Reverse Correlation Uncovers More Complete Tinnitus Spectra

Alec Hoyland, Nelson V. Barnett, Benjamin W. Roop, Danae Alexandrou, Myah Caplan, Jacob Mills, Benjamin Parrell, Divya A. Chari, and Adam C. Lammert†

***Abstract*—*Goal:* This study validates a method for characterizing the sounds experienced by tinnitus patients with potential for characterizing a wider range of sounds than currently possible. The approach is based on reverse correlation, an established behavioral method widely used in psychophysics for unconstrained characterization of internal perceptual representations. *Methods:* Ten normal-hearing subjects participated in a reverse correlation experiment, where they assessed the subjective similarity of random auditory stimuli and target tinnitus-like sounds ("buzzing" and "roaring"). Reconstructions of the targets were obtained by regressing subject responses on the stimuli, and were compared for accuracy to the frequency spectra of the targets using Pearson's r. *Results:* Reconstruction accuracy was significantly higher than chance across subjects: buzzing: 0.52 ± 0.27 (mean ± s.d.), $t(9) = 5.766$, $p < 0.001$; roaring: 0.62 ± 0.23, $t(9) = 5.76$, $p < 0.001$; combined: 0.57 ± 0.25, $t(19) = 7.542$, $p < 0.001$. *Conclusion:* Reverse correlation can accurately reconstruct non-tonal tinnitus-like sounds in normal-hearing subjects, indicating its potential for characterizing the sounds experienced by patients with non-tonal tinnitus.**

***Index Terms*—reverse correlation, tinnitus, behavioral assay**

***Impact Statement*—Characterization of tinnitus can inform tinnitus treatment by allowing for individualized sound therapies, leading to better outcomes for patients suffering from the cognitive and psychological effects of tinnitus.**

I. Introduction

TINNITUS—the perception of sound (*e.g.*, ringing, buzzing) in the absence of any corresponding external stimulus—affects over 25 million people in the U.S., with some estimates ranging up to 50 million [1], [2], a third of whom experience functional cognitive impairment and substantial reduction in quality of life [3]–[5]. Primary treatment options for tinnitus are currently limited by a lack of methods for accurately characterizing the full breadth of internal sounds experienced by patients. Clinical guidelines for tinnitus management frequently involve targeted exposure to tinnitus-like external sounds, either as a form of habituation therapy (termed *sound therapy*), or in the context of Cognitive Behavioral Therapy, where patients are encouraged to perceive their tinnitus as a neutral stimulus [6]. Critically, treatment outcomes have been repeatedly shown to improve when the external sounds used in sound therapy are closely informed by the internal tinnitus experience of the patient [7]–[16]. However, existing strategies for characterizing the tinnitus sounds, such as Pitch Matching (PM) methods, are best suited for patients with tonal tinnitus, who experience sounds resembling pure tones (*e.g.*, ringing). This group may represent between 50-80% of all tinnitus patients [17]–[20], with the other 20-50% experiencing nontonal (*e.g.*, buzzing, roaring) tinnitus sounds. There is a pressing need for characterization methods targeting tinnitus patients who are not served by current procedures [21]–[24].

Tinnitus sounds are presumed to be both complex and heterogeneous [20], [21], [25], although little has been firmly established about the specific characteristics. Therefore, we base our approach on the demonstrated capabilities of *reverse correlation* (RC), an established behavioral method [26]–[28] for probing internal representations of visual and auditory percepts that is unconstrained by prior knowledge about the percepts themselves. RC asks subjects to render subjective judgments (*e.g.*, "does this sound like your tinnitus?") over random stimuli that have no direct relation to the percepts of interest, from which correspondingly unbiased estimates of those objects can be produced. RC is closely related to Wiener theory [29], which has also inspired so-called white-noise approaches to system characterization in physiology [30], [31] and engineering more broadly [32].

Here, we validate RC as a method for characterizing all component frequencies of non-tonal tinnitus sounds by estimating the full psychoacoustic tinnitus spectrum (PTS). To that end, we asked normal-hearing participants to complete an augmented RC experiment where they had to compare random stimuli to a target tinnitus-like sound. Target sounds were subsequently used as the basis for validating the quality of the estimated PTS. Our results demonstrate, for the first time, that tinnitus-like sounds with complex spectra can be accurately estimated using RC. This discovery may open the door to developing a clinical assay based on RC for characterizing the full range of sounds experienced by tinnitus patients.

This work was supported in part by the University of Massachusetts Center for Clinical and Translation Science.

A. Hoyland is with the Department of Biomedical Engineering (BME) at Worcester Polytechnic Institute (WPI), Worcester, MA and with Clarifai, Inc., Wilmington, DE. N.V. Barnett, M. Caplan, J. Mills and A.C. Lammert are also with BME at WPI. B.W. Roop is formerly of the Neuroscience Program at WPI, now with MIT Lincoln Laboratory, Lexington, MA. D. Alexandrou is with the Stritch School of Medicine at Loyola University Chicago, IL. B. Parrell is with the Department of Communication Sciences and Disorders and the Waisman Center at University of Wisconsin, Madison, WI. D.A. Chari is with the University of Massachusetts Chan Medical School, Worcester, MA and the Massachusetts Eye and Ear Infirmary, Boston, MA. The authors declare no competing interests.

† Corresponding Author

II. Materials and Methods

Software code used for the experiments and analysis was written in MATLAB (Mathworks, Inc., Natick, Massachusetts) and is freely available at https://github.com/alec-hoyland/tinnitus-project/.

A. Stimuli

Stimuli were composed of $b = 8$ Mel-spaced frequency bins, which divide the frequency spectrum between $[100,\ 13,000]$ Hz into continuous segments of equal amplitude (*i.e.*, "rectangular" bins). While $b = 8$ bins were found to provide a good approximation to the target sounds used here (see Fig. 3), more bins results in finer spectral details in the reconstruction as well as increased computational complexity.

To generate a stimulus, we randomly select $[2, 7]$ bins with uniform probability to be "filled" with power 0 dB, while "unfilled" bins were assigned a power of -100 dB. Assigning random phase, we then take the inverse Fourier transform of the spectrum to yield a 500-ms stimulus waveform.

B. Target Sounds

Tinnitus-like target sounds were drawn from online examples maintained by the American Tinnitus Association [33]. The two examples selected for this study corresponded to those described as "buzzing" and "roaring." These sounds were selected for the complexity and complementary nature of their frequency spectra. These examples were downloaded and truncated to 500-ms in duration.

C. Experiment

We recruited $n = 10$ subjects with self-reported normal hearing to participate in the experiment. All procedures were approved by the UMass IRB. Subjects listened over earphones, and manually adjusted the loudness of presented sounds to a self-determined comfortable level. Presentation level was not recorded. Trials followed an $A - X$ presentation paradigm (Fig. 1), where subjects listened to a target sound (A) followed by a stimulus (X). A remained the same for every trial within a given experimental condition, while X was randomly generated for each trial. The task is designed to mimic the process by which a tinnitus patient might compare a random stimulus to their own tinnitus percept in the context of a RC experiment.

Subjects were told that some stimuli had the target sound embedded in them, and were instructed to respond "yes" to any such stimuli, and to respond "no" otherwise. Subjects performed trials in blocks of 100, with breaks between blocks. Subjects completed two (2) blocks of trials per experimental condition (*i.e.*, for a given target, A). The 400 total trials took subjects about 10-20 minutes to complete.

D. Reconstruction

A subject performing n RC trials with b frequency bins produces a stimulus matrix $\Psi \in \mathbb{R}^{b \times n}$ and a response vector $y \in \{1, -1\}^n$, where 1 corresponds to a "yes" response and -1 to a "no." RC classically assumes the subject response model:

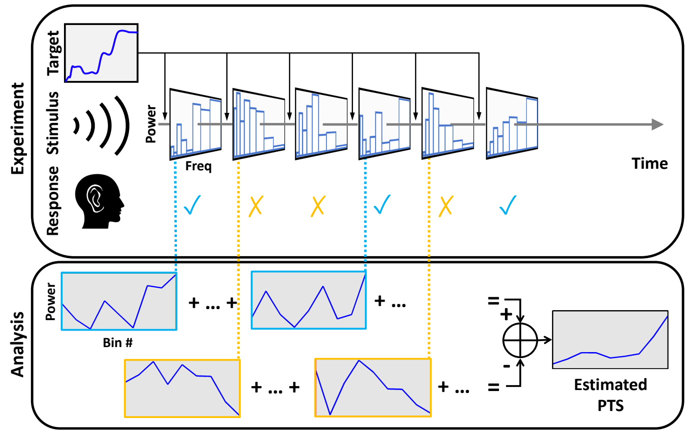

Fig. 1. Diagram of the experimental protocol. Subjects listen to a series of random stimuli, each preceded by a target sound. Subjects compare the stimulus to the target, and respond either "yes" or "no" depending on their perceived similarity. The recorded stimulus-response pairs are used to form an estimate of the target using a restricted regression procedure.

$$y = sign(\Psi^{\mathrm{T}} x), \tag{1}$$

where $x \in \mathbb{R}^{\mathrm{b}}$ represents the subject's internal representation of interest (*e.g.*, their tinnitus percept). This model can be inverted, yielding the following commonly-used formula for reconstructing internal representations in RC experiments:

$$\hat{x} = \frac{1}{n} \Psi y \tag{2}$$

This formula is a restricted form of the Normal equation—the least-squares solution to a linear regression problem—under the assumption that the stimulus dimensions are uncorrelated [27]. Intuitively, this implies that the reconstruction, $\hat{x}$, is the sum of all stimuli eliciting a "yes" response, after subtracting the sum of all stimuli eliciting a "no."

E. Validation

By adopting the $A - X$ paradigm for the present experiments, we have the opportunity to directly validate the reconstructions $\hat{x}_{buzzing}$ and $\hat{x}_{roaring}$ against their corresponding target sounds. To that end, we represent the power spectra of the target sounds as vectors $s_{buzzing} \in \mathbb{R}^{\mathrm{b}}$ and $s_{roaring} \in \mathbb{R}^{\mathrm{b}}$ using the same $b = 8$ frequency bins used in stimulus generation. The power in each bin is taken to be the mean power level for frequencies contained within that bin. The vectors $s_{buzzing}$ and $s_{roaring}$ can then be directly compared to their corresponding reconstructions $\hat{x}_{buzzing}$ and $\hat{x}_{roaring}$ using Pearson's r, representing a measure of reconstruction accuracy. One-sample t-tests were performed on the mean Fisher-transformed r values across subjects to assess significant differences from zero.

F. Synthetic and Random Subjects

For the purposes of establishing upper and lower bounds on human performance in this experiment, additional experiments were run *in-silico* with two simulated subjects, representing a participant who always gives *ideal* responses and one who always gives *random* responses. Each experiment ran for $n = 200$ trials and was repeated 1000 times.

The *ideal* subject gives responses following:

$$y_i = \begin{cases} 1 & \text{if } \Psi_i s \geq Q(0.5; \Psi s) \\ -1 & \text{otherwise} \end{cases} \tag{3}$$

for $i \in 1, \ldots, n$, where $Q(x, y)$ is the quantile function for $x \in [0, 1]$ of the similarity calculation Ψs, and Ψ_i is the i^{th} column of Ψ. The ideal subject represents an *upper bound* on task performance, where the subject has precise knowledge of every stimulus and where the reconstruction algorithm mirrors the subject response model assumed by RC.

The *random* subject chooses responses at random:

$$r_i = \begin{cases} 1 & \text{if } X > 0.5 \\ -1 & \text{otherwise} \end{cases} \tag{4}$$

where $X \in [0, 1]$ is a uniform random variable. The random subject represents a *lower bound* on task performance, where the subject ignores any knowledge of the stimulus, or indeed their own internal representation.

III. Results

Figure 2 shows the distribution of Pearson's r (*i.e.*, reconstruction accuracy) from human, ideal, and random subject responses in the buzzing and roaring experimental conditions. Accuracies from human responses were generally higher than those from random subject responses, with some accuracies approaching those from ideal subject responses. Accuracy from the random subject was 0.00 ± 0.44 (mean $\pm$ st.dev.) for buzzing and 0.00 ± 0.39 for roaring, while mean accuracy from human responses was significantly different from 0 in all conditions: buzzing: 0.52 ± 0.27, $t(9) = 5.766$, $p < 0.001$; roaring: 0.62 ± 0.23, $t(9) = 5.76$, $p < 0.001$; combined: 0.57 ± 0.25, $t(19) = 7.542$, $p < 0.001$. From Figure 2, it appears that the distribution of buzzing results differs from that of the roaring results, however the difference between buzzing and roaring is not statistically significant (two-way ANOVA across subjects ($F(13) = 2.94$, $p > 0.05$) and target signals ($F(1) = 2.44$, $p > 0.05$)). The buzzing ground truth spectrum contains significant high-frequency content, suggesting that the task may vary in difficulty for subjects with different audiograms (*i.e.*, for subjects with minor high-frequency hearing loss). Figure 3 shows example reconstructions from the best human responses (as quantified by reconstruction accuracy) overlaid on the target sound spectra, showing that those reconstructions capture many of the salient features of the target.

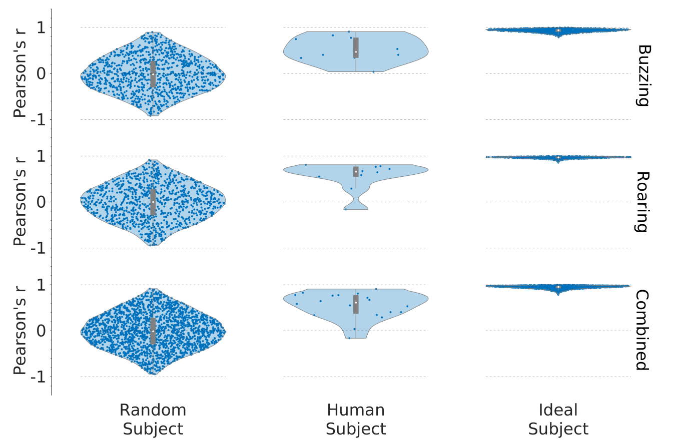

Fig. 2. Reconstruction accuracy for human subjects is significantly above baseline, but is not optimal. Reconstruction accuracies of random, human, and ideal subjects are shown as violin plots with box plots overlaid. The median is a white dot, the ordinate of the blue points are the Pearson's r values. The top row shows accuracies for the buzzing target, the middle row shows accuracies for the roaring target signal, and the bottom row shows results for both buzzing and roaring combined.

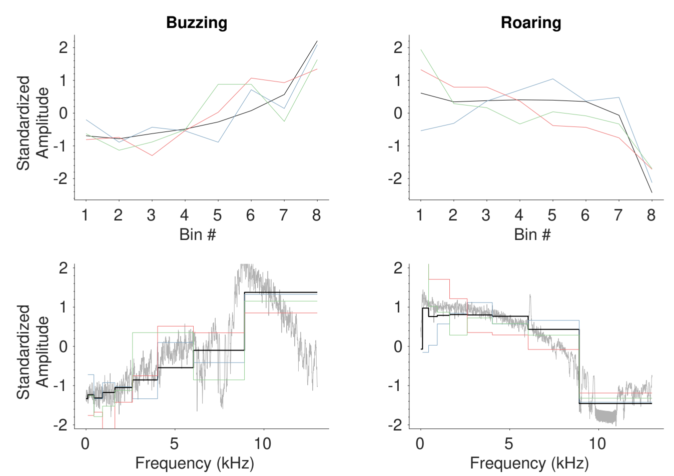

Fig. 3. Reconstructions of the PTS capture many salient features of tinnitus spectra. The left-hand column shows reconstructions for the buzzing target, while the right-hand column shows reconstructions for the roaring target. On each set of axes, the black trace indicates the target, while colored traces plot exemplar reconstructions from human subjects' data. In the top row, the target and reconstructions are shown as standardized power levels within each frequency bin. In the bottom row, those standardized power levels are mapped from the 8-dimensional bin domain to a 11025-dimensional frequency domain. In the bottom row, the unbinned power spectral densities (ground-truth) of the target signals are shown in gray.

IV. Conclusion

Our results show that RC can accurately reconstruct the frequency spectrum of certain relevant non-tonal tinnitus sounds. This indicates that RC may also be useful for characterizing the sounds experienced by patients with non-tonal forms of tinnitus, for whom existing tinnitus characterization methods are ill-suited. As such, this study may represent the first step in the development of a clinical assay to characterize a wider variety of tinnitus percepts than currently possible. The clinical utility of RC is further reinforced by its practical advantages, as an approach that is purely behavioral, requires no specialized equipment, and has a marginal setup time. Moreover, subjects in the present study were able to complete the required number of trials within 20 minutes, indicating that this procedure is feasible as an assay that could be conducted in a single clinical visit. Future work will focus on validating this approach with larger sample sizes, with more target signals, stricter sound pressure level (SPL) control, and in patients suffering from tinnitus.

Reconstruction accuracies observed here are well below those provided by an ideal subject's responses, which may be attributed to noisy human response characteristics. Whereas noisy responses are universally observed in applications of RC, this situation may nonetheless be improved by optimizing the experimental protocol, stimulus generation, and reconstruction method, none of which are guaranteed optimal in this early-stage validation study. For example, recent approaches to improving RC reconstruction methods have been shown to boost the efficiency, noise robustness and overall accuracy of RC [34], [35], and may also be applicable in the context of tinnitus.

ACKNOWLEDGMENT

A. H. thanks Drs. Srinivas Gorur-Shandilya and Mark Zielinski for informative discussions regarding data visualization.

REFERENCES

[1] J. M. Bhatt, H. W. Lin, and N. Bhattacharyya, "Prevalence, Severity, Exposures, and Treatment Patterns of Tinnitus in the United States," *JAMA Otolaryngology–Head & Neck Surgery*, vol. 142, no. 10, pp. 959–965, Oct. 2016.

[2] National Institute on Deafness and Other Communication Disorders, "Quick Statistics About Hearing," https://www.nidcd.nih.gov/health/statistics/quick-statistics-hearing.

[3] D. B. Ahmed, A. Ahmed, D. T. Akhtar, and D. S. Salim, "Impact of Tinnitus Perception on Psychological Distress in Male and Female Tinnitus Patients," *Foundation University Journal of Psychology*, vol. 1, no. 1, pp. 01–26, Jan. 2017.

[4] D. M. Nondahl, K. J. Cruickshanks, D. S. Dalton, B. E. K. Klein, R. Klein, C. R. Schubert, T. S. Tweed, and T. L. Wiley, "The Impact of Tinnitus on Quality of Life in Older Adults," *Journal of the American Academy of Audiology*, vol. 18, no. 3, pp. 257–266, Mar. 2007.

[5] S. Tegg-Quinn, R. J. Bennett, R. H. Eikelboom, and D. M. Baguley, "The impact of tinnitus upon cognition in adults: A systematic review," *International Journal of Audiology*, vol. 55, no. 10, pp. 533–540, Oct. 2016.

[6] D. E. Tunkel, C. A. Bauer, G. H. Sun, R. M. Rosenfeld, S. S. Chandrasekhar, E. R. Cunningham, S. M. Archer, B. W. Blakley, J. M. Carter, E. C. Granieri, J. A. Henry, D. Hollingsworth, F. A. Khan, S. Mitchell, A. Monfared, C. W. Newman, F. S. Omole, C. D. Phillips, S. K. Robinson, M. B. Taw, R. S. Tyler, R. Waguespack, and E. J. Whamond, "Clinical Practice Guideline: Tinnitus," *Otolaryngology–Head and Neck Surgery*, vol. 151, no. 2_suppl, pp. S1–S40, Oct. 2014.

[7] P. B. Davis, B. Paki, and P. J. Hanley, "Neuromonics Tinnitus Treatment: Third Clinical Trial," *Ear and Hearing*, vol. 28, no. 2, pp. 242–259, Apr. 2007.

[8] M. Landgrebe, A. Azevedo, D. Baguley, C. Bauer, A. Cacace, C. Coelho, J. Dornhoffer, R. Figueiredo, H. Flor, G. Hajak, P. van de Heyning, W. Hiller, E. Khedr, T. Kleinjung, M. Koller, J. M. Lainez, A. Londero, W. H. Martin, M. Mennemeier, J. Piccirillo, D. De Ridder, R. Rupprecht, G. Searchfield, S. Vanneste, F. Zeman, and B. Langguth, "Methodological aspects of clinical trials in tinnitus: A proposal for an international standard," *Journal of Psychosomatic Research*, vol. 73, no. 2, pp. 112–121, Aug. 2012.

[9] A. K. Nickel, T. Hillecke, H. Argstatter, and H. V. Bolay, "Outcome Research in Music Therapy," *Annals of the New York Academy of Sciences*, vol. 1060, no. 1, pp. 283–293, 2005.

[10] H. Okamoto, H. Stracke, W. Stoll, and C. Pantev, "Listening to tailor-made notched music reduces tinnitus loudness and tinnitus-related auditory cortex activity," *Proceedings of the National Academy of Sciences*, vol. 107, no. 3, pp. 1207–1210, Jan. 2010.

[11] R. Schaette, O. König, D. Hornig, M. Gross, and R. Kempter, "Acoustic stimulation treatments against tinnitus could be most effective when tinnitus pitch is within the stimulated frequency range," *Hearing Research*, vol. 269, no. 1, pp. 95–101, Oct. 2010.

[12] A. Stein, A. Engell, M. Junghoefer, R. Wunderlich, P. Lau, A. Wollbrink, C. Rudack, and C. Pantev, "Inhibition-induced plasticity in tinnitus patients after repetitive exposure to tailor-made notched music," *Clinical Neurophysiology*, vol. 126, no. 5, pp. 1007–1015, May 2015.

[13] P. A. Tass, I. Adamchic, H.-J. Freund, T. von Stackelberg, and C. Hauptmann, "Counteracting tinnitus by acoustic coordinated reset neuromodulation," *Restorative Neurology and Neuroscience*, vol. 30, no. 2, pp. 137–159, Jan. 2012.

[14] H. Wang, D. Tang, Y. Wu, L. Zhou, and S. Sun, "The state of the art of sound therapy for subjective tinnitus in adults," *Therapeutic Advances in Chronic Disease*, vol. 11, p. 2040622320956426, Jan. 2020.

[15] P. Neff, B. Langguth, M. Schecklmann, R. Hannemann, and W. Schlee, "Comparing three established methods for tinnitus pitch matching with respect to reliability, matching duration, and subjective satisfaction," *Trends in Hearing*, vol. 23, p. 2331216519887247, 2019.

[16] S. Schoisswohl, J. Arnds, M. Schecklmann, B. Langguth, W. Schlee, and P. Neff, "Amplitude modulated noise for tinnitus suppression in tonal and noise-like tinnitus," *Audiology and Neurotology*, vol. 24, no. 6, pp. 309–321, 2019.

[17] J. S. Turner JR, "Auditory dysfunction: Tinnitus," *Clinical Methods: The History, Physical, and Laboratory Examinations*, 1990.

[18] J. A. Henry, K. E. James, K. Owens, T. Zaugg, E. Porsov, and G. Silaski, "Auditory test result characteristics of subjects with and without tinnitus." *Journal of Rehabilitation Research & Development*, vol. 46, no. 5, 2009.

[19] O. C. Ukaegbe, F. T. Orji, B. C. Ezeanolue, J. O. Akpeh, and I. A. Okorafor, "Tinnitus and its effect on the quality of life of sufferers: A Nigerian cohort study," *Otolaryngology–Head and Neck Surgery*, vol. 157, no. 4, pp. 690–695, 2017.

[20] D. Vajsakovic, M. Maslin, and G. D. Searchfield, "Principles and Methods for Psychoacoustic Evaluation of Tinnitus," *Current Topics in Behavioral Neurosciences*, Feb. 2021.

[21] J. A. Henry, K. M. Reavis, S. E. Griest, E. J. Thielman, S. M. Theodoroff, L. D. Grush, and K. F. Carlson, "Tinnitus: An Epidemiologic Perspective," *Otolaryngologic Clinics of North America*, vol. 53, no. 4, pp. 481–499, Aug. 2020.

[22] J. A. Henry, ""Measurement" of Tinnitus," *Otology & Neurotology*, vol. 37, no. 8, p. e276, Sep. 2016.

[23] A. Noreña, C. Micheyl, S. Chéry-Croze, and L. Collet, "Psychoacoustic Characterization of the Tinnitus Spectrum: Implications for the Underlying Mechanisms of Tinnitus," *Audiology & neuro-otology*, vol. 7, pp. 358–69, Nov. 2002.

[24] L. E. Roberts, G. Moffat, and D. J. Bosnyak, "Residual inhibition functions in relation to tinnitus spectra and auditory threshold shift," *Acta Oto-Laryngologica*, vol. 126, no. sup556, pp. 27–33, 2006.

[25] M. B. Meikle, T. A. Creedon, and S. E. Griest, "Tinnitus Archive: Archive Background and Development," http://www.tinnitusarchive.org/appendix/archiveBackgroundAndDevelopment/, 2004.

[26] A. Ahumada and J. Lovell, "Stimulus Features in Signal Detection," *The Journal of the Acoustical Society of America*, vol. 49, no. 6B, pp. 1751–1756, Jun. 1971.

[27] F. Gosselin and P. G. Schyns, "Superstitious Perceptions Reveal Properties of Internal Representations," *Psychological Science*, vol. 14, no. 5, pp. 505–509, Sep. 2003.

[28] W. O. Brimijoin, M. A. Akeroyd, E. Tilbury, and B. Porr, "The internal representation of vowel spectra investigated using behavioral response-triggered averaging," *The Journal of the Acoustical Society of America*, vol. 133, no. 2, Feb. 2013.

[29] N. Wiener, *Nonlinear Problems in Random Theory*, 1966.

[30] D. Ringach and R. Shapley, "Reverse correlation in neurophysiology," *Cognitive Science*, vol. 28, no. 2, pp. 147–166, 2004.

[31] P. Z. Marmarelis and V. Z. Marmarelis, "The White-Noise Method in System Identification," in *Analysis of Physiological Systems: The White-Noise Approach*, ser. Computers in Biology and Medicine, P. Z. Marmarelis and V. Z. Marmarelis, Eds. Boston, MA: Springer US, 1978, pp. 131–180.

[32] L. Ljung, *System Identification: Theory for the User*. Pearson Education, Dec. 1998.

[33] "Measuring Tinnitus — American Tinnitus Association," Aug. 2022.

[34] B. W. Roop, B. Parrell, and A. C. Lammert, "A Compressive Sensing Approach for Inferring Cognitive Representations with Reverse Correlation," *bioRxiv*, 2021.

[35] A. Compton, B. W. Roop, B. Parrell, and A. C. Lammert, "Stimulus whitening improves the efficiency of reverse correlation," *Behavior Research Methods*, Aug. 2022.