Reverse Correlation Uncovers More Complete Tinnitus Spectra

Alec Hoyland, Nelson Barnett, Benjamin W. Roop, and Adam C. Lammert[†]

***Abstract*—— *Goal*: Make tinnitus characterization better. *Methods*: Use reverse correlation and compressed sensing. *Results*: Cool and fun results! Much data, very wow! *Conclusions*: Alec and Adam are cool and smart.**

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

***Impact Statement*—30 words on significance**

I. Introduction

TINNITUS—the perception of sound (*e.g.* ringing, buzzing) in the absence of an external stimulus—affects over 25 million people in the U.S., with some estimates ranged up to 50 million, a third of which experience functional cognitive impairment and substantial reduction in quality of life [1], [2]. Primary treatment options for tinnitus are currently limited by a lack of methods for accurately characterizing the internal sounds experienced by patients. Tinnitus treatment typically involves *sound therapy*, a form of habituation therapy, which involves target exposure to external sounds to attenuate the perception of tinnitus or to encourage patients to perceive their tinnitus as a neutral stimulus [3]. 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 [4]–[7], specifically, its component frequencies that constitute the *psychoacoustic tinnitus spectrum* (PTS). However, existing methods for characterizing the PTS rely on reductionist assumptions concerning the nature of tinnitus sounds (*e.g.* that they are pure tones or have small-width Gaussian spectra) and produce characterizations that are correspondingly bias and incomplete when compared to the spectral variety of tinnitus percepts—less than 50% of tinnitus patients report their tinnitus sounding like “ringing” [2]. There is a pressing need for methods to more completely characterize the PTS [1], [8], [9], to further improve treatment outcomes for patients suffering from tinnitus.

This work was supported in part by the University of Massachusetts Center for Clinical and Translation Science Pilot Project Program 2022, grant no. GRANT NUMBER.

Mr. Hoyland is a graduate student in the Department of Biomedical Engineering at Worcester Polytechnic Institute (WPI) and a research scientist at Clarifai Inc. Mr. Nelson Barnett is a research assistant at WPI. Mr. Benjamin W. Roop is a research scientist at MIT Lincoln Laboratory. Dr. Lammert is an Assistant Professor in the Department of Biomedical Engineering at WPI. The authors declare no competing interests.

[†] Corresponding Author

Code is freely available at https://github.com/alec-hoyland/tinnitus-project/. Data are available upon request.

We utilize *reverse correlation* approach to characterize the PTS more completely, without the strong biases introduced by existing methods. Reverse correlation is a widely-accepted method for unconstrained and unbiased estimation of latent neural representations (*e.g.* neural receptive fields) based on the white-noise method for black-box system identification [10], [11]. This method has been used to characterize the psychophysical processes of perception in vision (*e.g.* faces) and audition (*e.g.* phonemes) [12]–[14] all on the basis of stimulus-response data [15], [16]. In reverse correlation, subjects are presented with richly-varying random stimuli (*e.g.* white noise) and make simple “yes/no” responses about whether they perceive a particular signal (*e.g.* their tinnitus percept). Internal representations, such as the PTS, can be estimated by regressing subject responses against the stimuli over many trials. Despite widespread use in characterizing neural and cognitive representations, reverse correlation has never been used for tinnitus characterization.

In the present study, we validate a reverse correlation-based PTS reconstruction assay on healthy hearing subjects via a template-matching experiment, in which subjects compare the reverse correlation stimuli to a target tinnitus sample sound. We benchmark the accuracy and reliability of the reconstructions, using the target tinnitus signal as a ground truth. High accuracy and reliablity measures relative to low numbers of trials per subject provide evidence for the feasibility of the reverse correlation paradigm for uncovering spectral representations of tinnitus.

Our long-term goal is to improve outcomes for patients suffering from tinnitus by providing a validated clinical assay, based on the demonstrated capabilities of the reverse correlation approach, that clinicians can use to accurately and efficiently characterize the individualized perceptual experience of tinnitus. Our guiding hypothesis is that reverse correlation will produce PTS estimates that patients will consistently report as being similar to their own tinnitus experience. The rationale for this work is that accurate and efficient characterization of the PTS can inform individualized tinnitus treatment with enhanced habituation therapies.

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/. The code depends on [17]–[19].

A. Stimulus Generation

To generate stimuli, we partition the frequency space $f \in [100,\ 13,000]$ Hz into $b = 8$ mel-spaced frequency bins so that each bin is perceptually different to a listener [20] (*cf.* Supplementary Information). All frequencies in the same bin have the same amplitude. For each stimulus, we randomly select $[2, 7]$ bins with equal probability to be "filled" with power 0 dB. Unfilled bins are set to -100 dB. The inverse Fourier transform of the spectrum yields a 500-ms stimulus waveform.

B. Reconstruction Analysis

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

The linear regression solution is given by the normalized inner product of the stimulus matrix and the responses (Eq. 1). This is a restricted form of the normal equation, the least-squares solution to the linear regression problem, under the assumption that the stimulus dimensions are uncorrelated [21]. Intuitively, this implies that the reconstruction, x, is a linear combination of the randomly-generated stimuli that lies closer in simularity to the "yes" stimuli than the "no" ones [13].

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

Reconstruction accuracy was quantified by Pearson's r between the binned target signal spectrum and the reconstructed binned spectrum.

C. Synthetic and Random Subjects

Two *in-silico* experiments were run with synthetic and random subjects. Each experiment ran for $n = 200$ trials and was repeated 1000 times.

The synthetic subject used a template-matching algorithm. Given the target signal spectrum (for either "buzzing" or "roaring"), $s \in \mathbb{R}^t$, and the stimulus matrix in the spectral domain, $\Psi \in \mathbb{R}^{n \times t}$, each element of the synthetic response vector, $y_s \in 1, -1^n$ is defined as:

$$r_i = \begin{cases} 1 & \text{if } \Psi_i^{\mathrm{T}} s \geq Q(0.5; \Psi_i^{\mathrm{T}} s) \\ -1 & \text{otherwise} \end{cases} \tag{2}$$

for $i \in 1, ..., n$, where $Q(x, y)$ is the quantile function for $x \in [0, 1]$, given the empirical distribution of the simularity calculation $Psi^{\mathrm{T}} s$. Thus, the 50% most-similar stimuli receive a "yes" response and the 50% least-similar stimuli receive a "no" response. Due to the nature of this calculation, the synthetic subject has full-knowledge of all unbinned stimuli spectra as well as the unbinned spectrum of the target signal before making responses. The synthetic subject represents a *best-case* scenario, where the subject has precise knowledge of every stimulus and where the reconstruction algorithm mirrors the template-matching algorithm.

The random subject chooses responses at random, with a 50% probability of "yes":

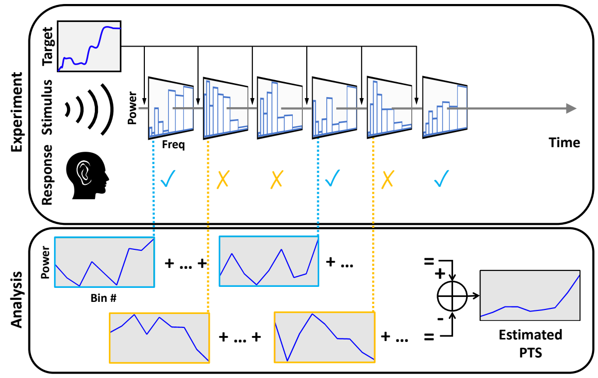

Fig. 1. Diagram of the experimental paradigm. The subject listens to a priming target signal, then a stimulus. They compare their mental model of the stimulus to their mental model of the target signal, before making a binary choice about the two signals' similarity (either "yes" it is similar or "no" it is not).

$$r_i = \begin{cases} 1 & \text{if } X > 0.5 \\ -1 & \text{otherwise} \end{cases} \tag{3}$$

where $X \in [0, 1]$ is a uniform random variable. The random subject reflects choosing responses completely randomly — leading to low expected reconstruction accuracy. Theoretically, this should be the *worst-case* scenario. Human and synthetic subjects are expected to perform better.

D. Experiment

We recruited $n = 10$ subjects for the experiment with healthy hearing from $[100,\ 13,000]$ Hz. Subjects used over-the-ear headphones and manually adjusted loudness using sample stimuli before performing the task.

The subjects performed an AX paradigm binary choice task in 2 blocks of 100 trials with breaks between blocks, for a total of 200 trials per experimental condition. The experiment typically took subjects 10-15 minutes to complete. For each trial in the experiment, the subject listened to a 500-ms auditory target signal followed by one of the 500-ms randomly generated stimuli. The subject was instructed to press the J key if the stimulus sounded similar to the target signal and the F key if the stimulus did not (Fig. 1).

The target signals were drawn from online examples from the American Tinnitus Association, representing the range of tinnitus experiences (American Tinnitus Association, Vienna, Virginia). We selected two example tinnitus waveforms from their website: "buzzing" and "roaring". We represented the target spectra using the same $b = 8$ frequency bins used in stimulus generation, where the power in each bin was taken to be the mean power level in the spectrum for frequencies contained within that bin. In this way, the AX experiment mimicked comparing a randomly-generated stimulus to an internal perception of tinnitus, however the known target signals unify across subjects and provide a gold-standard to benchmark against.

III. Results

Figure 2 reports Pearson's r scores for $n = 10$ human, $n = 1000$ synthetic, and $n = 1000$ random subjects. Most human subjects perform significantly better than the random baseline and some approach the synthetic subject's empirical upper bound. The distributions of human subject r are significantly different from the baseline random results (Mann-Whitney U test, $p < 0.001$).

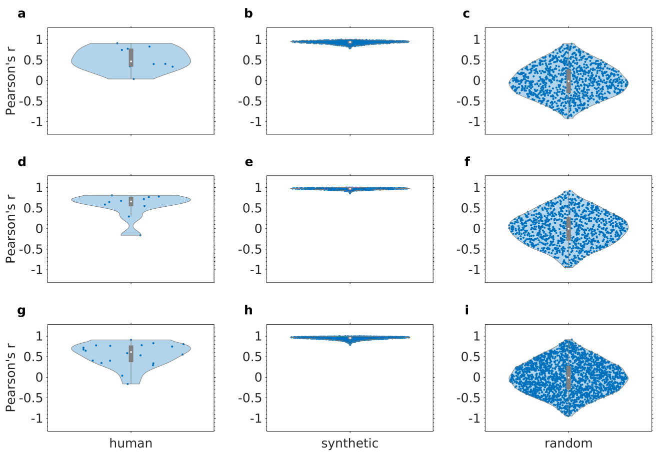

Fig. 2. Reconstruction accuracy for human subjects is significantly above baseline, but is not optimal. The reconstruction accuracy of human, synthetic, and random subjects are shown as violin plots with box plots overlaid. The median is a white dot, the ordinate of the blue points are the true Pearson's r values — the abscissa has no meaning. **a-c** show violin plots for the buzzing target signal, **d-f** show results the roaring target signal, and **g-i** show results for the all data combined. Human subjects perform significantly better than random chance and some approach the empirical maximum of the synthetic subject results (Mann-Whitney U test, $p < 0.001$).

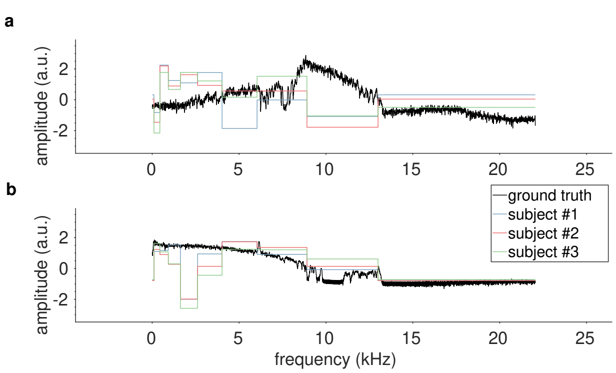

Fig. 3. Reconstructions of the PTS capture many salient features of tinnitus spectra. **(a)** shows buzzing results and **(b)** shows roaring results. In each, the black trace indicates the ground truth, unbinned frequency spectrum of the ATA target signal. The colored traces plot exemplar human subject results, as binned reconstructions mapped from 8-dimensional bin space to 11025-dimensional frequency space. Since the bins span from 100 to 13,000 Hz, in the reconstructions $(13000, 22000]$ Hz is always 0.

IV. Conclusion

We applied reverse correlation to the novel domain of psychoacoustic tinnitus spectrum reconstruction and improved reconstruction performance (PTS) up to 2x using compressed sensing. Using our reconstruction algorithm, the experiment required only 2,000 trials for quality reconstruction results which is a 10x improvement over reverse correlation results in similar domains [13]. Subjects finished the experiment within two hours, indicating that this procedure is feasible as an "outpatient medical test" to characterize the PTS of a subject, a crucial step in diagnosis and treatment [1], [8], [9]. After fine-tuning, we will use this algorithm to characterize the PTSes of clinical tinnitus patients to help in the treatment of patients and to further understanding of tinnitus.

Results from human subjects are far below the synthetic baseline, indicating that some improvements can be made to boost human performance. The simplicity of the mathematical model for the synthetic subject (and its favorability towards the problem) leads to a synthetic subject that accounts for each tonotopic bin with equal attentiveness, does not suffer fatigue, and never changes its threshold for positive vs. negative responses. The experiment asks subjects to rate each stimulus as similar or dissimilar to the target tinnitus signal. This can lead to "threshold drift," where the decision threshold between positive and negative assignments to stimuli change as the experiment goes along. While the AX paradigm protects against this effect somewhat by playing the target sound before each decision stimulus, subjects self-reported feeling unsure that they were making consistent decisions. In future investigations, we will use a two alternate forced choice (2AFC) paradigm, where a subject is provided two stimuli after the cue, and chooses the more similar one. All chosen stimuli get positive responses assigned and non-chosen stimuli get negative responses assigned (CITE?). Reconstructions can proceed normally, using both the positive and negative responses.

mention closed-loop/ML?

Acknowledgment

A. H. thanks Drs. Srinivas Gorur-Shandilya and Mark Zielinski for informative discussions regarding data visualization. The authors thank Mr. Jacob Mills and Ms. Myah Caplan for their assistance in data collection.

References

[1] 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.

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

[3] P. J. Jastreboff, "25 years of tinnitus retraining therapy," *HNO*, vol. 63, no. 4, pp. 307–311, Apr. 2015.

[4] 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.

[5] 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.

[6] 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.

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

[9] 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.

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

[11] G. M. Ljung and G. E. P. Box, "On a measure of lack of fit in time series models," *Biometrika*, vol. 65, no. 2, pp. 297–303, Aug. 1978.

[12] 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.

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

[14] 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.

[15] 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.

[16] P. Neri and D. M. Levi, "Receptive versus perceptive fields from the reverse-correlation viewpoint," *Vision Research*, vol. 46, no. 16, pp. 2465–2474, Aug. 2006.

[17] B. Bechtold, "Violin Plots for Matlab," Dec. 2022.

[18] M. Koch, "Yaml," Dec. 2022.

[19] S. Gorur-Shandilya, "Srinivas.gs_mtools," Jul. 2022.

[20] R. Polfreman, "Sound Spotting - a Frame-Based Approach," Jan. 2001.

[21] F. B. Martin and G. Zyskind, "On Combinability of Information from Uncorrelated Linear Models by Simple Weighting," *The Annals of Mathematical Statistics*, vol. 37, no. 5, pp. 1338–1347, Oct. 1966.