Reverse Correlation Uncovers More Complete Tinnitus Spectra

Alec Hoyland, Nelson Barnett, Benjamin W. Roop, Danae Alexandrou, Myah Caplan, Jacob Mills, Ben Parell, Divya A. Chari, 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 any corresponding external stimulus—affects over 25 million people in the U.S., with some estimates ranging up to 50 million, a third of whom 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 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 [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]. However, existing methods for characterizing the tinnitus sounds—Pitch Matching (PM) methods—are best suited for patients with tonal tinnitus, who experience sounds resembling pure tones (*e.g.*, ringing). This group represents only half of all tinnitus patients [2], with others experiencing nontonal (*e.g.*, buzzing, roaring) tinnitus sounds CITE. Moreover, PM fails for up to 30% of patients with tonal tinnitus CITE. There is a pressing need for characterization methods targeting tinnitus patients for who are not served by current characterization methods [1], [8], [9].

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

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. Barnett is a post-baccalaureate research associate at WPI. Mr. Roop is a research scientist at MIT Lincoln Laboratory. Ms. Caplan is a research assistant at WPI. Mr. Mills is a research assistant at WPI. 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.

Tinnitus sounds are presumed to be both complex and heterogeneous [1], [2], 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 [10]–[12] 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 Weiner theory CITE, which has also inspired so-called white-noise approaches to system characterization in physiology [13]–[15] and engineering more broadly [16].

Here, we validate RC as a method for characterizing non-tonal tinnitus sounds, specifically for estimating the full psychoacoustic tinnitus spectrum (PTS), including all component frequencies of the tinnitus sound. 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, a result that may open the door the developing a clinical assay based on RC for characterizing the full range of sounds experienced by tinnitus patients.

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. 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 Figure 3), the number of bins would need to be increased in proportion to the level of detail in the frequency spectrum.

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 power of -100 dB. Assigning

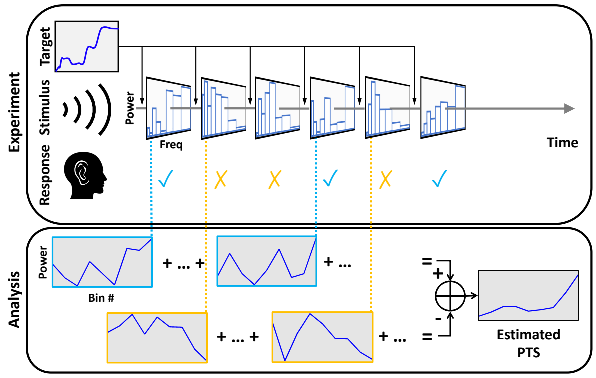

Fig. 1. Diagram of the experimental paradigm. The subject listens to a priming target signal, then a stimulus. They compare their mental model of the stimulus to their mental model of the target signal, before making a binary choice about the two signals' similarity (either "yes" it is similar or "no"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

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 CITE. 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. Trials followed an $A-X$ presentation paradigm (Figure 1), where subject 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 (Figure 1).

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 response "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 $\Psi \in \mathbb{R}^{n\times b}$ 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:

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

where $x \in R^{\mathrm{t}}$ 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 linear regression problem—under the assumption that the stimulus dimensions are uncorrelated [11], [20]. 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 compare the reconstructions $\hat{x}_{buzzing}$ and $\hat{x}_{roaring}$ for the experimental conditions to their corresponding target sound for the purposes of validation. To that end, we represent the power spectra of the target sounds as vectors $s_{buzzing} \in \mathbb{R}^8$ and $s_{roaring} \in \mathbb{R}^8$ using the same $b = 8$ frequency bins used in stimulus generation, where 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, with 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:

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

where $Q(x,y)$ is the quantile function for $x \in [0,1]$ of the simularity calculation $Psi^{\mathrm{T}} s$. The ideal subject represents a *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 experiments conditions. Accuracies from human responses were generally higher than those from random subject responses, with some accuracies approaching those from ideal subject responses. Mean accuracy from the random subject responses was 0 in all conditions, while mean accuracy from human responses was significantly different from 0 in all conditions: buzzing: $t(9) = 5.766$, $p < 0.001$; roaring: $t(9) = 5.76$, $p < 0.001$; combined: $t(19) = 7.542$, $p < 0.001$. Figure 3 shows example reconstructions from human responses overlaid on the target sound spectra, showing that those reconstructions capture many of the salient features of the the target.

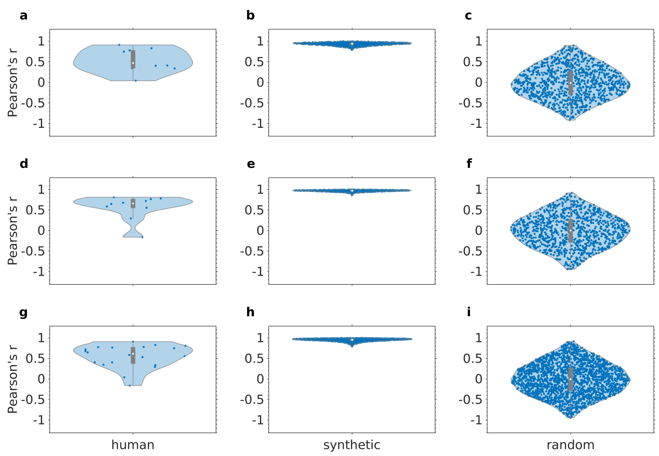

Fig. 2. Reconstruction accuracy for human subjects is significantly above baseline, but is not optimal. The reconstruction accuracy of human, synthetic, and random subjects are shown as violin plots with box plots overlaid. The median is a white dot, the ordinate of the blue points are the true Pearson's 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.

IV. Conclusion

Our results show that RC can accurately reconstruct the frequency spectrum of complex, non-tonal tinnitus-like sounds, indicating 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 in patients suffering from tinnitus.

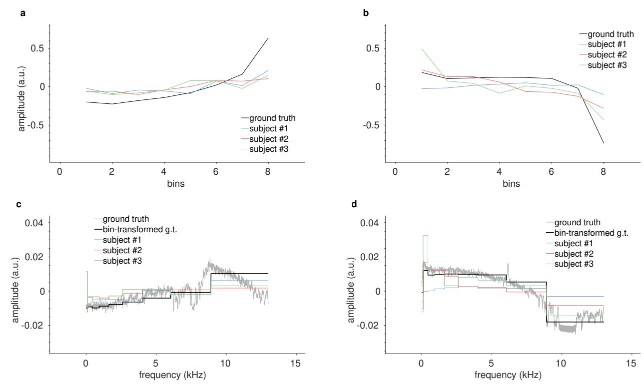

Fig. 3. Reconstructions of the PTS capture many salient features of tinnitus spectra. **(a)** shows buzzing results and **(b)** shows roaring results. In each, the black trace indicates the ground truth, unbinned frequency spectrum of the ATA target signal. The colored traces plot exemplar human subject 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.

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 CITE: comptonLammert2022, roopLammertPreprint, 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. 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] 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.

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

[12] 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.

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

[14] 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.

[15] 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.

[16] 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.

[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] 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.