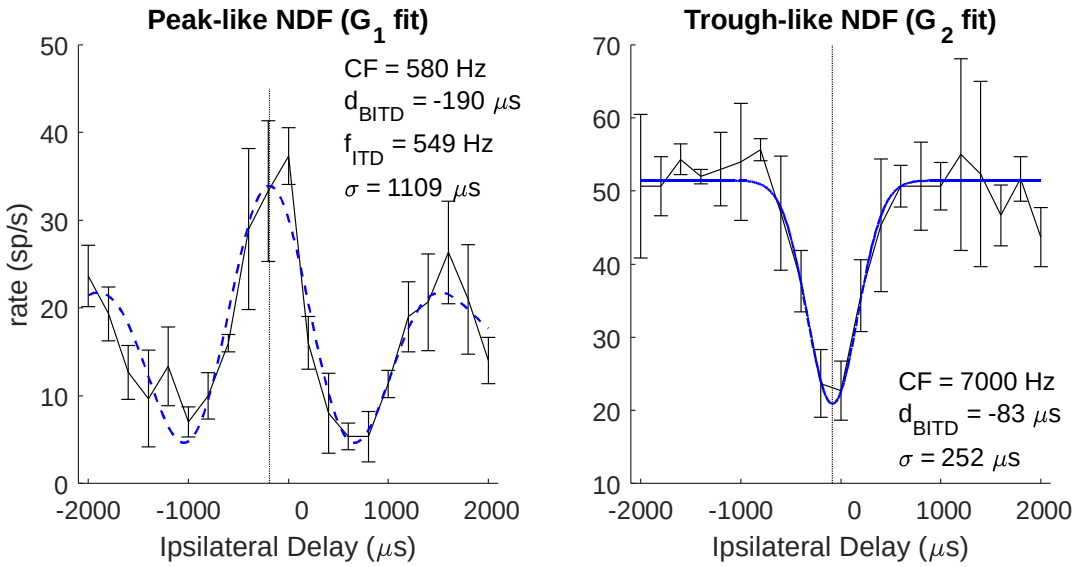

Figure 3.1: Example neural ITD responses (black solid curve) and fitted Gabor function (blue dashed curve) for peak-like (left) and trough-like (right) NDFs. Vertical dotted line indicates the best ITD. The neuron's CF, Best ITD, ITD tuning frequency (f_{ITD}) are described in the text.

MTF shape classification

The MTF shape was classified with rules designed to be simple and to agree with the qualitative descriptions of these functions. Enhancement or suppression was identified with the Mann-Whitney test as significantly higher or lower response rates than the response to unmodulated noise at two or more neighboring modulation frequencies. The presence or absence of enhancement or suppression was used to classify the MTF into the following four types: all-pass (AP, no enhancement or suppression), band-enhanced (BE, only enhancement), band-suppressed (BS, only suppression), and hybrid (both enhancement and suppression, at different modulation frequencies). The MTF was classified as unresponsive if response rates to all stimulus presentations were lower than 5 spikes/s.

Rate analysis

Average rates excluding 20-ms onset responses were calculated for all stimulus types. For TIN stimuli, at each noise and tone level (i.e. SNR), a rate-based receiver-operating-characteristic (ROC, Egan, 1975) was calculated using rate responses for all noise and tone-plus-noise presentations. The percent-correct performance was estimated from the area under the ROC curve. Note that rates in response to tone-plus-noise stimuli could be either higher or lower than rates in response to noise-alone stimuli, so the minimum percent correct was limited to 50%, regardless of the direction of change in rate. Linear interpolation was used to find the lowest SNR with 70.7% correct as the neural threshold, corresponding to human threshold estimated with a two-down, one-up tracking procedure (Levitt, 1971). For some sensitive neurons, the lowest SNR tested was above threshold, and the lowest SNR tested is indicated as threshold in the result figures for these neurons.

3.5 Results

Responses to both N_0S_0 and N_0S_π stimuli were recorded from 136 isolated single units (111 units were the same as presented in the rate analysis of Chapter 2). The distribution of the CFs is shown in Fig. 3.2. All units were tested using a tone frequency within 1/3-octave of the neuron's CF. Based on the MTF categorization criteria described above, there were 40 BE units (29.4%), 62 BS units (45.6%), 12 hybrid units (8.8%) and 22 AP units (16.2%).

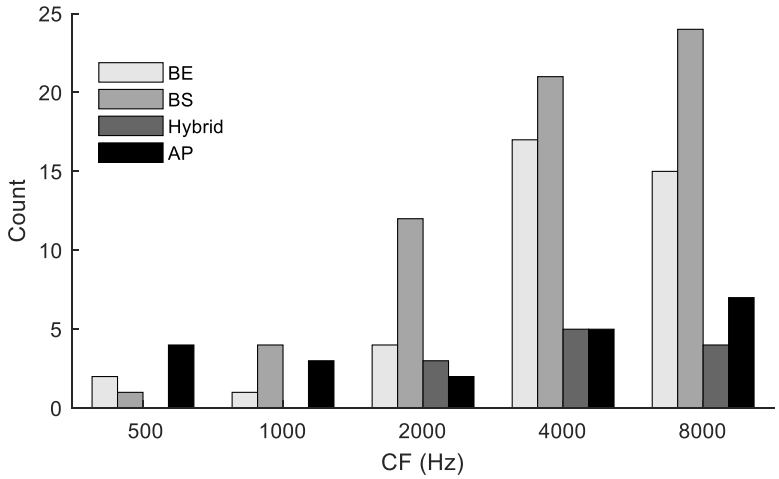

Figure 3.2: Distribution of MTFs across CF (in one-octave bins) for the units presented in this study. Gray shades from light to dark indicate units with band-enhanced (BE), band-suppressed (BS), hybrid and all-pass (AP) MTF shapes. Two neurons with CF of 12.1k were included in the last bin for simplicity. Most MTF types were represented across the range of CFs, although hybrid MTFs were not observed at the lower CFs.

Responses of single neurons

Figure 3.3 shows responses of two example neurons. Neuron 1 (top row) had a BE MTF and Neuron 2 had a BS MTF (bottom row). Recall that based on the envelope hypothesis, adding a tone to narrowband gaussian noise flattens the envelope and neural fluctuation. Neurons with BE MTF are excited by fluctuations and therefore were expected to have decreasing rate with increasing SNR. On the contrary, neurons with BS MTF are suppressed by fluctuations and therefore were expected to have increasing rate with increasing SNR. Neuron 1 had decreasing rate versus SNR at all noise levels, and Neuron 2 had increasing rate versus SNR at all noise levels; both responded as expected, supporting the envelope-based hypothesis. Note that average rate changed at lower SNR

for the N_0S_π condition than for the N_0S_0 condition, indicating lower neural thresholds, consistent with psychophysical results.

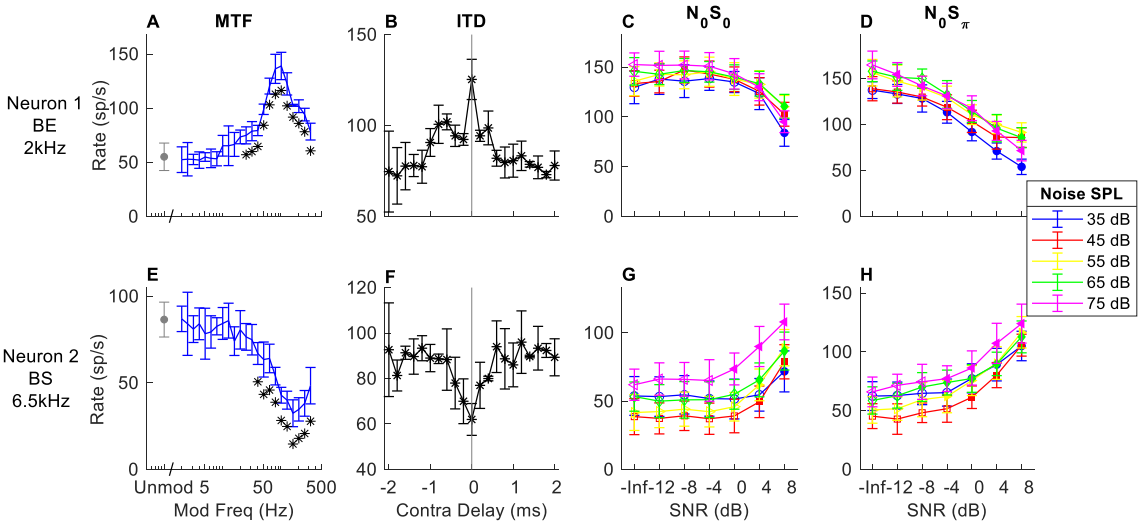

Figure 3.3: Responses of two example neurons (top and bottom row respectively). A and E: MTF, response rates to amplitude-modulated noise; stars indicate modulation frequencies that had rates significantly different from the unmodulated condition. B and F: ITD sensitivity, response rates vs. time delay in contralateral side (negative indicates ipsilateral side has delay). C, D, G and H: responses to N_0S_0 and N_0S_π stimuli at different noise levels (different symbols) vs. SNR (from left to right); filled symbols indicate supra-thresholds. Errorbars indicate standard deviation. MTF shape and tone frequency for TIN stimuli (close to CF) are shown on the left. The example BE neuron had decreasing rate upon addition of a tone for both N_0S_0 and N_0S_π, while the example BS neuron had increasing rate for both conditions.

Neural responses to N_0S_0 stimuli have previously been simply described as having increasing rate as a function of tone level (Jiang et al., 1997a; Ramachandran, May, Davis, 2000), possibly based on the assumption that neurons respond more strongly to increasing stimulus energy (i.e. upon addition of a tone). However, Neuron 1 had decreasing rate versus SNR, which would not be expected based on stimulus energy. The shape of NDF has been used to explain changes in neural responses for the N_0S_π

condition (Jiang et al., 1997a, 1997b): a diotic noise masker has zero ITD; adding a dichotic tone introduces a non-zero ITD. Neurons with peak-like NDF respond most strongly to near-zero ITDs, and thus are expected to have decreasing rate with increasing SNR. In contrast, neurons with trough-like NDF are expected to have increasing rate with increasing SNR. Responses to N_0S_π stimuli of Neuron 1 and Neuron 2 can also be explained by their NDF shape: Neuron 1 had peak-like NDF shape and decreasing rate versus SNR for the N_0S_π condition; Neuron 2 had trough-like NDF and increasing rate versus SNR.

Responses in the IC can also be complicated and not well explained by a single hypothesis. Figure 3.4 shows six other example neurons with different MTF and NDF shapes. Neurons 3 and 4 both had BE MTFs and decreasing rate versus SNR for the N_0S_0 condition at most noise levels, as expected. However, for the N_0S_π condition, Neuron 3 had decreasing rate versus SNR that could be explained by its MTF shape, but not its trough-like NDF. In contrast, Neuron 4 had an increasing rate versus SNR that could be explained by its NDF shape, but not by its MTF shape. Neurons 5, 6 and 7 all had BS MTFs, and thus were expected to have increasing rate versus SNR, but the responses of these neurons differ. Neuron 5 had increasing rate versus SNR for both N_0S_0 and N_0S_π condition, which could be explained by its BS MTF, but not by its peak-like NDF. The MTF of Neuron 6 did not explain responses to either N_0S_0 or N_0S_π stimuli, but responses to N_0S_π stimuli (decreasing rate) could be explained by its peak-like NDF. Neuron 7 also had decreasing rate versus SNR, which could not be explained by either MTF or NDF shape. Neuron 8 had an all-pass MTF, and responses to N_0S_π stimuli could be explained by the peak-like NDF.

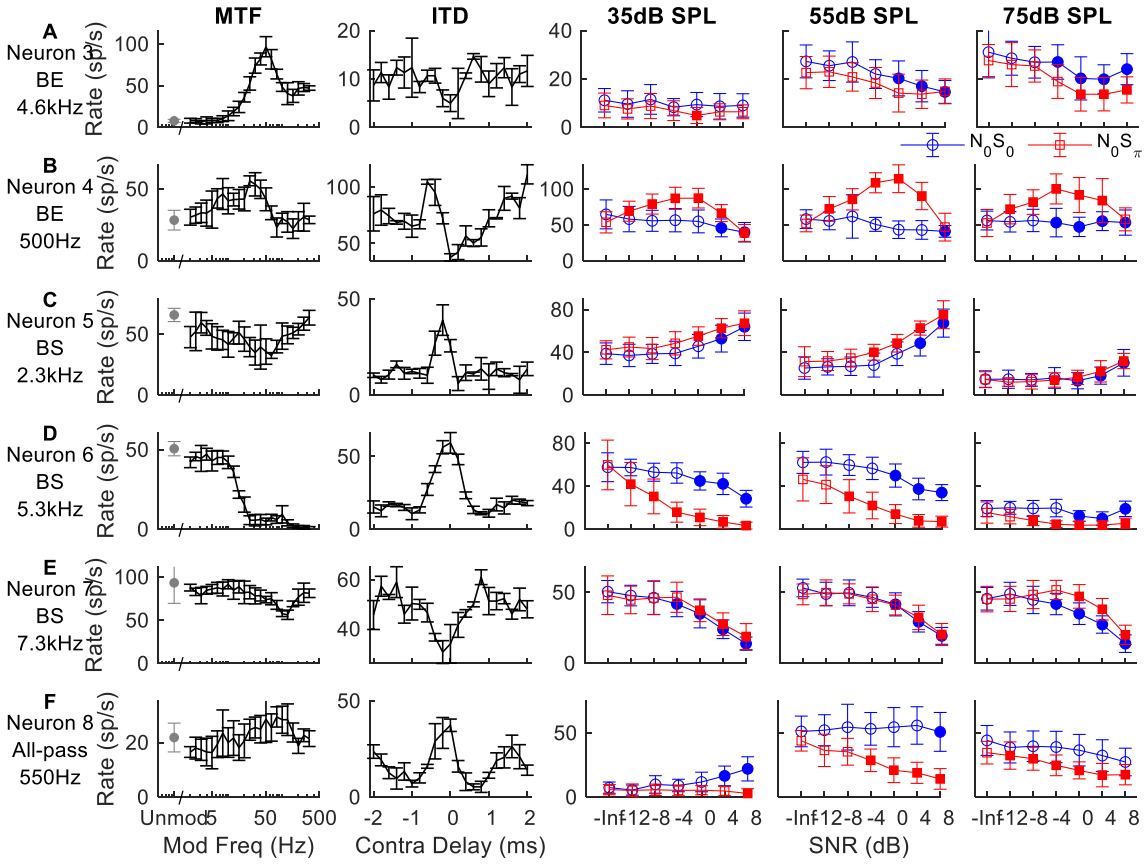

Figure 3.4: Responses of six example neurons (A-F). The left two columns show the neuron's MTF and ITD sensitivity, respectively. The right three columns show the neuron's response to N_0S_0 (blue circles) and N_0S_π (red squares) TIN stimuli at noise levels of 35, 55, and 75 dB SPL, respectively; filled symbols indicate supra-threshold responses. MTF shape and tone frequency of TIN stimuli (close to CF) are shown on the left.

Evaluating the envelope-based hypothesis for the population of IC responses

Responses of all neurons in the population were first evaluated by the correlation between maximum changes in rate elicited by amplitude modulations (using responses in the MTFs) and by addition of tone to a masker (using the TIN responses). For SAM responses, the maximum average rate difference in the MTF, between the maximum or minimum of the MTF and the response to the unmodulated stimulus, was determined for

each neuron. IC neurons with BE MTFs were assigned a positive maximum rate difference (Fig. 3.5, triangles), and BS neurons had a negative maximum rate difference (Fig. 3.5, squares). AP neurons had a small maximum rate difference that could be either positive or negative (Fig. 3.5, circles). The maximum rate difference of hybrid neurons depended on whether the enhancement or suppression was stronger (Fig. 3.5, diamonds). Similarly, for TIN responses, the maximum difference in average rate across all SNRs tested and that of the noise-alone stimulus was determined for each neuron. Correlations between rate differences due to addition of a tone and due to amplitude modulation were significantly correlated for both N_0S_0 (Fig. 3.5A) and N_0S_π conditions (Fig. 3.5B). The correlation coefficient was higher for the N_0S_π condition than for the N_0S_0 condition, possibly because the SNR range limited the change in rate elicited by a tone, and this limitation was reduced when the threshold was at the lower end of the SNR range (i.e. for the N_0S_π condition).

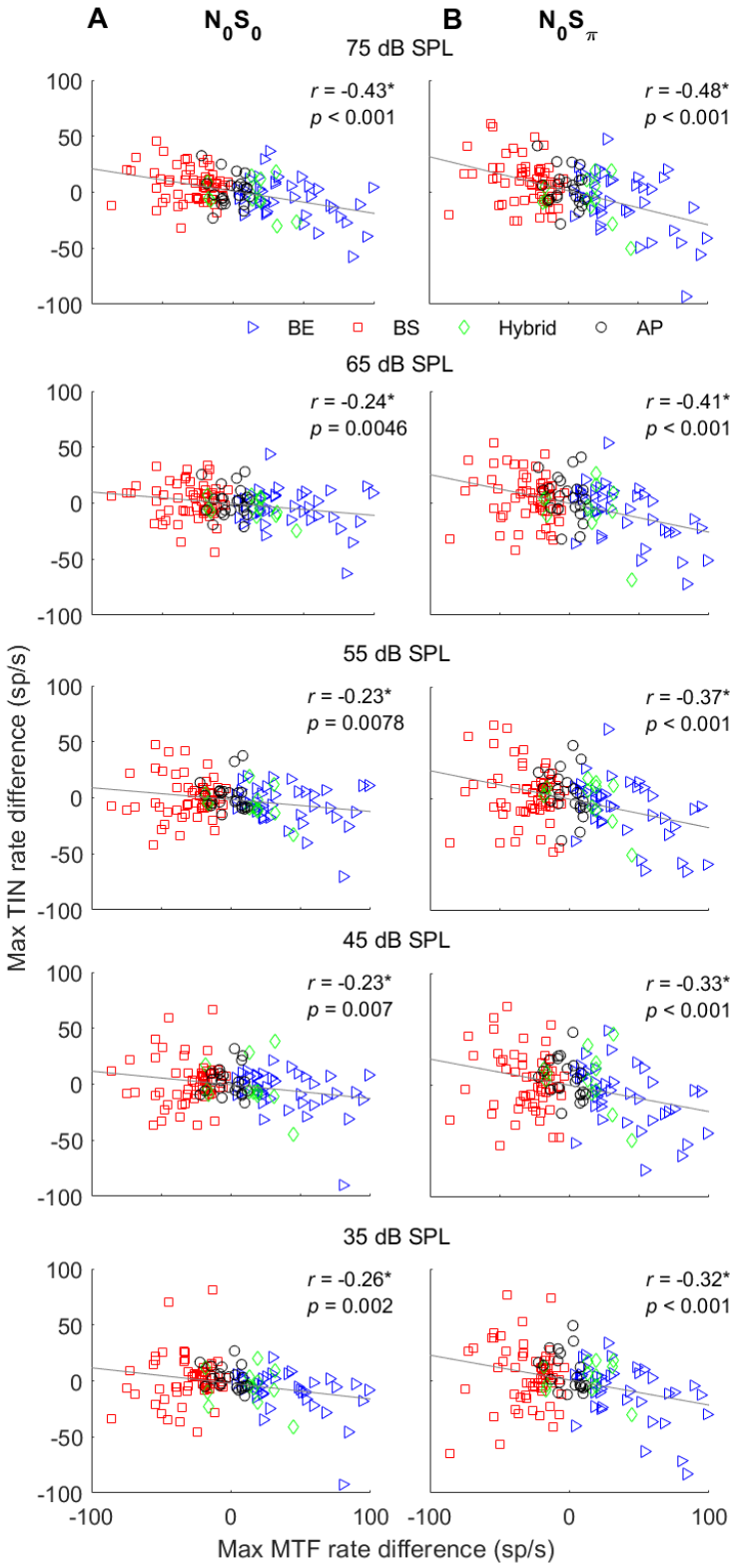
A
N_0S_0
B
N_0S_π
75 dB SPL
r = -0.43*
p < 0.001
r = -0.48*
p < 0.001
BE
BS
Hybrid
AP
65 dB SPL
r = -0.24*
p = 0.0046
r = -0.41*
p < 0.001
55 dB SPL
r = -0.23*
p = 0.0078
r = -0.37*
p < 0.001
45 dB SPL
r = -0.23*
p = 0.007
r = -0.33*
p < 0.001
35 dB SPL
r = -0.26*
p = 0.002
r = -0.32*
p < 0.001
Max TIN rate difference (sp/s)
Max MTF rate difference (sp/s)
100
50
0
-50
-100

Figure 3.5: Correlation between rate differences elicited by addition of a diotic (A) or dichotic (B) tone and by amplitude modulation (MTF). Neurons with different MTF types are shown in different symbols (see legend). BE and BS neurons had the largest rate differences in the MTFs, and thus the symbols for these neurons are located on the right and left side of each plot, as expected. Correlation coefficients and p values are shown in each plot; a star indicates that the correlation coefficient was significant after Bonferroni correction ($p < 0.01$). Rate differences in TIN responses were significantly correlated with those in MTFs for both N_0S_0 and N_0S_π conditions at all noise levels, consistent with the envelope-based hypothesis.

Another way to evaluate population responses was to calculate the proportion of IC neurons that had the expected rate-change direction in response to TIN stimuli based on the MTF shape (Fig. 3.6). Recall that the envelope-based hypothesis predicts decreasing rate for BE neurons (black) and increasing rate for BS neurons (gray). Also recall that rate-change direction was defined at the threshold for each neuron. For the N_0S_0 condition, the proportion of neurons with increasing/decreasing average rates were significantly different between the BE and BS groups at noise levels of 35 dB SPL ($\chi^2(1,1) = 8.98, p = 0.0027$), 65 dB SPL ($\chi^2(1,1) = 4.00, p = 0.046$) and 75 dB SPL ($\chi^2(1,1) = 11.95, p = 0.0005$), but not at noise levels of 45 dB SPL ($\chi^2(1,1) = 1.35, p = 0.24$) and 55 dB SPL ($\chi^2(1,1) = 1.70, p = 0.19$). For the N_0S_π condition, the proportions of neurons with increasing/decreasing average rates were significantly different between BE and BS group only at the noise level of 75 dB SPL ($\chi^2(1,1) = 4.52, p = 0.03$), but not at other levels: $\chi^2(1,1) = 3.50, 0.62, 3.04, 2.25$, and $p = 0.06, 0.43, 0.08, 0.13$ at noise levels of 35 to 65 dB SPL, respectively. Thus, BE and BS neurons responded as predicted by the envelope-based hypothesis for the N_0S_0 condition at most noise levels, but only at the highest tested noise level for the N_0S_π condition. The conclusions based on this MTF-

categorization method and on the correlation method (Fig. 3.5) were not the same for low-to-medium noise levels, possibly because the MTF-categorization method emphasized the number of neurons in the population with expected responses, whereas the correlation method emphasized the size of the rate change in responses of single neurons. This discrepancy will be further discussed later; in general, both analyses showed that envelope-based cues were most important at the highest noise level.

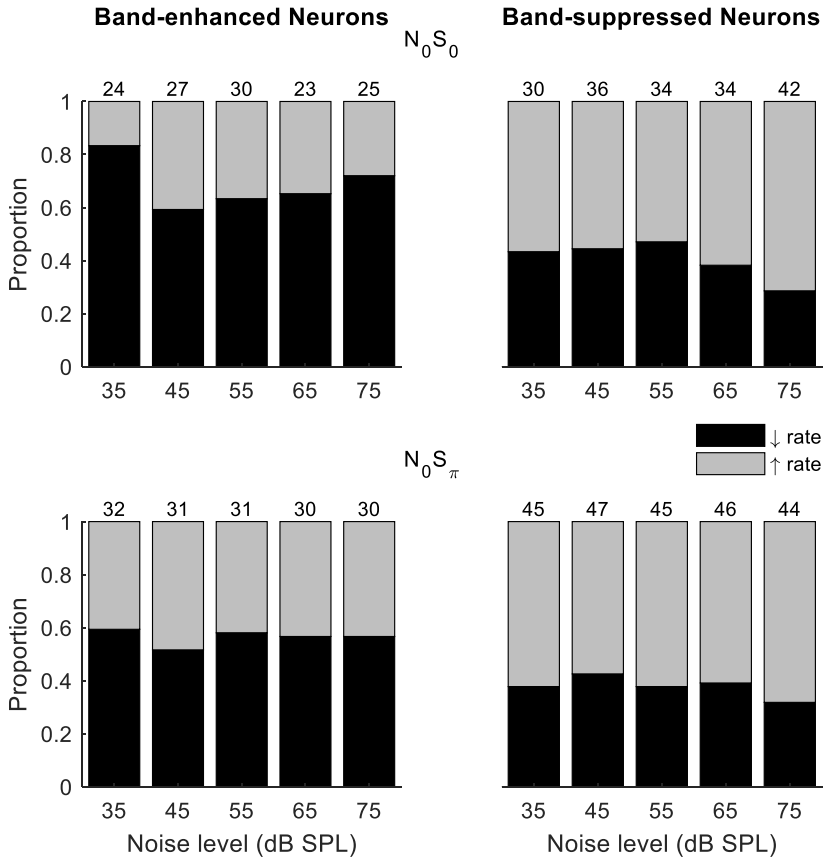

Figure 3.6: Proportions of increasing/decreasing average rate versus SNR of BE (left) and BS (right) neurons for both N_0S_0 (top) and N_0S_π (bottom) conditions. Only neurons that elicited measurable thresholds are included; the number is shown at the top of each bar. Gray indicates increasing rate vs. SNR at threshold, and black indicates decreasing rate. BE and BS neurons responded as expected based on MTF shape at most noise levels for the N_0S_0 condition, but only at 75 dB SPL for the N_0S_π condition.

Evaluating the ITD-based hypothesis for the population of IC responses

The shape of the NDF has previously been used to explain TIN detection thresholds for the N_0S_π condition, thus neural thresholds were also examined for different types of NDF: peak-like, trough-like, or insensitive. Recall that the ITD-based hypothesis predicts that neurons with peak-like NDF would have decreasing rate as a function of SNR for the N_0S_π condition, and neurons with trough-like NDF would have increasing rate versus SNR. The ITD-based hypothesis was also examined by the NDF-categorization and correlation methods. The proportion of neurons with each rate-change direction at threshold was computed for each NDF group (Fig. 3.7). For the N_0S_0 condition, a large proportion of neurons with peak- or trough-like NDFs had decreasing rate with increasing SNR, which has not been previously reported (Jiang et al., 1997a). For the N_0S_π condition, a χ^2 test showed that the proportion of neurons with increasing or decreasing rate vs. SNR was significantly different between neurons with peak- and trough-like ITD sensitivity at noise levels of 35 to 65 dB SPL ($\chi^2(1,1) = 17.23, 20.96, 13.34, 8.28$, all $p < 0.005$); but not at 75 dB SPL ($\chi^2(1,1) = 1.81;\ p = 0.18$). For neurons with trough-like ITD sensitivity, the number of neurons that had increasing rate at threshold (as expected) decreased with increasing noise level (Fig. 3.7, bottom middle plot). For neurons that were insensitive to ITD, an increasing number of neurons had increasing rate at threshold at higher noise levels for both N_0S_0 and N_0S_π conditions.

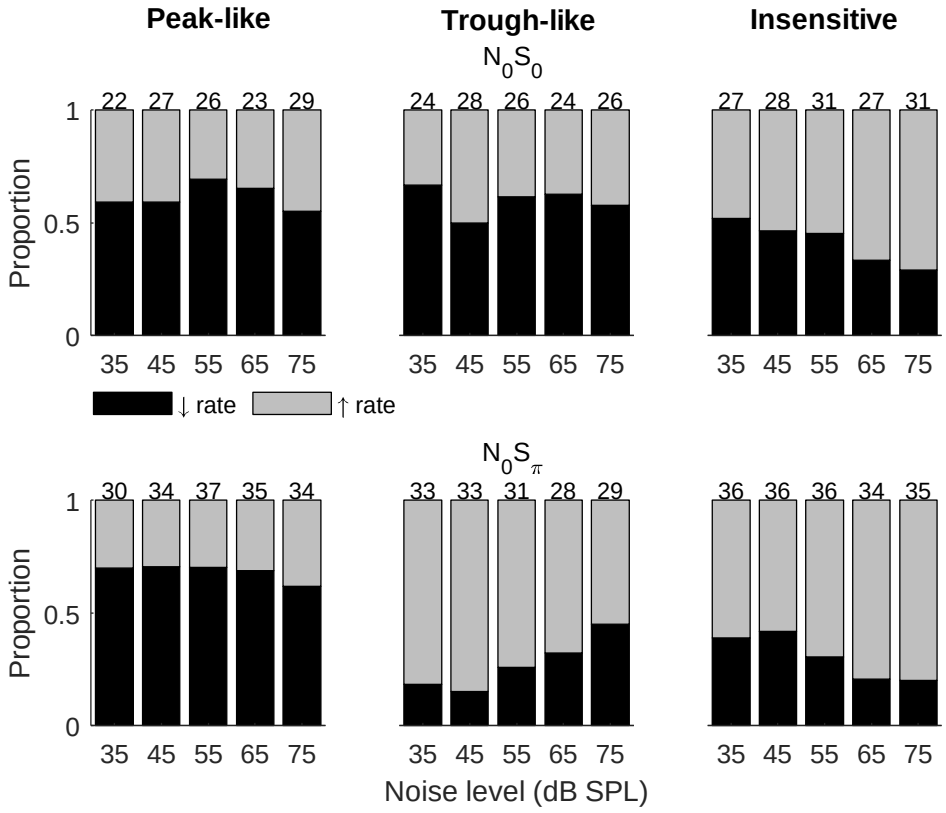

Figure 3.7: Proportions of rate-change direction at neural threshold for neurons with peak-like (left), trough-like (middle) NDF and ITD-insensitive (right) for the N_0S_0 (top) and N_0S_π (bottom) conditions. Gray indicates increasing rate and black indicates decreasing rate at threshold. Only neurons with measurable thresholds are shown here; the number of neurons is shown at the top of each bar. For the N_0S_π condition, the proportion of neurons with increasing/decreasing rate vs. SNR was significantly different between neurons with peak- or trough-like NDF at most noise levels.

The correlation between the maximum rate difference in response to N_0S_0 and N_0S_π stimuli a nd the maximum rate difference in the NDF was also computed to evaluate the ITD-based hypothesis. The difference between average rate in response to zero ITD and mean rate in response to absolute ITDs larger than 1 ms was computed for each neuron. For TIN responses, it is typical to compare the difference in rate between

responses to N_0 to N_0S_π for neurons with different NDF types (Fig. 3.7; Jiang et al., 1997); however, such evaluation may ignored changes in response to diotic tone (i.e. N_0S_0 to N_0S_π instead of N_0 to N_0S_π). Therefore, the difference between average rates in response to N_0S_0 and N_0S_π stimuli was computed at each SNR, and a test for a significant correlation was carried out between the maximum difference between diotic and dichotic TIN stimuli and the maximum rate difference in the NDF (Fig. 3.8). The correlation was significant at all noise levels, supporting the ITD-based hypothesis. However, the correlation coefficient decreased with increasing noise level. The results of NDF-categorization and correlation methods were consistent, suggesting that the role of ITD was important, but that its importance decreased as masker level increased.

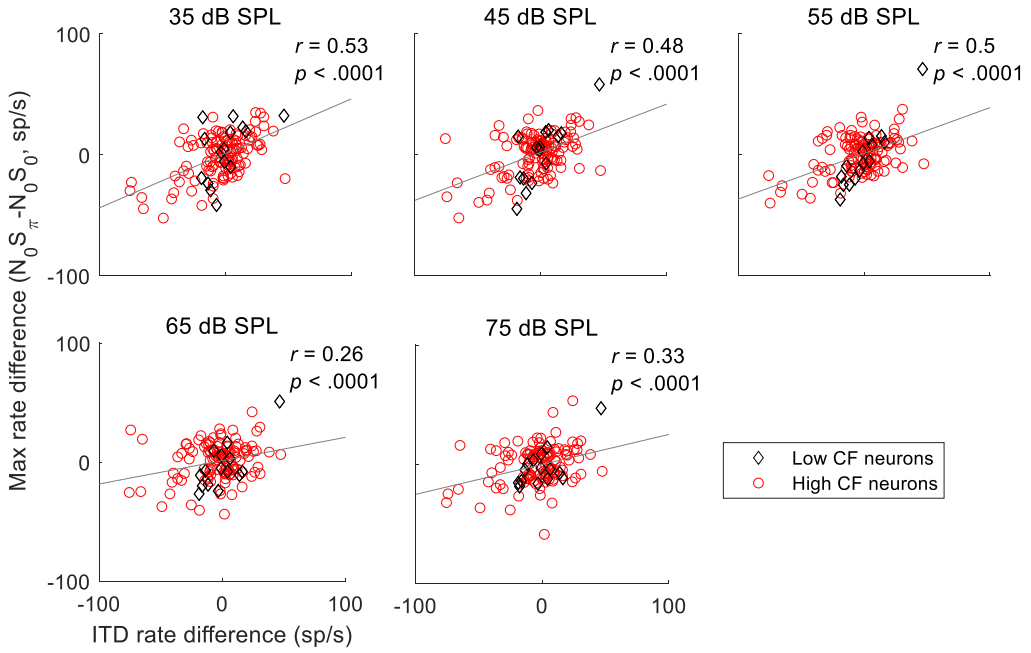

Figure 3.8: Correlation between maximum rate differences of binaural TIN responses (diotic and dichotic) and NDF at each noise levels. The correlation coefficient and p-value are shown in the plot. The maximum rate difference between responses to diotic and dichotic condition was significantly correlated with the maximum rate difference of NDF, supporting the ITD-based hypothesis.

Even though the results above supported the ITD-based hypothesis, they leave a few unanswered questions. First, in previous studies, the NDF shape was examined for responses to N_0S_π stimuli either at a low tone frequency (e.g., 500 Hz, Jiang et al., 1997) or for a broadb frequency range (e.g., broadband chirp, Lane & Delgutte, 2005), neither of which examined the role of stimulus frequency. In addition, ITD sensitivity is assumed to be important for sound localization at low frequencies (Middlebrooks & Green, 1991). Therefore, whether NDF shape *only* explains responses to low-frequency N_0S_π stimuli is of interest. Second, it is also important to examine whether ITD-based predictions are more successful when the relation between the best ITD and the stimulus frequency is

taken into account. Third, adding an out-of-phase tone to diotic noise not only introduces ITDs, but also ILDs, yet ILD has not been considered previously in physiological studies of N_0S_π responses. These questions will be addressed briefly in the following sections.

Low- vs. high-CF neurons. To study the role of frequency in the ITD-based results, neurons with CFs below 1.5 kHz were re-examined (Fig. 3.9). Due to the small number of units, the χ^2 test was not applicable, so a qualitative description is presented here. For neurons with peak-like NDFs, all low-CF neurons had increasing rate upon addition of a tone at threshold at low noise levels for the N_0S_0 condition, and most neurons had increasing rate for the N_0S_π condition. For neurons with trough-like NDFs, most neurons had increasing rate upon addition of either a diotic or dichotic tone at threshold. The ITD-based prediction explained the rate-change direction for almost all N_0S_π responses in low-CF neurons with trough-like NDFs (e.g., 3/3 at 65 dB SPL), as compare to explaining only 19/28 of the rate-change directions at 65 dB SPL for the entire population (Fig. 3.7, middle bottom plot).

Data from Fig. 3.7 were re-analyze for units with CF > 1.5 kHz (not shown), the proportions of rate-change direction for neurons with peak- or trough-like NDF were still significantly different at noise levels of 35-65 dB SPL ($\chi^2(1,1) = 17.11, 16.12, 8.38, 6.17, 0.91;\ p < 0.001, p < 0.001,\ p = 0.0038, p = 0.013, p = 0.34$). There was a decreasing trend in the proportion of units with rate-change directions that were explained by the NDF shape with increasing noise level. Therefore, NDF shape best explained neural responses of both low- and high-CF units at low and medium noise levels.

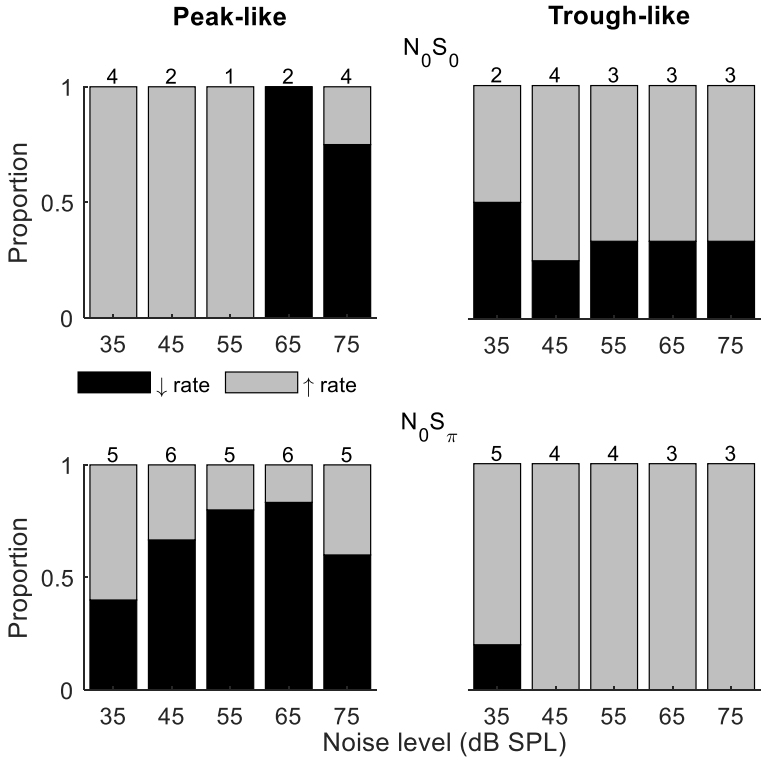

Figure 3.9: Proportion of rate-change direction for low-CF neurons with peak-like (left) and trough-like (right) ITD sensitivity for the N_0S_0 (top) and N_0S_π (bottom) conditions. Only neurons with CF lower than 1.5 kHz and measurable thresholds are shown here. Format same as in Figs. 3.6 and 3.7.

The role of best ITD. One concern for predicting neural responses to N_0S_π stimuli is related to the potential interaction between tone frequency and the best/worst ITD. For example, for a neuron with a peak-type NDF having a best ITD of 200 µs, a dichotic tone at 2.5 kHz (with a period of 400 µs) would result in an increase in rate, rather than to the "expected" decrease in rate for a peak-type NDF. To account for this potential interaction, the correlation was computed between the change in rate in the NDF for ITDs corresponding to interaural phase differences (IPDs) of 0° and 180° at the tone frequency and the change in rate elicited by addition of a dichotic tone at 8-dB SNR. Note that an

IPD of 180° may result in either a positive or a negative ITD; in this test, the sign of the ITD (positive or negative) was chosen to be consistent with the sign of the best/worst ITD. The correlation was not significant ($r = 0.04, p = 0.71$), indicating that the rate-change direction in TIN responses could not be explained by the direction of change in rate between responses to ITDs corresponding to 0° to 180° delays of the tone frequency. Considering that many stimulus frequencies, and CFs, were beyond the limit of fine-structure phase-locking, neurons with low CFs (less than 1.5 kHz) were examined separately, but that correlation was also not significant ($r = 0.21, p = 0.53$, Fig. S1). Possible reasons for this result will be discussed below.

Rate differences in response to N_0S_π stimuli and binaural cues. The rate differences in response to ITDs or ILDs was quantified by the difference between the maximum and minimum response rates over the range of stimuli tested. The maximum change in rate in response to N_0S_π stimuli, regardless of the direction of the rate change as a function of SNR, was significantly correlated to the maximum rate differences in both ITD and ILD responses (Fig. 3.10, A and B), explaining a small but significant proportion of the variance (i.e. r^2). The significant correlation between the maximum rate differences for N_0S_π responses and rate differences for both ITD and ILD responses could be because 1) adding a dichotic tone not only introduces ITDs, but also ILDs; and/or 2) the dynamic ranges of ITD and ILD responses were significantly correlated (Fig. 3.10C). Changes in neural responses to N_0S_π could be due to a combination of ITD and ILD sensitivities; this effect is a potential topic for future studies.

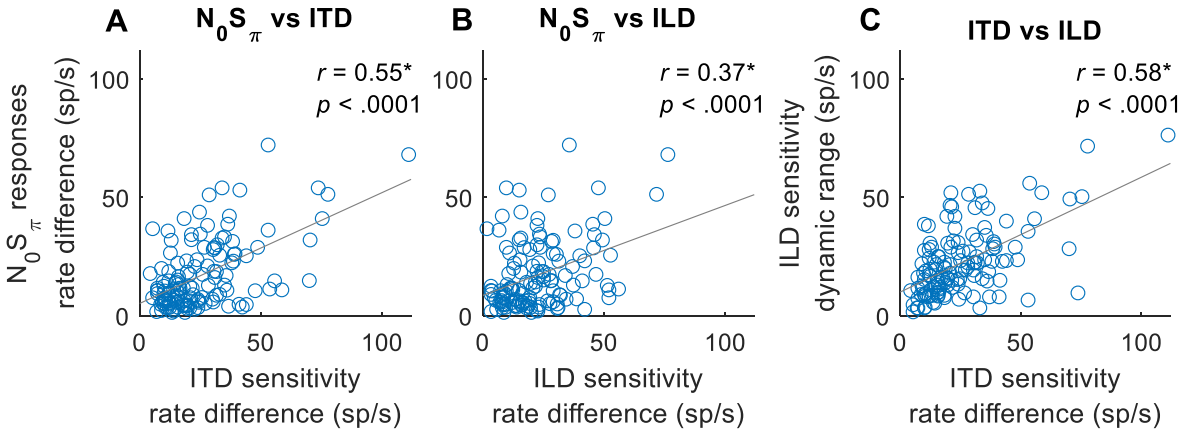

Figure 3.10: Correlation between rate differences of responses to ITD, ILD, and N_0S_π at 65 dB SPL (as title indicated). Correlation coefficients and *p*-value are shown on the top right of each panel; star indicates that the correlation coefficient was significant after Bonferroni correction ($p < 0.017$). Solid gray line indicates linear regression.

Evaluating the IAC-based hypothesis for the population of IC responses

Adding a dichotic tone to diotic noise introduces both ITD and ILD cues, as well as interaural decorrelation, but the changes in these cues differ for different tokens of noise waveforms as well as for different SNRs. For example, the ITD of a N_0S_π stimulus would become dominated by the ITD of the added tone with increasing tone level, but the effective ITD of a N_0S_π stimulus with a low-SNR tone (e.g. at threshold) is hard to estimate, and varies with the noise token due to the phase interaction between the noise and tone. Additionally, unlike a pure tone, the instantaneous ITD of N_0S_π stimuli vary throughout the duration of the stimulus waveform, and thus these stimuli do not have an "overall" ITD. Therefore, prediction of the rate-change direction upon addition of a tone at threshold based on ITD (or ILD) sensitivity may not be as simple as has been previously hypothesized. On the other hand, the effect of interaural decorrelation could be studied with a more straightforward method. To examine the effect of decorrelation, average rates

were recorded in response to 1/3-octave diotic (N_0) and uncorrelated (N_u) gaussian noise for 68 neurons. The N_u noises presented at the two ears were simply independent narrowband noise tokens. The correlation between the difference in average rate in response to the N_0S_π condition (the difference between average rates in response to noise-alone and to N_0S_π at 0-dB SNR) and the difference in average rates in response to the N_0 and N_u conditions was significant at all noise levels (Fig. 3.11), supporting the IAC-based hypothesis. The correlation was the strongest at a noise level of 65 dB SPL, the level at which the N_0 and N_u noise were presented. At 65 dB SPL, additional analyses of the rate differences in responses to N_0S_π stimuli at SNRs of -8 to 8 dB relative to the noise-alone condition were all significantly correlated to the rate difference between the responses to N_u and N_0 noise, with correlation coefficients ranging from 0.71 to 0.84, and *p* values all less than 0.0001 (significant after Bonferroni correction, not shown). The significant correlation coefficients at *all* SNRs and noise levels indicated that, in general, the direction and size of the changes in rate in response to N_0S_π stimuli were explained by the change in the stimulus from N_0 towards N_u.

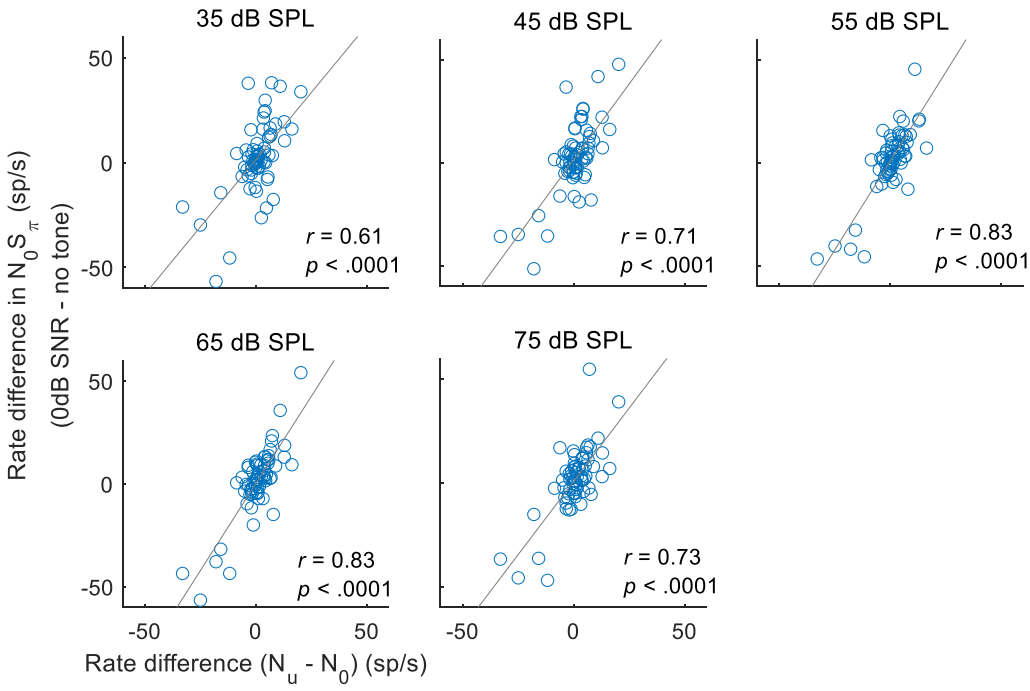

Figure 3.11: Correlation between rate differences elicited by addition of dichotic tone (N_0S_π) at 0 dB SNR and difference in rate between responses to N_0 and N_u condition. Correlation coefficient and *p*-value are shown on the top left; star indicates that the correlation coefficient was significant after Bonferroni correction ($p < 0.0014$). Solid line indicates linear regression.

Rate-based neural thresholds

Rate-based thresholds of all units for the N_0S_0 and N_0S_π conditions at five noise levels were computed and compared with behavioral data from previous studies (Fig. 3.12). There was no clear trend in the numbers of units with increasing or decreasing rate-change direction across frequency, for either the N_0S_0 or N_0S_π condition, except a weak trend of more units with increasing rate at the lowest noise level tested (bottom row). The lowest rate thresholds across frequency were lower for the N_0S_π condition than for the N_0S_0 condition, as expected.

The lowest rate thresholds at 500 Hz matched the mean rabbit behavioral detection threshold at the same frequency (Zheng et al., 2002). Compared with human thresholds, the lowest rate thresholds for the N_0S_0 condition were close to human thresholds across frequencies, but the lowest rate thresholds for the N_0S_π condition only matched human thresholds at high frequencies. Human thresholds from Goupell (2012) are slightly lower than van de Par et al. (1999) at some frequencies, possibly due to differences in paradigm and stimulus bandwidths. Note that stimuli used in previous studies have slightly different parameters from this study: stimuli in Zheng et al. (2002) had 200-Hz bandwidth (vs. 116 Hz in this study) and an overall level of 63 dB SPL; stimuli in van de Par et a. (1999) had bandwidths of 100 Hz, 250 Hz, 500 Hz and 1 kHz (vs. 116 Hz, 232 Hz, 463 Hz and 926 Hz in this study) for center frequencies of 500 Hz, 1 kHz, 2 kHz and 4 kHz, and with overall level of 70 dB SPL; stimuli in Buss et al. (2003) had 50-Hz bandwidth and overall noise levels of 42, 57 and 72 dB SPL; stimuli in Goupell (2012) had bandwidths of 78 Hz, 240 Hz, 456 Hz and 888 Hz (vs. 116 Hz, 463 Hz, 926 Hz and 1852 Hz in this study) for center frequencies of 500 Hz, 2 kHz, 4 kHz and 8 kHz. However, despite the discrepancies among stimuli, in general, the lowest rate-based thresholds could explain human thresholds for the N_0S_0 condition across all frequencies tested and for the N_0S_π condition at high frequencies. Note that the thresholds of most sensitive neurons across frequencies did not vary qualitatively across noise levels, consistent with human thresholds tested at multiple noise levels (Buss et al., 2003) and with roving-level paradigm (Henning et al., 2005).

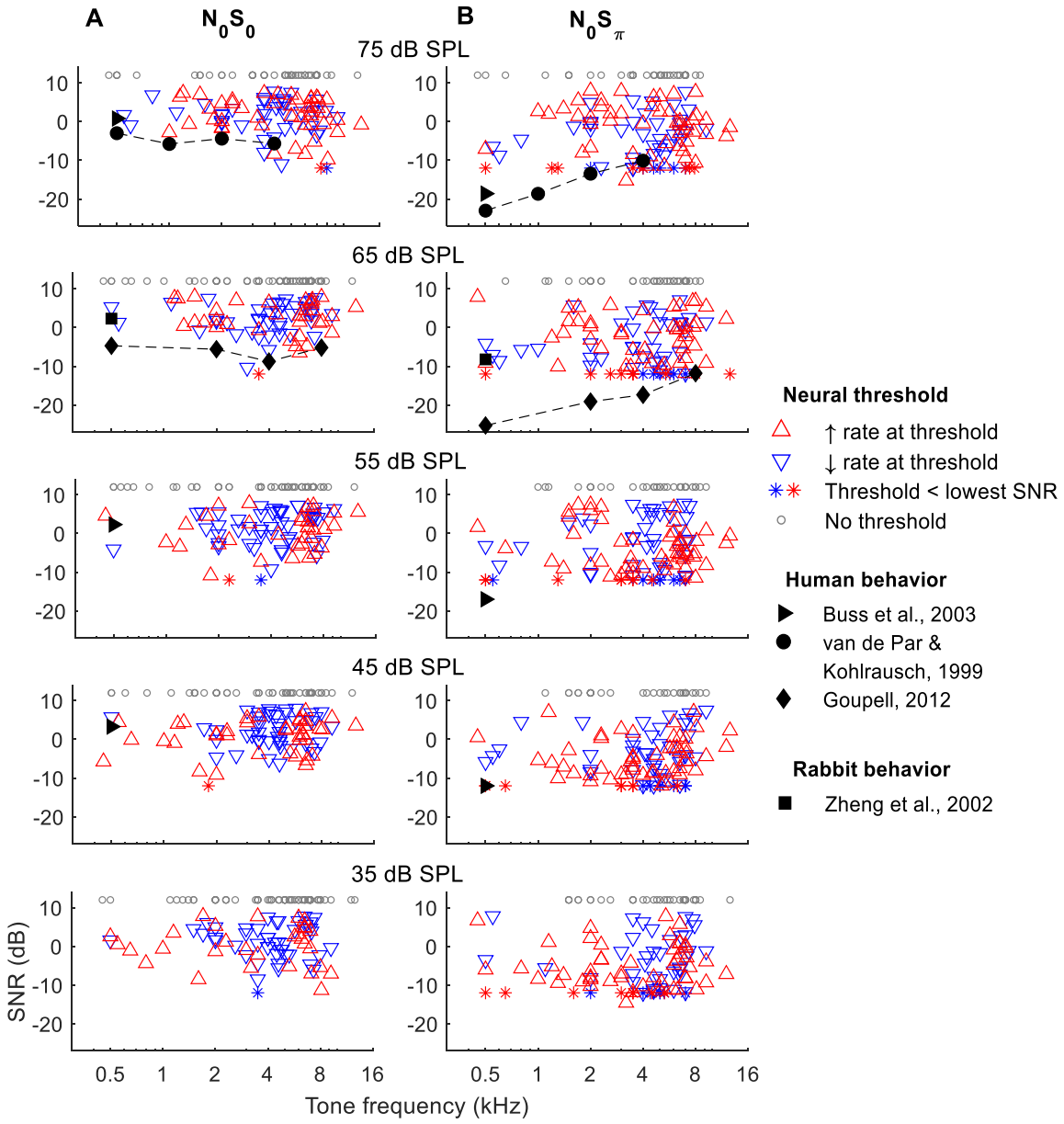

Figure 12: Rate-based threshold for N_0S_0 (A) and N_0S_π (B) conditions. Thresholds of most sensitive neurons across frequencies matched human behavioral data for the N_0S_0 condition but had a trend different from human for the N_0S_π condition. Neural thresholds at 500 Hz matched rabbit behavioral data for both conditions.

Because envelope-, ITD-, and IAC-based cues were all shown to explain TIN responses, it is interesting to ask whether these characteristics could predict neural

thresholds. For envelope-based cues, there was no significant correlation between thresholds and rate differences elicited by amplitude modulation (calculated as in Fig. 7) ($r = -0.072, p = 0.54$). For ITD-based cues, neurons that have more sensitive ITD tuning were expected have lower N_0S_π thresholds. For example, one would expect that a neuron would have a lower N_0S_π detection threshold if the response rate changed more quickly as a function of ITD. This hypothesis was examined in ITD-sensitive neurons with NDFs that were fitted with the G_2 function as high-CF neurons were the major population (87% of sensitive neurons). The level of sensitivity to ITD tuning was quantified by A/σ (the amplitude of the fitted function divided by the standard deviation of the gaussian function). However, the correlation was not significant between the level of sensitivity of ITD tuning and the N_0S_π threshold at a noise level of 65 dB SPL ($r = 0.14, p = 0.31$). For IAC-based cues, there was no significant correlation between the N_0S_π threshold at noise level of 65 dB SPL and the rate difference between responses to N_0 and N_u conditions (calculated as in Fig. 11) ($r = -0.058, p = 0.71$). Unexpectedly, none of the cues were directly related to neural thresholds.

Rate-based neural BMLDs

Neural BMLDs were evaluated in two ways: using the BMLDs of individual neurons and using the BMLDs calculated from the N_0S_0 and N_0S_π thresholds of the neural population. For BMLDs of single neurons (Fig. 3.13), only neurons with measurable thresholds for both N_0S_0 and N_0S_π conditions are plotted, together with human BMLDs (Buss et al., 2003; Goupell, 2012; van de Par & Kohlrausch, 1999). There was no clear association observed between small or negative BMLDs and rate-change direction for either N_0S_0 or N_0S_π conditions, in contrast to a previous report (Jiang et al., 1997a). There was also no clear pattern of same (open symbols) or opposite (filled symbols) rate-change

directions for N_0S_0 and N_0S_π conditions across frequency (i.e. thresholds were similar for upward and downward triangles). Overall, there were more neurons with the same rate-change directions than with opposite rate-change directions (more open symbols than filled symbols) between N_0S_0 and N_0S_π conditions. Among neurons with opposite rate-change directions across conditions, more neurons had decreasing rate at threshold for the N_0S_0 condition (more filled downward than upward triangles). At 500 Hz, single-neuron BMLDs were close to human BMLDs at noise levels of 45 and 65 dB SPL, but not at other noise levels. At 1 kHz and above, the maximum single-neuron BMLDs were larger than human BMLDs. The maximum BMLDs were similar across noise levels, as well as across frequencies, unlike human BMLDs that decrease substantially with increasing frequency (Goupell, 2012; van de Par & Kohlrausch, 1999).

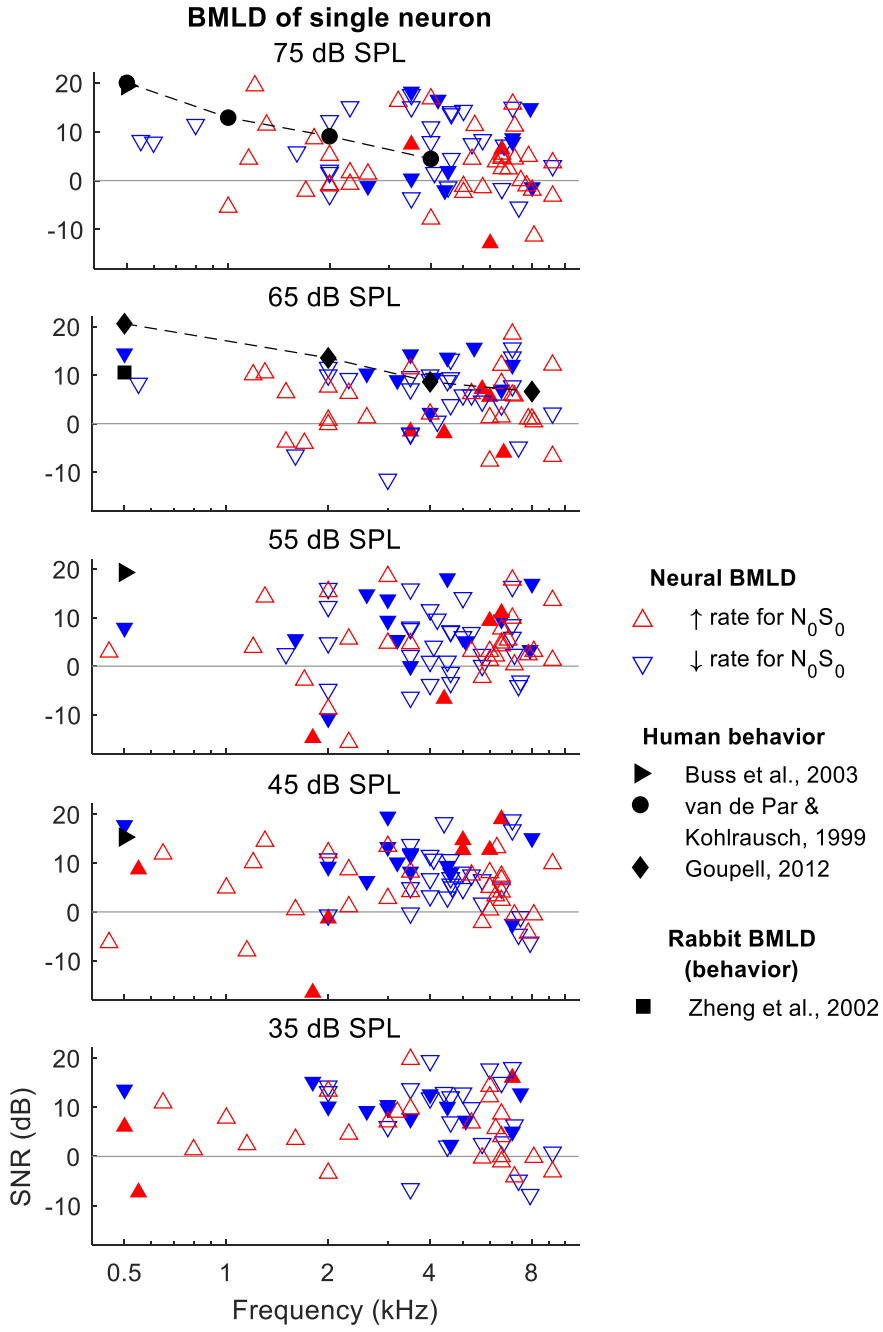

Figure 3.13: BMLDs calculated based on single-neuron thresholds for neurons with measurable thresholds for both N_0S_0 and N_0S_π conditions. Open triangles indicate that the direction of change in rate vs. SNR at threshold for the N_0S_π condition was the *same* as for the N_0S_0 condition, whereas filled triangles indicate *opposite* direction of change in rate at threshold for the N_0S_0 and N_0S_π conditions.

To calculate BMLDs of the neural population, neural thresholds for the most sensitive subset of neurons were calculated for 0.5, 1, 2, 4, and 8 kHz for the N_0S_0 or N_0S_π conditions. For each frequency, the threshold was based on the lowest 10^{th} percentile within a one-octave range centered at that frequency. Due to the limited SNR range that was tested, many sensitive neurons were suprathreshold (70.7% correct) at the lowest tested SNR, especially for the N_0S_π condition. To reduce the number of neurons for which the BMLD was limited in this way, individual thresholds were recalculated using a criterion of 79.1% correct (Fig. 3.14, squares and diamonds). Thresholds at 55 to 75 dB SPL had similar patterns and were plotted together in Fig. 14, which shows that neural population thresholds for both N_0S_0 (blue solid line) and N_0S_π conditions (red dashed line) did not vary across frequency. Human thresholds were moved up by 4 dB to align the means of the human and N_0S_0 thresholds of the population, to better compare the trend across frequency (Fig. 14). Human N_0S_π thresholds increase as a function of frequency, whereas thresholds of the neural population did not. Therefore, human and neural BMLDs had different trends across frequency: human BMLDs decrease with increasing frequency, whereas neural BMLDs did not. The neural BMLDs of the population were smaller than the maximum single-neuron BMLDs but had a similar trend across frequency.

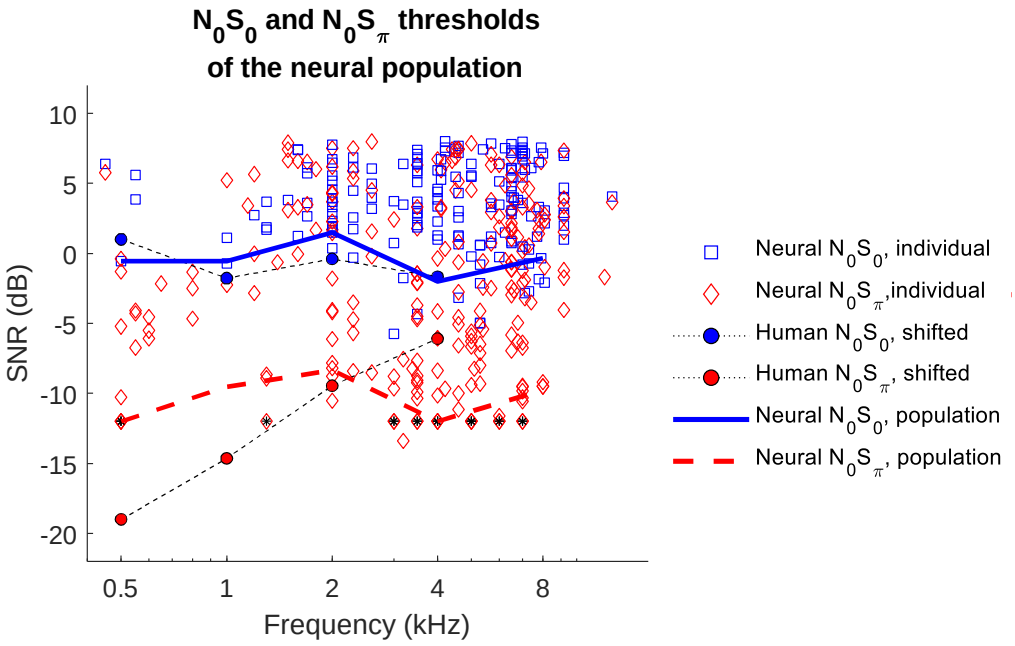

Figure 3.14: N_0S_0 (solid blue line) and N_0S_π thresholds (dashed red line) of the neural population across frequency. Individual neural thresholds at **79.1%** correct for N_0S_0 (blue square) and N_0S_π (red diamond) conditions, for noise levels of 55-75 dB SPL are shown. Star indicates that the threshold was lower than the lowest measured SNR. Human detection thresholds are from van de Par et al. (1999) and shifted up by 4 dB for comparison with neural thresholds, which were computed using a higher criterion. Neural BMLDs had a different trend across frequency compare to human BMLDs.

3.6 Discussion

In the current study, single-neuron responses to TIN stimuli were recorded in the IC for both N_0S_0 and N_0S_π conditions over a wide range of target frequencies, as well as noise and tone levels. Envelope-, ITD-, and IAC-based hypotheses for TIN detection were tested. Adding a tone flattens the envelope of a narrowband gaussian noise waveform, peripheral nonlinearities effectively increase this flattening, and IC neurons are tuned to amplitude modulations, as described by MTFs. Therefore, for neurons that are excited by

neural fluctuations (BE MTFs), average rates were expected to decrease upon addition of a tone, whereas for neurons that are suppressed by fluctuations (BS MTFs), average rates were expected to increase, regardless of whether the tone was in or out of phase. On the other hand, for the N_0S_π condition, adding an out-of-phase tone to a diotic noise waveform introduces time-varying ITDs, so neurons with peak-like NDFs were expected to have decreasing rates, whereas neurons with trough-like NDFs were expected to have increasing rates as a function of SNR. For the N_0S_π condition, adding an out-of-phase tone to a diotic noise waveform also causes interaural decorrelation. Therefore, neurons that had higher average rates in response to N_u than to N_0 noise were expected to have higher average rates for the tone-plus-dichotic-tone condition than for the noise-alone condition.

Results showed that for the N_0S_0 condition, the envelope-based hypothesis explained neural responses upon addition of a tone at most noise levels, but the results for the N_0S_π condition were predicted by envelope-based cues only at the highest noise level tested (75 dB SPL). Additionally, for the N_0S_π condition, the ITD-based hypothesis better explained neural response rates elicited by an out-of-phase tone at the low-to-medium noise levels than at the higher noise level. Changes in rate due to interaural decorrelation were significantly correlated with changes in rate upon addition of an out-of-phase tone to identical noise at all noise levels. Because predictions of rate-change direction based on MTF shape for the N_0S_0 condition were discussed in Chapter 2, the current Discussion will focus on the results for the N_0S_π condition and BMLDs.

Comparison with previous physiological studies

There have been a limited number of physiological studies of neural responses to both N_0S_0 stimuli and N_0S_π stimuli, especially in the IC (Jiang et al., 1997a, 1997b; Lane

& Delgutte, 2005). Lane and Delgutte (2005) used chirp as detection target, and Jiang et al. (1997b) also used tone as detection target, so the results here were compared with Jiang et al. (1997a). There were a few differences between the stimuli used in the current study and in Jiang et al. (1997a), which may explain the differences in the results between the two studies. First, responses were only recorded for tone frequency of 500 Hz in Jiang et al. (1997a). Even though only neurons with low CF (<1.5 kHz) were recorded, the best frequency of a neuron with CF of 1 kHz is an octave above 500 Hz, and responses to 500-Hz tones could be essentially different from responses to CF tones. For example, the response of a model auditory-nerve fiber (Zilany et al., 2014) is saturated in response to a CF tone at 65 dB SPL, but not in response to a 65-dB-SPL tone one octave below CF. Therefore, when the tone frequency is far from CF, the AN fiber's rate will vary with stimulus sound level (i.e. sound level). The difference between CF and target tone frequency could explain the finding that the majority of neurons in Jiang et al. (1997a) had increasing rate with increasing tone level for the N_0S_0 condition, whereas many neurons had decreasing rate versus SNR in the current study.

Second, many neurons in the current study did not have measurable thresholds due to the limited range of SNRs tested, but finer steps and a wider range of SNRs were used in Jiang et al. (1997), so thresholds were measurable for almost all neurons. However, it is worth noting that a 20-dB range of SNRs were tested in this study, neurons without a measurable threshold over this SNR range were insensitive to addition of a tone. A threshold might have been measured if a SNR high enough had been tested, but it could also be that the neuron was simply responding to high-intensity sound, not to tone-plus-noise.

Third, the noise masker bandwidth in Jiang et al. (1997a) was from 50 Hz to 5 kHz with a noise level of 65 dB SPL, whereas the current study used 1/3-octave noise centered at the tone frequency and a wide range of noise levels, including 65 dB SPL. The difference in noise masker bandwidth represents a large difference in noise spectrum level: 28 dB SPL in Jiang et al. (1997a), versus 44 dB SPL for the 500-Hz target tone tested at the noise level of 65 dB SPL in the current study. This difference may stimulate neurons differently, especially at low stimulus frequencies. Even though neurons respond to a wide frequency range at high sound levels (Ruggero, 1992), the excitation frequency range is usually asymmetric and spreads more to lower frequencies than to higher frequencies (Schmiedt, 1989). Therefore, for low CF neurons (e.g., CF of 1 kHz), the wide frequency range of the noise masker used in Jiang et al. (1997a) (e.g., the frequency range of noise masker above 2 kHz) may not have effectively excited the neurons. Additionally, due to the cochlear compression, using a high overall level but wide noise bandwidth may not mask tones in the same way as a narrow noise bandwidth but the same spectrum level (Almishaal, Bidelman, & Jennings, 2017).

The role of the MTF in N_0S_π responses. The shape of the MTF has been shown to explain the direction of rate change upon addition of an in-phase tone to diotic noise (N_0S_0) (Chapter 2). For the N_0S_π condition, predictions of rate changes upon addition of the dichotic tone based on rate differences in the MTF were successful at all noise levels (Fig. 3.5), whereas predictions of rate-change direction based on MTF shape were only successful at the highest noise level tested, 75 dB SPL). This discrepancy may be due to the different emphases of the two analysis methods used to explore this hypothesis: the categorization method emphasized the number of neurons that had the expected rate-change direction near threshold, whereas the correlation method emphasized the size of

the rate change. Therefore, although the number of neurons with the expected rate-change direction based on the MTF shape was lower than expected, but for that the neurons that did respond as expected, the changes in rate were correlated between MTF and N_0S_π responses. However, overall the results suggested that envelope-based cues were most important at the higher noise levels.

The role of the NDF in N_0S_π responses. The NDF was shown to explain rate responses to N_0S_π stimuli at most noise levels, but both the proportion of neurons with the expected rate-change directions, and the amount of variance explained in the rate changes decreased with increasing noise level (Figs. 3.7 and 3.8). The prediction based on the NDF was reexamined based on each neuron's rates in response to ITDs corresponding to IPDs of 0° and 180° at the tone frequency. One reason for this re-analysis is that if a half period of the tone frequency was smaller than a neuron's best ITD, the prediction of the N_0S_π responses would differ from that based on the NDF shape alone. Specifically, as the SNR varies from a noise-alone stimulus to a high-level-tone in noise, the IPD of an N_0S_π stimulus changes from 0° to nearly 180°. Therefore, the rate difference between responses to IPDs equivalent to 0° and 180° (or -180° if the best ITD were negative) was hypothesized to be correlated with the rate difference elicited by addition of an out-of-phase tone with the highest SNR tested in N_0S_π stimuli. However, no such correlation was found, even for neurons with low CFs. There were several possible explanations for this result: First, the instantaneous ITD of the N_0S_π stimulus varies across the duration of the stimulus due to the interaction between the tone and narrowband noise, but only the ITD corresponding to the 180° phase at the tone frequency was used in this test. Additionally, the NDF used in this study was in response to wideband noise, but the noise masker for the N_0S_π condition was narrowband noise. Nevertheless, this result

suggested that predicting neural responses, or even just the rate-change direction, did not depend solely on the shape of the NDF.

The role of IAC in N_0S_π responses. Adding an out-of-phase tone reduces the IAC (e.g., Bernstein & Trahiotis, 2017). However, whereas IAC cues have been one focus of modeling studies, they have not previously been tested in physiological studies. As predicted by the IAC-based hypothesis, the change in rate elicited by an out-of-phase tone was significantly correlated with the rate difference between responses to N_0 and N_u noise (Fig. 3.11); the large proportion of variance explained (50% to 71%) suggested an important role of the IAC in physiological N_0S_π responses.

Proposed relationship between envelope, ITD, and IAC cues. Results showed that envelope-, ITD-, and IAC-based cues all explained a proportion of neural responses to N_0S_π stimuli (maximum 23%, 28%, and 71%). The ITD-based hypothesis explained neural responses at low-to-medium noise levels, whereas the envelope-based hypothesis explained at high noise level; the IAC-based hypothesis, on the other hand, explained neural responses at all noise levels and the largest proportion of variance in rate responses among the three cues at 65 dB SPL, at which N_0 and N_u noise responses were collected. However, these cues may not be independent. For example, the decreasing trend in the proportion of results explained by ITD-based hypothesis could be that envelope ITDs dominated responses of the high CF neurons, which were the majority neurons in the population studied here, but the fluctuation amplitudes in AN responses saturate (i.e. flatten) at higher sound levels , and thus binaural differences of neural representation of the stimulus envelope would also decrease with increasing sound level, which would explain a weaker effect of envelope ITDs at high levels. Also, at high

frequencies, IAC-cues have been proposed to be envelope-based (Bernstein & Trahiotis, 1996; Durlach, 1964).

Some effort has been made to separate the role of IAC and ITD in binaural detection (Culling, 2011; van der Heijden & Joris, 2010), but the debate is not resolved. Based on results from these studies, both ITD and ILD cues are proposed to contribute to interaural decorrelation. Adding an out-of-phase tone not only introduces ITDs, but also ILDs; additionally, the added binaural cues are time-varying. However, there has been little discussion about the role of ILD responses in predicting N_0S_π responses. One reason could be that there is no clear prediction of a change in rate based on ILD sensitivity, as both positive and negative ILDs are elicited by the out-of-phase tone. Yet ILD could be important – the dynamic range of ILD responses was correlated not only to that of N_0S_π responses, but also to the dynamic range of ITD responses (Fig. 3.10). Fluctuations of ITD in an N_0S_π stimulus increase with increasing tone level, whereas fluctuations of ILD first increase and then decrease as tone level increases (Appendix A). Therefore, interaural decorrelation could be due to a nonlinear combination of ITD and ILDs cues: both ITD and ILD cues affect IAC at low tone levels, whereas at high tone levels (e.g., above 4 dB SNR), ITD cues dominate IAC. This proposed idea is consistent with the previous modeling study (Mao & Carney, 2014) in which ITD cues are shown to dominate in stimuli with low modulation depths (e.g., tone-plus-noise), and the combination of ITD and ILD cues dominate in stimuli with high modulation depths (e.g., noise). In that study, the nonlinear combination of ITD and ILD cues is described as the slope of the interaural envelope difference (SIED), whereas detection in the N_0S_π condition at high frequencies has been proposed to be explained by the envelope-based IAC (Bernstein & Trahiotis,

1996; Durlach, 1964). Thus, the SIED cue is hypothesized to be a specific implementation for an envelope-based IAC in explaining N_0S_π responses.

Neural BMLDs vs. human BMLDs. Rate-based thresholds were estimated for both N_0S_0 and N_0S_π conditions in order to estimate neural BMLDs over a range of frequencies and noise levels. For the N_0S_0 condition, the lowest rate-based thresholds across frequency could explain human detection thresholds. For the N_0S_π condition, the lowest rate-based thresholds across frequency had a different trend from human detection thresholds: neural thresholds were higher (i.e. worse) than human thresholds at low frequencies, and lower (i.e. better) than human thresholds at high frequencies. Many neurons had BMLDs as large as 20 dB. BMLDs estimated based on the most sensitive units in the neural population and estimates of maximum BMLDs for single neurons only varied slightly across frequency, whereas human BMLDs decrease substantially with increasing frequency. BMLDs estimated for the neural population were shown to be slightly lower than maximum single-neuron BMLDs across all frequencies, because individual neurons with the lowest thresholds in either the N_0S_0 or N_0S_π condition did not always have the lowest thresholds in the other condition.

Rate-based neural thresholds were similar across noise levels, consistent with human psychophysical studies (Buss et al., 2003). Human BMLDs have been shown barely affected by the roving-level paradigm, in which stimulus levels randomly vary from interval to interval (Henning et al., 2005). Similar patterns of rate-based neural BMLDs across noise levels could explain the level-resistant of human listeners.

Future directions. The current study showed that envelope-, ITD-, and IAC-based hypotheses all explained a proportion of neural responses to N_0S_π stimuli. However, there

are many questions left to be answered in future studies. For example, what cues affect a neuron's threshold, as opposed to response rates? Neural threshold was not correlated with any of the tested characteristics, even for those that were expected to be most highly related, such as sensitivity to ITDs. A better understanding of how the characteristics of the NDF in high-CF neurons are related to N_0S_π responses also requires further study. High-CF neurons have ITD sensitivity (Yin et al., 1984), but the factors that affect different aspects of the NDF in high-CF neurons (e.g., prominence of peak or trough, best ITD) have not been studied. These factors may play an important role in understanding N_0S_π responses at high frequencies. Another direction that could be taken in the future is to understand SIED cues and their potential relationship with IAC for the N_0S_π condition. The SIED cues have been shown to better predict human N_0S_π detection thresholds than ITD or ILD alone, and SIED is a nonlinear combination of ITD and ILD (Mao & Carney, 2014). Many IC neurons in this study were shown to be sensitive to the SIED cue (Appendix B). Therefore, the SIED cue could be a potential direction for future physiological studies. IC neurons receive afferent inputs from almost all lower levels in the auditory pathway (Casseday et al., 2002), so the combination of several types of neural sensitivity to different features of sounds may be required to explain neural responses to complex sound stimuli.

Appendix

A *Binaural cues as a function of SNR*

To better understand neural responses as a function of SNR for the N_0S_π condition, IPD and ILD fluctuation cues were computed for stimuli used in this study (Fig. 3.15) with the same calculation as in Goupell and Harmann (2006). Briefly, instantaneous IPDs and

ILDs were calculated; fluctuation refers to the standard deviation of the instantaneous cues. IPD fluctuation increases with increasing SNR, whereas ILD fluctuation changes non-monotonically. Note that these cues were stimulus-based, and there was no substantial difference observed related to tone frequency.

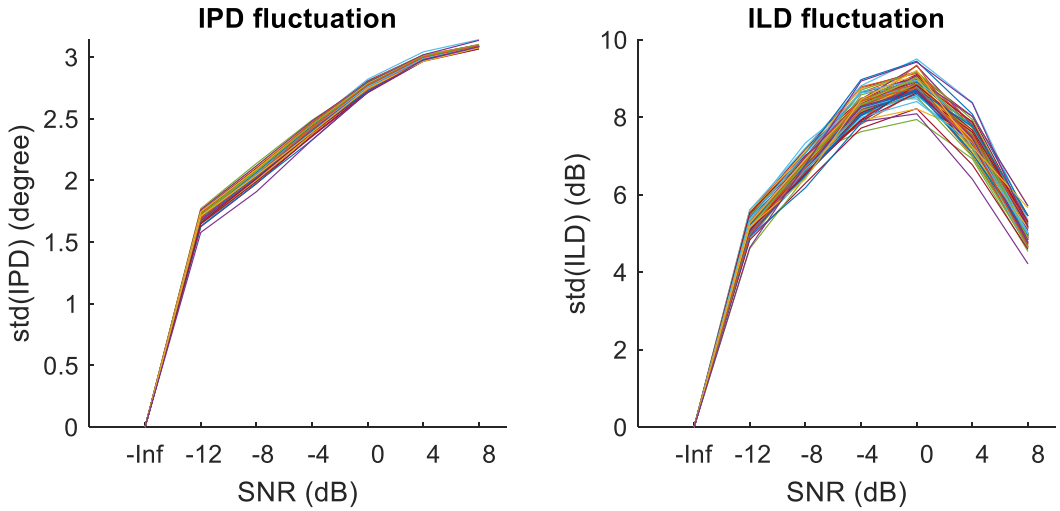

Figure 3.15: IPD (left) and ILD fluctuation (right) of stimuli used in this study. Calculation was the same as in Goupell and Harmann (2006). Note that the trend of IPD and ILD fluctuation as a function of SNR did not vary qualitatively with tone frequencies.

B *Sensitivity to the SIED*

Tuning to the SIED was determined by a whitened reverse-correlation based on the SIED of dichotic stimuli consisting of uncorrelated narrowband noises at the two ears (N_u stimuli). Reverse correlation (revcor), also known as a spike-triggered average or first-order Wiener kernel, has been used to estimate a neuron's linear receptive field (e.g., Lewis et al., 2002) based on responses to white gaussian noise. However, the SIEDs of the stimuli were not white gaussian noise, and thus resulted in a biased revcor. The biased

revcor was “whitened” by an automatic smoothness determination (ASD) process implemented in Aoi and Pillow (2017). A baseline revcor was estimated based on randomized spike timing. Figure 3.16 shows the SIED revcor of an example neuron. Most IC neurons recorded in the current study had a substantial SIED revcor functions, suggesting that they are tuned to this stimulus feature.

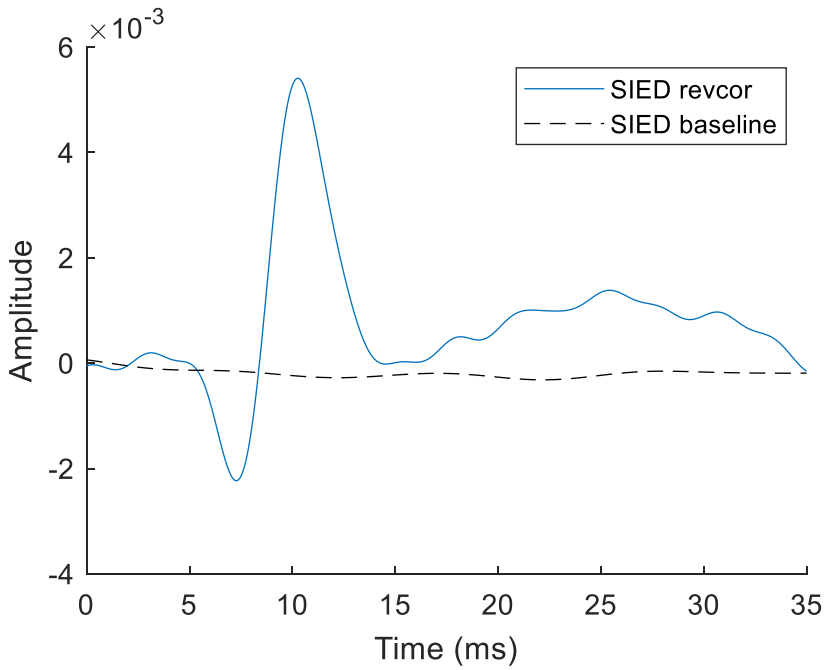

Figure 3.16: Reverse correlation (revcor) based on SIED (solid blue) and baseline revcor (dashed black) computed using shuffled spike times, as an estimate for the revcor functions noise floor. Spike times were at time 0, thus the x-axis labels indicate time preceding the spikes.

Bibliography

Almishaal, A., Bidelman, G. M., & Jennings, S. G. (2017). Notched-noise precursors improve detection of low-frequency amplitude modulation. *J Acoust Soc Am, 141*(1), 324. doi:10.1121/1.4973912

Bernstein, L. R., & Trahiotis, C. (1996). On the use of the normalized correlation as an index of interaural envelope correlation. *J Acoust Soc Am, 100*(3), 1754-1763. doi:10.1121/1.416072

Bernstein, L. R., & Trahiotis, C. (1997). The effects of randomizing values of interaural disparities on binaural detection and on discrimination of interaural correlation. *J Acoust Soc Am, 102*(2 Pt 1), 1113-1120. doi:10.1121/1.419863

Bernstein, L. R., & Trahiotis, C. (2017). An interaural-correlation-based approach that accounts for a wide variety of binaural detection data. *Journal of the Acoustical Society of America, 141*(2), 1150-1160. doi:10.1121/1.4976098

Buss, E., Hall, J. W., 3rd, & Grose, J. H. (2003). The masking level difference for signals placed in masker envelope minima and maxima. *J Acoust Soc Am, 114*(3), 1557-1564. doi:10.1121/1.1598199

Cant, N. B., & Oliver, D. L. (2018). Overview of Auditory Projection Pathways and Intrinsic Microcircuits. *Mammalian Auditory Pathways: Synaptic Organization and Microcircuits, 65*, 7-39. doi:10.1007/978-3-319-71798-2_2

Carney, L. H. (2018). Supra-threshold hearing and fluctuation profiles: Implications for sensorineural and hidden hearing loss. *Journal of the Association for Research in Otolaryngology*, *19*(4), 331-352.

Casseday, J. H., Fremouw, T., & Covey, E. (2002). The inferior colliculus: a hub for the central auditory system. In *Integrative functions in the mammalian auditory pathway* (pp. 238-318): Springer.

Colburn, H. S. (1977). Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. *J Acoust Soc Am, 61*(2), 525-533. doi:10.1121/1.381294

Culling, J. F. (2011). Subcomponent cues in binaural unmasking. *J Acoust Soc Am, 129*(6), 3846-3855. doi:10.1121/1.3560944

Durlach, N. I. (1964). Note on Binaural Masking-Level Differences at High Frequencies. *Journal of the Acoustical Society of America, 36*(3), 576-&. doi:Doi 10.1121/1.1919006

Egan, J. P. (1975). *Signal detection theory and ROC-analysis*. Academic press.

Hirsh, I. J. (1948). The influence of interaural phase on interaural summation and inhibition. *The Journal of the Acoustical Society of America*, *20*(4), 536-544.

Goupell, M. J. (2012). The role of envelope statistics in detecting changes in interaural correlation. *J Acoust Soc Am, 132*(3), 1561-1572. doi:10.1121/1.4740498

Goupell, M. J., & Hartmann, W. M. (2006). Interaural fluctuations and the detection of interaural incoherence: bandwidth effects. *J Acoust Soc Am, 119*(6), 3971-3986. doi:10.1121/1.2200147

Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J Acoust Soc Am, 115*(2), 833-843. doi:10.1121/1.1639908

Jiang, D., McAlpine, D., & Palmer, A. R. (1997a). Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions. *Journal of Neurophysiology, 77*(6), 3085-3106. Retrieved from <Go to ISI>://WOS:A1997XJ17700019

Jiang, D., McAlpine, D., & Palmer, A. R. (1997b). Detectability index measures of binaural masking level difference across populations of inferior colliculus neurons. *Journal of Neuroscience*, *17*(23), 9331-9339.

Lane, C. C., & Delgutte, B. (2005). Neural correlates and mechanisms of spatial release from masking: single-unit and population responses in the inferior colliculus. *Journal of Neurophysiology, 94*(2), 1180-1198. doi:10.1152/jn.01112.2004

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. doi:10.1152/jn.1988.60.6.1799

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J Acoust Soc Am, 49*(2), Suppl 2:467+. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/5541744

Liu, L. F., Palmer, A. R., & Wallace, M. N. (2006). Phase-locked responses to pure tones in the inferior colliculus. *Journal of Neurophysiology, 95*(3), 1926-1935. doi:10.1152/jn.00497.2005

Mao, J., & Carney, L. H. (2014). Binaural detection with narrowband and wideband reproducible noise maskers. IV. Models using interaural time, level, and envelope differences. *J Acoust Soc Am, 135*(2), 824-837. doi:10.1121/1.4861848

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annu Rev Psychol, 42*, 135-159. doi:10.1146/annurev.ps.42.020191.001031

Nelson, P. C., & Carney, L. H. (2007). Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *Journal of neurophysiology*, *97*(1), 522-539.

Ramachandran, R., Davis, K. A., & May, B. J. (2000). Rate representation of tones in noise in the inferior colliculus of decerebrate cats. *Journal of the Association for Research in Otolaryngology*, *1*(2), 144-160.

Richards, V. M. (1992). The detectability of a tone added to narrow bands of equal-energy noise. *The Journal of the Acoustical Society of America*, *91*(6), 3424-3435.

Ruggero, M. A. (1992). Responses to sound of the basilar membrane of the mammalian cochlea. *Curr Opin Neurobiol, 2*(4), 449-456. doi:10.1016/0959-4388(92)90179-o

Schmiedt, R. A. (1989). Spontaneous rates, thresholds and tuning of auditory-nerve fibers in the gerbil: comparisons to cat data. *Hear Res, 42*(1), 23-35. doi:10.1016/0378-5955(89)90115-9

van de Par, S., & Kohlrausch, A. (1999). Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters. *J Acoust Soc Am, 106*(4 Pt 1), 1940-1947. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10530018

van der Heijden, M., & Joris, P. X. (2010). Interaural correlation fails to account for detection in a classic binaural task: dynamic ITDs dominate N0Spi detection. *J Assoc Res Otolaryngol, 11*(1), 113-131. doi:10.1007/s10162-009-0185-8

Wagner, H., Takahashi, T., & Konishi, M. (1987). Representation of interaural time difference in the central nucleus of the barn owl's inferior colliculus. *Journal of Neuroscience, 7*(10), 3105-3116. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/3668618

Yin, T. C., Kuwada, S., & Sujaku, Y. (1984). Interaural time sensitivity of high-frequency neurons in the inferior colliculus. *J Acoust Soc Am, 76*(5), 1401-1410. doi:10.1121/1.391457

Zheng, L., Early, S. J., Mason, C. R., Idrobo, F., Harrison, J. M., & Carney, L. H. (2002). Binaural detection with narrowband and wideband reproducible noise maskers: II. Results for rabbit. *Journal of the Acoustical Society of America, 111*(1), 346-356. doi:Doi 10.1121/1.1423930

Zheng, Y., & Escabi, M. A. (2013). Proportional spike-timing precision and firing reliability underlie efficient temporal processing of periodicity and envelope shape cues. *Journal of Neurophysiology, 110*(3), 587-606. doi:10.1152/jn.01080.2010

Chapter 4. Summary and Discussion

The goal of this thesis is to understand how well stimulus envelopes and neural fluctuations can predict the direction of change in neural response rate upon addition of a tone to noise in the mammalian inferior colliculus (IC). The results show that an envelope-based hypothesis could predict the rate-change direction in most neurons with measurable thresholds for the N_0S_0 condition and at high noise levels for the N_0S_π condition. Additionally, the results show that other characteristics were important to understand responses to N_0S_π stimuli: the shape of responses to interaural time differences (ITDs), noise delay function (NDF), could predict the rate-change direction in most neurons with measurable thresholds for the N_0S_π condition at low to medium noise levels; rate changes due to changes in interaural correlation (IAC) could explain a large proportion of variance in rate changes elicited upon addition of an out-of-phase tone to identical noise.

4.1 Summary of novel results

Narrowband gaussian noise (GN) has a fluctuating envelope and the envelope becomes flatter upon addition of a tone. On the contrary, low-noise noise (LNN) has a flat envelope and the envelope becomes more fluctuating upon addition of a low-level tone. Therefore, predictions for an envelope-based hypothesis are opposite for these two types of noise masker, and thus make them a nice pair to examine the envelope-based hypothesis. Results showed that for most neurons with measurable thresholds, the direction of the change in average rate could be predicted by the MTF shape for both GN and LNN maskers.

An envelope-based hypothesis could explain the rate-change direction at threshold for many neurons for the N_0S_0 condition, a natural extension of the study is to examine if this envelope-based hypothesis could explain the direction of changes in rate at threshold for the N_0S_π condition as well. Results show that the envelope-based hypothesis could explain the rate-change direction at threshold of most neurons at high levels, and of most neurons with high characteristic frequency (CF) at most noise levels. A neuron's NDF shape has been proposed to explain the rate-change direction in responses to N_0S_π stimuli, so the ITD-based hypothesis was also examined. Results show that at low noise levels, classical ITD sensitivity could prediction neural responses at low noise levels, but at the highest noise level tested, prediction of the envelope-based hypothesis outperformed prediction of ITD-based hypothesis. Even though NDF has been assumed to be the physiological basis of IAC, the current study directly evaluated the effect of interaural decorrelation due to addition of an out-of-phase tone. Results show that correlation between rate difference between N_0 and N_u noise (decrease in IAC) and rate changes upon addition of an out-of-phase tone was significantly correlated, and the correlation coefficient was higher than the correlation coefficient between dynamic ranges of NDF and responses to N_0S_π stimuli.

4.2 Limitations of current studies

This thesis provided physiological evidence for the hypothesis that IC neurons respond to TIN stimuli based on changes in the envelope. However, because these are the first studies in the effort to extend the neural coding of envelope to non-periodic stimuli, there were some limitations in this study that could be pursued in the future.

One limitation of this study is the identification of MTF shape. In this thesis, sinusoidal-amplitude-modulated (SAM) sounds with noise carriers were used to obtain the MTFs. The shape of MTF was identified in order to predict the direction of the change in response rate of tone-plus-GN and tone-plus-LNN. However, our preliminary data show that different carriers may change the shape of MTF, and the difference between carriers may vary from neuron to neuron. For SAM sounds with pure tone carriers, the neuron's CF has often been used for the frequency of the tone carrier (e.g., Nelson & Carney, 2007; Rose & Capranica, 1985). However, it is sometimes difficult to identify the CF based on the complex response maps that IC neurons have. These aspects of MTFs were not studied.

In this thesis, to understand the effect of sound level, neural responses to TIN stimuli were recorded for a wide range of noise levels. However, due to the limited daily recording time (2 hr), only a few SNRs were selected to estimate neural thresholds. Even though the selected SNRs had a range of over 20 dB, for some sensitive neurons, the lowest SNR tested was higher than the threshold. For some insensitive neurons, this SNR range was still not wide enough to estimate the threshold. On the other hand, if these neurons did not have measurable thresholds over 20-dB range of SNR, they could simply be insensitive to addition of a tone.

Envelope-related cues have not been previously considered for the N_0S_π condition. This study examined the envelope-based hypothesis. Results show that a neuron's MTF shape could explain the direction of change in rate with increasing SNR. The envelope cue examined was monaural cue, and thus binaural processing was not involved. The types of ITD sensitivity have been used to explain neural responses to N_0S_π stimuli (e.g., Jiang, McAlpine, & Palmer, 1997; van der Heijden & Joris, 2010), and thus were examined

in this study. Results showed that the types of ITD sensitivity explained the direction of changes in rate versus SNR, consistent with the hypothesis. Additionally, results showed that the rate difference between responses to diotic and dichotic noise could better predict the rate difference between responses to noise-alone stimuli and N_0S_π stimuli than the dynamic range of ITD responses. There are two possible reasons that caused this difference. First, adding an out-of-phase tone introduces both ITDs and interaural levels differences (ILDs), but ILD sensitivity has not been previously considered. One solution may be to combine ITD and ILD sensitivities, but this is beyond the scope of the current study. Second, the dynamic range of ITD responses was calculated over the entire range of ITDs tested, but may be wider than the range of ITDs in N_0S_π stimuli. Some effort was done to match the range of ITDs in N_0S_π stimuli, but majority high-CF neurons in the study do not phase lock to the fine structure of high frequency stimuli. Thus, ITDs calculated for N_0S_π stimuli based on fine-structure would not be accurate, but there is also no previous knowledge about how to calculate the proper ITDs for N_0S_π stimuli for high-CF neurons. The role of IAC was also examined in this study and was shown to be important in understanding responses to N_0S_π stimuli. However, the role of IAC could be further explored by examining rate changes elicited with different level of IAC between N_u noises at the two ears, and how these rate changes correlate with responses to N_0S_π stimuli.

4.3 Envelope-based modeling approach

Modeling is an important aspect in hearing research. A good model can not only simulate neural responses or human performance for one stimulus, but also helps to predict responses or performance for other stimuli. Dau, Puschel, and Kohlrausch (1996) developed an envelope-based model, but this model was stimulus-based and unable to

model responses of single neurons. Carney, Li, and McDonough (2015) presented a physiologically-realistic model, including stages of auditory nerve, cochlear nucleus, and the inferior colliculus (IC). The IC was represented by the same-frequency inhibition and excitation (SFIE) model, which can simulate an IC neuron with MTF. The AN-SFIE model is an implementation of the envelope-encoding strategy that was mentioned in the Introduction section: the envelope is preserved in the neural fluctuations of high-spontaneous-rate AN fiber, enhanced in the cochlear nucleus, and reflected in responses rate in the IC. The envelope-based hypothesis was tested in single-neuron responses, so it is interesting to test if the AN-SFIE model, which captures the sensitivity of IC neurons to envelope, can simulate IC neuron responses to other stimuli such as TIN. This work was done in addition to the thesis work and is appended after this chapter (Fan, Henry, & Carney, 2018; see appendix). Briefly, excitatory and inhibitory frequencies and the corresponding response latencies of each neuron were extracted from the second-order Wiener kernel that was estimated from responses to white noise. These parameters were used to fit the AN-SFIE model to predict responses to sinusoidal-amplitude-modulated (SAM) sounds, tone-in-GN stimuli, and white noise. Results show that this model could simulate neural rate responses to SAM sounds with both noise and tone carriers and tone-in-GN stimuli in many neurons, and the second-order Wiener kernel that was estimated from model responses was also similar to the kernel estimated from neural responses. The importance of this work is that envelope-based model can not only simulate responses to one type of stimulus, but also to other stimuli. These results suggested that if the characteristics of a neuron can be simulated with a model, this neuron's response to other stimuli that have not been recorded with is likely to be predicted by the model.

Additionally, modeling responses will help predict and understand population responses, as simultaneous single-neuron recordings for a large population of neurons are difficult.

4.4 Slope of interaural envelope differences

In the N_0S_π condition, ITD sensitivity was shown to explain the direction of change in rate upon addition of a tone to a noise masker, but the maximum rate change upon addition of a dichotic tone was less correlated to ITD dynamic range than the rate difference due to interaural decorrelation. Thus, neural responses to N_0S_π stimuli could be more than what ITD sensitivity predicts. Because adding a dichotic tone to noise introduces both ITDs and ILDs, combination of ITD and ILD sensitivity could improve the prediction. Although there has not been modeling study focusing on combining ITD and ILD cues in TIN studies, a previous modeling study for N_0S_π stimuli showed that the slope of interaural envelope differences (SIED) is a nonlinear combination of ITD and ILD (Mao & Carney, 2014). The SIED is a cue based on binaural envelopes. As the name indicated, the SIED takes the difference between the slopes of stimulus envelope at the two ears. Preliminary results showed that many IC neurons are sensitive to SIED, but further studies are still needed to better understand this cue. If SIED could explain neural responses better than ITD, neural responses to both N_0S_0 and N_0S_π stimuli in the IC may be explained by a concise model based on envelopes, which would be consistent with IC neural sensitivity to envelope that are not present at earlier auditory processing levels.

Bibliography

Carney, L. H., Li, T., & McDonough, J. M. (2015). Speech Coding in the Brain: Representation of Vowel Formants by Midbrain Neurons Tuned to Sound Fluctuations. *eNeuro, 2*(4). doi:10.1523/ENEURO.0004-15.2015

Dau, T., Puschel, D., & Kohlrausch, A. (1996). A quantitative model of the "effective" signal processing in the auditory system. I. Model structure. *J Acoust Soc Am, 99*(6), 3615-3622. doi:10.1121/1.414959

Fan, L. C., Henry, K. S., & Carney, L. H. (2018). Challenging One Model With Many Stimuli: Simulating Responses in the Inferior Colliculus. *Acta Acustica United with Acustica, 104*(5), 895-899. doi:10.3813/Aaa.919249

Jiang, D., McAlpine, D., & Palmer, A. R. (1997). Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions. *Journal of Neurophysiology, 77*(6), 3085-3106.

Mao, J., & Carney, L. H. (2014). Binaural detection with narrowband and wideband reproducible noise maskers. IV. Models using interaural time, level, and envelope differences. *J Acoust Soc Am, 135*(2), 824-837. doi:10.1121/1.4861848

Nelson, P. C., & Carney, L. H. (2007). Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *Journal of Neurophysiology, 97*(1), 522-539. doi:10.1152/jn.00776.2006

Rose, G. J., & Capranica, R. R. (1985). Sensitivity to amplitude modulated sounds in the anuran auditory nervous system. *Journal of Neurophysiology, 53*(2), 446-465. doi:10.1152/jn.1985.53.2.446

van der Heijden, M., & Joris, P. X. (2010). Interaural correlation fails to account for detection in a classic binaural task: dynamic ITDs dominate N_0S_{pi} detection. *J Assoc Res Otolaryngol, 11*(1), 113-131. doi:10.1007/s10162-009-0185-8

Appendix: Challenging one model with many stimuli: simulating responses in the inferior colliculus

Langchen Fan[1)], Kenneth S. Henry[2)3)], Laurel H. Carney[1)3)]

[1)] Department of Biomedical Engineering,

[2)] Department of Otolaryngology,

[3)] Department of Neuroscience, University of Rochester

(Reprinted from Acta Acustica united with Acustica, issue 104, volume 5, page 895-899)

A.1 Abstract

Existing models to explain human psychophysics or neural responses are typically designed for a specific stimulus type and often fail for other stimuli. The ultimate goal for a neural model is to simulate responses to many stimuli, which may provide better insights into neural mechanisms. We tested the ability of modified same-frequency inhibition-excitation models for inferior colliculus neurons to simulate individual neuron responses to both amplitude-modulated sounds and tone-in-noise stimuli. Modifications to the model were guided by receptive fields computed with 2nd-order Wiener kernel analysis. This approach successfully simulated many individual neurons' responses to different types of stimuli. Other neurons suggest limitations and future directions for modeling efforts.

A.2 Introduction

The inferior colliculus (IC) is a critical center in the auditory system – all ascending pathways converge in the IC, en route to the thalamus and cortex (Cant & Oliver, 2018). Neurons in the IC are rate-tuned to sinusoidally amplitude-modulated (AM) sounds, as described by modulation transfer functions (MTFs, discharge rate vs. AM frequency). IC neurons with band-enhanced (BE) MTFs have increased rate when their neural inputs contain fluctuations near the best modulation frequency (BMF); those with band-suppressed (BS) MTFs have decreased rate for a range of fluctuation frequencies.

Adding a tone to narrowband Gaussian noise flattens the stimulus envelope (Richards & Nekrich, 1993), as reflected in peripheral responses (Carney, 2018). Therefore, in response to tone-in-noise (TIN) stimuli, band-enhanced IC neurons are expected to have decreasing response rate with addition of a tone, and band-suppressed IC neurons are expected to have increasing rate. These predicted changes in rate in response to TIN stimuli based on MTF types are consistent with neural responses of many BE and BS neurons, but not all (Fan & Carney, 2017). Computational models may provide insights into these complex neural responses.

Many existing models have been developed to explain human psychophysical results and responses of different auditory neurons. These models are usually based on a single stimulus type; however, in general the same neurons respond to many different stimulus types. Attempting to simulate neural responses to different stimulus types with one model will better explain how single neurons function and provide better population simulations in the future. This study was an initial step in this direction. The same-frequency inhibition-excitation (SFIE) model has previously been shown to simulate IC

responses to AM sounds at different modulation frequencies for both BE and BS MTF shapes (Carney, Li, & McDonough, 2015; Nelson & Carney, 2007). We generalized the SFIE model by varying the frequency tuning of neural inputs, input delays, and the number of inputs to match the response characteristics of single neurons. These modifications were guided by the neuron's receptive field obtained from the Fourier transform of the 2nd-order Wiener kernel. Each neuron-specific model was then tested for its ability to simulate the neuron's responses to AM stimuli, TIN stimuli, and white noise. For this initial study, the inputs to the model were provided by an AN model for cat (Zilany, Bruce, & Carney, 2014), as a detailed model for the rabbit periphery is not available.

A.3 Methods

A.3.1 Physiological methods

Extracellular recordings were made with tetrodes in the IC of awake Dutch-belted rabbits and sorted offline to isolate single units. The current study focused on twenty neurons for which the pattern of excitation and inhibition in the receptive field (RF) could be readily used to specify the model structure and parameters (see below).

Detailed physiological methods are described in Carney et al. (Carney, Zilany, Huang, Abrams, & Idrobo, 2014). Briefly, neural responses to pure tones from 250 Hz to 20 kHz were recorded to determine characteristic frequency (CF). Neural responses to sinusoidal AM noise and tones, TIN stimuli, and long-duration white noises were recorded. AM tones had carrier frequencies at CF, modulation frequencies from 2 to 350 Hz (for low CFs) or 1024 Hz (for high CF neurons) and were presented at 70 dB SPL. AM wideband noise (0.1-10 kHz) had modulation frequencies from 2 to 350 Hz and was presented at a

spectrum level of 30 dB SPL. All AM stimuli had 1-sec durations. TIN stimuli had a tone near CF and 1/3-octave bands of noise centered near CF, with 0.3-sec durations. Overall noise level for TIN stimuli varied from 35 to 75 dB SPL in 10-dB steps; tones were presented at signal-to-noise ratios (SNRs) ranging from -12 dB to 8 dB. White noises (0.1 – 20kHz) were 2-sec in duration, presented at 65 dB SPL.

Average discharge rates were computed for responses to AM sounds and TIN stimuli. Second-order Wiener kernels were calculated by multiplying the instantaneous spike rates and the outer product of time-reversed pre-spike stimulus segments [9] (Figure 1). The 1-D Fourier transform of the kernel provides an estimate of the neuron's RF. Excitation and inhibition in the RF can be identified with singular value decomposition (Lewis, Henry, & Yamada, 2002). Smoothing and peak-finding algorithms were applied to determine the center frequency and latency of the excitation and inhibition. Second-order Wiener-kernel derived RFs are an alternative to spectrotemporal RFs (Eggermont, Aertsen, & Johannesma, 1983; Escabi & Schreiner, 2002).

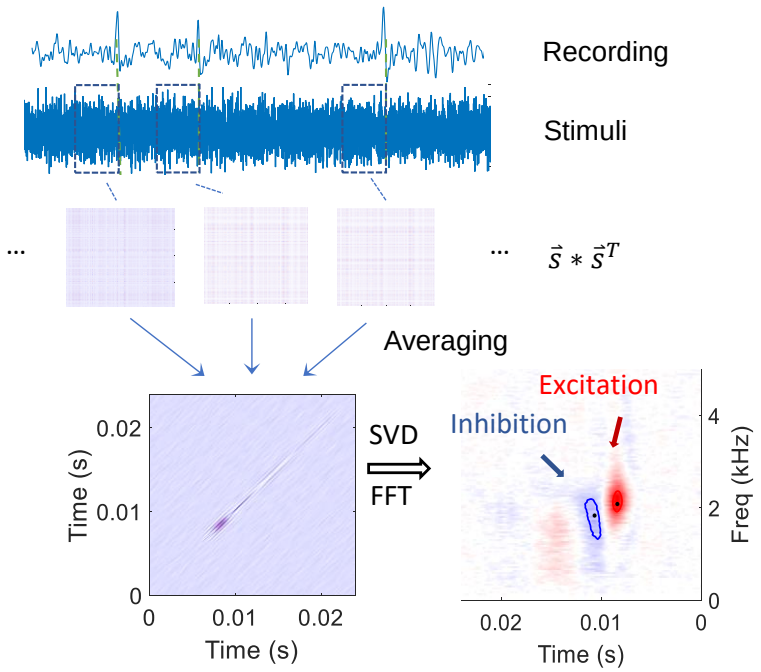

Figure A.1: Calculation of the receptive field (RF) using the 2nd-order Wiener kernel. The kernel was calculated by averaging the product of instantaneous spike rates with the outer product of pre-spike stimulus epochs. A 1-D Fourier transform yielded the RF. Singular-value decomposition (SVD) was used to identify excitation and inhibition in the RF (Lewis et al., 2002).

A.3.2 Modeling methods

The original SFIE model is comprised of a model of an auditory-nerve (AN) fiber (Zilany et al., 2014), brainstem (cochlear nucleus, CN) and IC neurons (Carney et al., 2015; Nelson & Carney, 2004). The CN model always receives excitatory and delayed inhibitory inputs from a single AN model and, therefore, has the same CF as the AN model. The IC neurons in the SFIE model have one excitatory and one delayed inhibitory input. Here, we generalized the SFIE model by removing restrictions on the CFs and numbers of excitatory and inhibitory inputs (Fig. 2). Relative timing between the first strong excitation frequency

(f_{exc1}) and other excitation or inhibition frequencies (f_{exc2}, f_{inh1} and in some cases f_{inh2}) were identified from the RF.

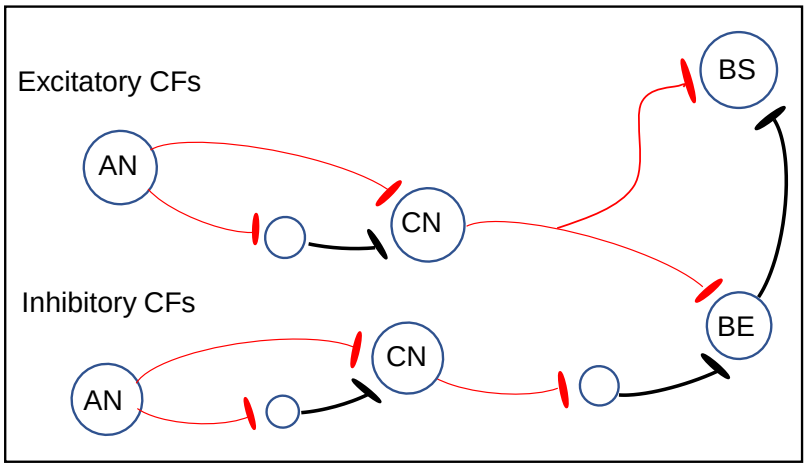

Figure A.2: Illustration of generalized SFIE model. The auditory nerve (AN) model provides excitatory (red, thin line) and inhibitory (black, thick line, via an interneuron) inputs to the cochlear nucleus (CN) model. The CFs and number of excitatory and inhibitory inputs to the inferior colliculus (IC) are not limited, but for simplicity, this diagram includes only one of each. The band-enhanced (BE) IC neuron receives excitatory and inhibitory inputs from CN; the band-suppressed (BS) IC neuron receives excitatory input(s) from the CN and is inhibited by a BE neuron

Latencies based on the RF were used for all modified IE models. Four variations of the model were considered: a) one excitatory and one inhibitory input with the same CF; b) one excitatory and one inhibitory input with different CFs, based on the RF; c) a second inhibitory input, with CF matched to the excitatory input, was added to b); and d) up to two excitatory and two inhibitory inputs with CFs selected based on the RF. Responses were also simulated with the original SFIE model as a reference. The strength of the inhibitory (str_{inh}) and second excitatory (str_{exc2}, when present) inputs to the IC were adjusted relative to the strongest excitatory input (str_{exc}) based on the number of excitatory and inhibitory

inputs: for modifications a) and b), str_{inh} = 1.3; for c), str_{inh} = 0.8 (both inhibitory inputs); for d), str_{exc} = 1, str_{inh} = 1.3 when two excitatory and inhibitory inputs were used; str_{exc1} = 1, str_{exc2} = 0.8, str_{inh} = 2.2 for two excitatory and inhibitory inputs; str_{inh} = 0.8 for one excitatory and two inhibitory inputs. Note that once the parameters for a given model were selected, the model was fixed to simulate responses to all stimulus types.

White noises used for simulations had the same statistics and parameters as in the neural recordings. AM sounds and TIN stimuli for simulations were identical to those used for recordings, except that sound levels of all stimuli used as model inputs were reduced by 10 dB, because the models were slightly more sensitive than the neurons. Model performance was evaluated using Pearson correlations between model and neural MTFs and TIN responses. For TIN, correlation coefficients were calculated for each overall noise level. Internal noise was contributed by the random variations in the AN model responses; no additional internal noise was added.

A.4 Results

Figure 3 a-d shows an IC neuron with a band-enhanced MTF. The neuron's TIN responses decreased with increasing SNRs, consistent with the prediction based on MTF type. Model results (modification b) are shown in Figure 3 e-h. Excitatory and inhibitory frequencies used in the model were 2083 Hz and 1833 Hz, respectively; the inhibitory input was delayed by 2.3 ms. Simulations for both MTFs and TIN stimuli followed the trends in neural responses. The correlation between MTF data and simulation was significant for noise carrier (r = 0.68, p <0.001) and tone carrier (r = 0.19, p = 0.19). For TIN stimuli, correlation coefficients were calculated for each noise level. Correlation coefficients for five overall noise levels (from low to high) were: 0.82, 0.96, 0.90, 0.51, 0.65;

p values were 0.01, <0.001, 0.04, 0.12, 0.058, respectively. Correlations were significant for all datasets except the TIN responses at the two highest noise levels. The model receptive field (RF) had excitatory and inhibitory regions that were similar to the neural RF (Figure 3 d and h).

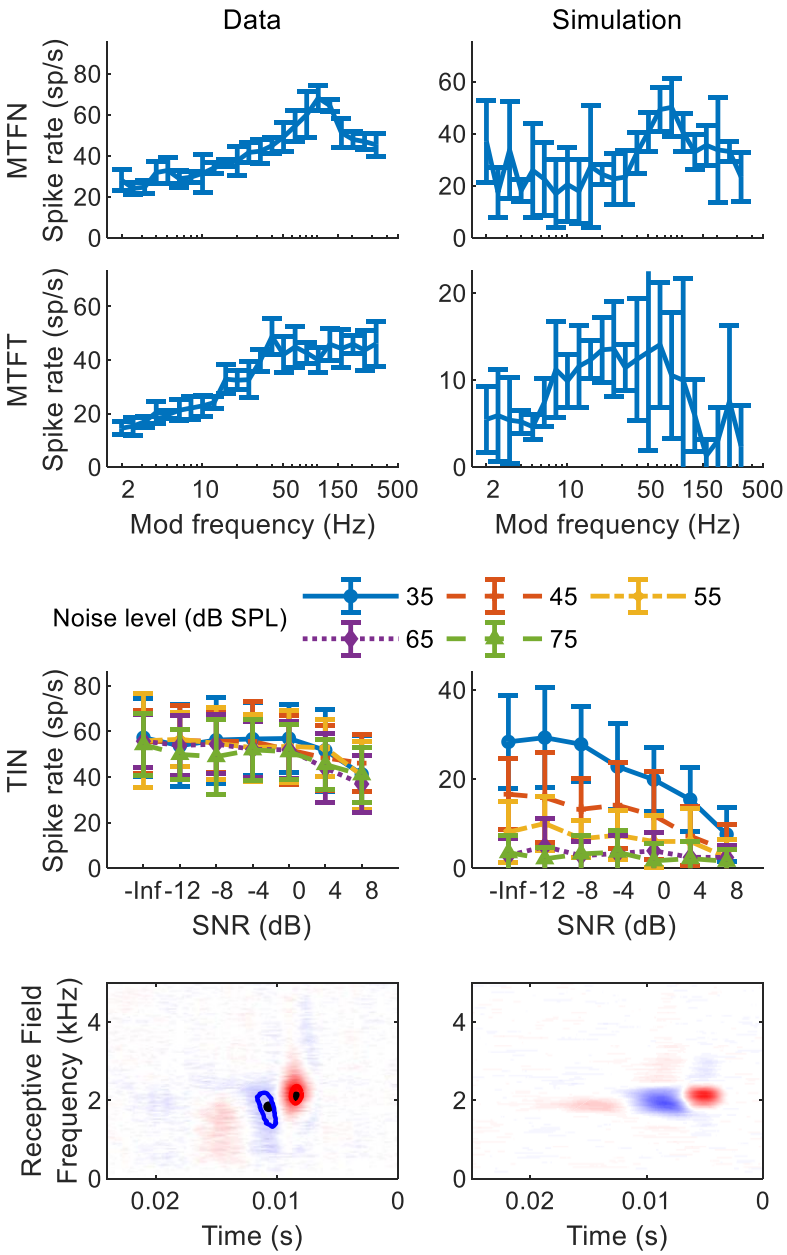

Figure A.3: Example of a band-enhanced IC neuron. MTFs (rate vs. modulation frequency) for a) noise and b) tone carriers; errorbars show standard deviation. c) Neural responses to TIN stimuli (rate vs. SNR) for several noise levels. d) Receptive field (time reversed) calculated with the 2nd-order Wiener-kernel analysis – the center of excitation and inhibition used in modified models are marked by black dots; e) to h): simulations with modified IE model for the same stimuli as in a) - d).

Figure 4 shows an IC neuron with band-suppressed MTF for both noise and tone carriers (a and b). The neuron's TIN response rates increased with increasing SNRs, consistent with the prediction for this MTF type. Model results (modification c, one excitation and two inhibitions) are shown in Figure 4 e-h. The excitation CF was 800 Hz; one inhibition at 800 Hz and one at 960 Hz were used. Both inhibitory delays were 4 ms. Simulations of this neuron's MTFs successfully replicated both the shape and the lowest point in the band-suppressed MTF. The correlations between MTF data and simulations were significant for both noise carrier (r = 0.77, p < 0.001) and tone carrier (r = 0.82, p <0.001). For TIN stimuli, correlation coefficients for five overall noise levels (from low to high) were: -0.37, 0.91, 0.84, 0.46, 0.96; p values were 0.80, 0.002, 0.008, 0.15, <0.001, respectively. The model RF had similar excitatory and inhibitory frequencies as the neural RF, but missed some details. For example, the frequency range of excitation and inhibition were much wider in the neural RF than in the model RF. The simulated 2nd-order Wiener kernel had inhibition at frequencies higher than the excitation, possibly due to high-frequency suppression in the model AN responses. In this example neuron, the simulated kernel also had excitation at frequencies higher than the inhibition. This pattern was possibly related to the model structure: band-suppressed MTFs were the result of inhibition from a band-enhanced neuron (Fig. 2); therefore, inhibition of the BE neuron facilitated responses of the BS neuron.

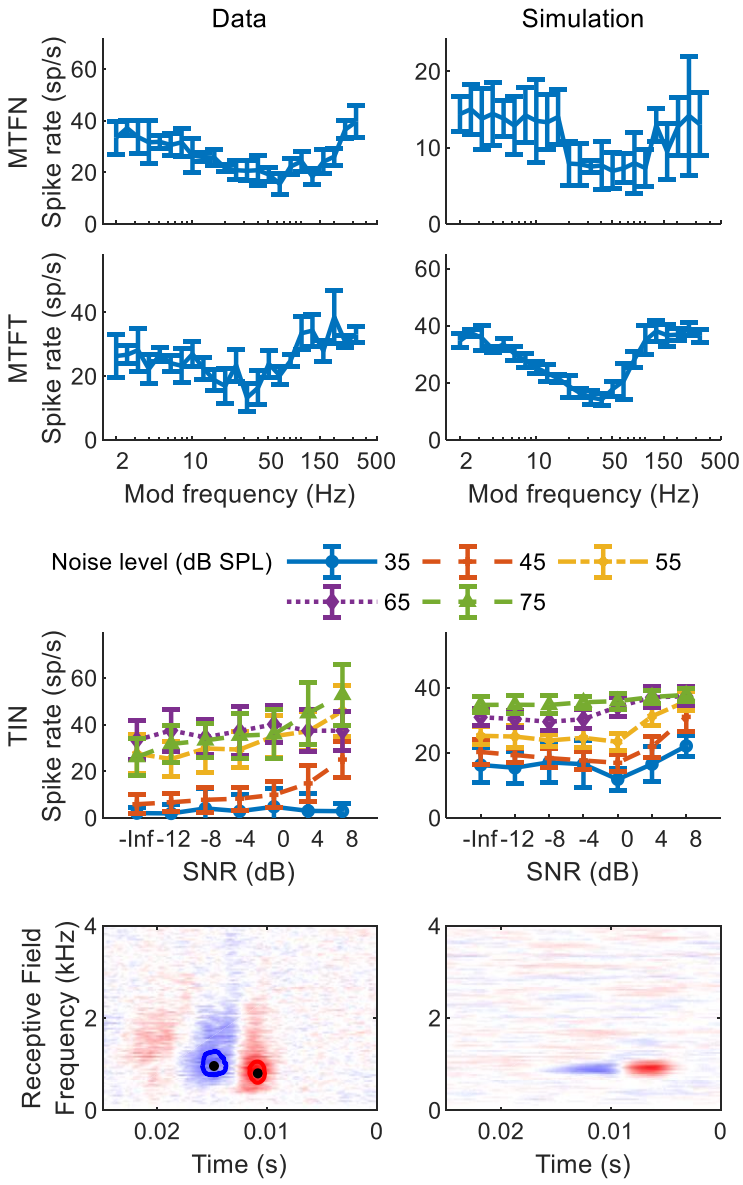

Figure A.4: Example of a band-suppressed IC neuron. Format same as Figure 3.

To determine whether modification significantly improved model performance, a paired t-test was performed between the results of the SFIE model and results of each of the modified models for each stimulus type. For the TIN stimuli, the average correlation coefficient across five noise levels was used in the test. None of the modified models had

significantly higher predictive value than the original SFIE model for any stimulus type. Thus, the performance of the general model structure was relatively robust, and it was not strongly influenced by fine-tuning the model parameters based on the RF of an individual neuron. Because internal noise was included in the AN model, correlations between model and neural responses varied slightly across simulations and models. The percentage of simulations that were significantly correlated to multiple responses was calculated based on average results of two rounds of simulations with the five models described above. Model responses were significantly correlated to a single neuron's response to at least one type of MTF (with tone or noise carrier, or both) and at least three noise levels for TIN in 42.7% of cases. Model responses were significantly correlated to both types of MTF and at least three TIN levels in 17.4% of cases.

A.5 Discussion

An SFIE-type model with fixed parameters can simulate a single neuron's responses to different stimulus types, for IC neurons with both band-enhanced and band-suppressed MTFs. The shape and best-modulation frequency of the MTF could be replicated for many neurons in the current study. In response to tone-in-noise stimuli, trends in the modeled responses with increasing tone level at different overall noise levels agreed with the neural responses for many neurons. For responses to white noise, the general pattern of excitation and inhibition was simulated, but discrepancies between the simulated and neural responses require further study.

Although the SFIE model was originally developed to understand modulation tuning in IC neurons, there were some neurons for which the MTFs were not successfully simulated. This discrepancy might be due to differences between the frequency and

latency picked from the receptive field and the set of values used to describe the standard BE MTFs in the original SFIE model. Large changes in these parameters may not be compatible with the simple structure of the SFIE model; for example, large changes in the latency of inhibition can result in discontinuities in the SFIE model's impulse response.

In response to TIN stimuli, simulated rates decreased more strongly with increasing overall noise level than was observed in IC BE neurons. This decrease in rate in response to TIN stimuli is influenced by saturation of the inner hair cell in the AN model (Carney, 2018), thus differences between simulated and neural responses may be due to differences between AN model properties and the rabbit AN responses (see below).

Failure to predict some neural responses could be due to limitations in interpreting patterns of the receptive field. We used the centers of the excitation and inhibition RF areas to specify the frequency and latency of model inputs. However, in many cases, the excitation and inhibition spanned a wider frequency range in the RF, which may reflect wider tuning in the rabbit periphery in comparison to the cat tuning in the AN model. Whether the bandwidth of peripheral filters plays a role in these modeling results could be tested in the future using an AN model adapted to the rabbit. Also, 2nd-order Wiener kernel analysis only shows the "net" excitation or inhibition at one frequency and time. Therefore, the frequency and timing parameters estimated from the RF for use in the SFIE models may not be accurate estimates of the underlying excitatory and inhibitory response components.

Patterns of excitation and inhibition in the RF can be complicated. In most neural RFs, excitation has a shorter latency than inhibition. However, in a small number of neurons that were not included in this study, the latency of the inhibition was shorter than

that of excitation. Simply reversing the order of the excitation and inhibition in the models used here did not yield successful simulations of these neurons. Other neurons (not included) had RFs with several excitatory and inhibitory bands from which model parameters were not easily specified. The complex patterns in the RFs provide a strong challenge for simplified neural models of the type used here.

Our results also show that the SFIE model without modifications performed approximately as well as the modified models. Thus, although the SFIE model performance was relatively robust, it was not sensitive to changes in the input frequencies and latencies based on the RF that were hypothesized to improve simulations. Alternative models that combine excitation and inhibition should be explored in the future, including coincidence detectors that receive excitatory and inhibitory inputs (Krips & Furst, 2009).

Acknowledgement

This research is supported by NIH-R01-DC010813 and NIH-R00-DC013792.

Bibliography

Cant, N. B., & Oliver, D. L. (2018). Overview of Auditory Projection Pathways and Intrinsic Microcircuits. *Mammalian Auditory Pathways: Synaptic Organization and Microcircuits, 65*, 7-39. doi:10.1007/978-3-319-71798-2_2

Carney, L. H. (2018). Supra-Threshold Hearing and Fluctuation Profiles: Implications for Sensorineural and Hidden Hearing Loss. *J Assoc Res Otolaryngol, 19*(4), 331-352. doi:10.1007/s10162-018-0669-5

Carney, L. H., Li, T., & McDonough, J. M. (2015). Speech Coding in the Brain: Representation of Vowel Formants by Midbrain Neurons Tuned to Sound Fluctuations. *eNeuro, 2*(4). doi:10.1523/ENEURO.0004-15.2015

Carney, L. H., Zilany, M. S., Huang, N. J., Abrams, K. S., & Idrobo, F. (2014). Suboptimal use of neural information in a mammalian auditory system. *Journal of Neuroscience, 34*(4), 1306-1313. doi:10.1523/JNEUROSCI.3031-13.2014

Eggermont, J. J., Aertsen, A. M., & Johannesma, P. I. (1983). Quantitative characterisation procedure for auditory neurons based on the spectro-temporal receptive field. *Hear Res, 10*(2), 167-190. doi:10.1016/0378-5955(83)90052-7

Escabi, M. A., & Schreiner, C. E. (2002). Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. *Journal of Neuroscience, 22*(10), 4114-4131. doi:20026325

Fan, L., & Carney, L. H. (2017). Neural responses in the inferior colliculus to diotic tone-in-noise stimuli support detection based on envelope and neural fluctuations. Abstract, ARO.

Krips, R., & Furst, M. (2009). Stochastic properties of coincidence-detector neural cells. *Neural Comput, 21*(9), 2524-2553. doi:10.1162/neco.2009.07-07-563

Lewis, E. R., Henry, K. R., & Yamada, W. M. (2002). Tuning and timing of excitation and inhibition in primary auditory nerve fibers. *Hear Res, 171*(1-2), 13-31. doi:10.1016/s0378-5955(02)00290-3

Nelson, P. C., & Carney, L. H. (2004). A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *J Acoust Soc Am, 116*(4 Pt 1), 2173-2186. doi:10.1121/1.1784442

Nelson, P. C., & Carney, L. H. (2007). Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *Journal of Neurophysiology, 97*(1), 522-539. doi:10.1152/jn.00776.2006

Richards, V. M., & Nekrich, R. D. (1993). The incorporation of level and level-invariant cues for the detection of a tone added to noise. *J Acoust Soc Am, 94*(5), 2560-2574. doi:10.1121/1.407368

Zilany, M. S., Bruce, I. C., & Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. *J Acoust Soc Am, 135*(1), 283-286. doi:10.1121/1.4837815