Reverse Correlation Characterizes More Complete Tinnitus Spectra in Patients

Nelson V. Barnett , Alec Hoyland, Divya A. Chari, Benjamin Parrell , and Adam C. Lammert

Abstract—Goal: **We validate a recent reverse correlation approach to tinnitus characterization by applying it to individuals with clinically-diagnosed tinnitus.** *Methods:* **Two tinnitus patients assessed the subjective similarity of their non-tonal tinnitus percepts and random auditory stimuli. Regression of the responses onto the stimuli yielded reconstructions which were evaluated qualitatively by playing back resynthesized waveforms to the subjects and quantitatively by response prediction analysis.** *Results:* **Subject 1 preferred their resynthesis to white noise; subject 2 did not. Response prediction balanced accuracies were significantly higher than chance across subjects: subject 1: 0.5963, subject 2: 0.6922.** *Conclusion:* **Reverse correlation can provide the foundation for reconstructing accurate representations of complex, non-tonal tinnitus in clinically diagnosed subjects. Further refinements may yield highly similar waveforms to individualized tinnitus percepts.**

***Index Terms—*Reverse correlation, tinnitus, behavioral assay, psychophysics.**

Impact Statement— **Characterization of tinnitus sounds can help clarify the heterogeneous nature of the condition and link etiology to subtypes and treatments.**

I. Introduction

TINNITUS, the internal presence of an externally nonexistent sound, is a pervasive condition [1], [2] that varies greatly in perception, duration, severity, and etiology [3], [4]. Understanding heterogeneity of tinnitus perception is a

longstanding goal in tinnitus research that may open the door to establishing subtypes and endotypes, improving treatment options and outcomes [5], [6], [7], [8], and identifying specific behavioral and neurological mechanisms [9], [10], [11]. Accurate characterization of individualized tinnitus sounds enables a better understanding of this heterogeneity [12], [13], [14].

Reverse correlation (RC) is a widely established method for estimating internal perceptual representations using random stimuli and subjective responses [15], [16], [17]. Recent work has shown that RC works well for reconstructing the frequency spectra of tinnitus-like percepts in healthy controls [18]. Here, we assess the viability of RC for reconstructing tinnitus spectra in clinically-diagnosed tinnitus subjects. Two tinnitus patients completed an RC experiment, after which we conducted reconstruction validation via response prediction analysis [19] and subjective resynthesis assessments. This work demonstrates the effectiveness of RC for estimating tinnitus percept spectra in clinical subjects.
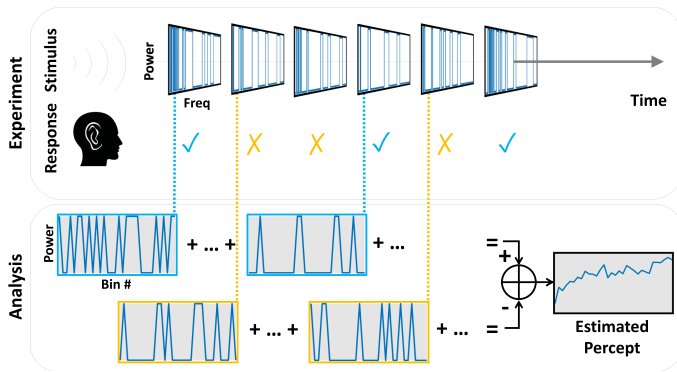
II. Materials and Methods

A. Subjects

Two subjects with clinically diagnosed tinnitus were recruited through the UMass Audiology clinic and referred to WPI for this study. Subjects completed a questionnaire regarding the frequency, severity, and characteristics of their tinnitus. Both subjects rated their tinnitus severity as 5 on a scale from 1 (most mild) to 10 (most severe). Subject 1 reported their tinnitus as "ringing, roaring, and humming." Subject 2 reported hearing "static, hissing, ocean waves, and whooshing." Both subjects had pure-tone hearing thresholds < 60 dB HL between 250 and 2000 Hz (i.e., no severe hearing loss in that frequency range).

B. Stimuli

For each stimulus, the frequency space between $[100, 13000]$ Hz was divided into $b = 32$ Mel-spaced bins. 32 bins provides a higher resolution than the 8 used in [18]. $[6, 16]$ bins were randomly assigned 0 dB, the rest -100 dB. All frequencies were assigned random phase, after which inverse Fourier transform of the spectrum yielded a 500 ms stimulus waveform.[1]

---

[1]Software for the experiments and analysis was written in MATLAB and is freely available at https://github.com/The-Lammert-Lab/tinnitus-reconstruction

Fig. 1. Diagram of the experimental protocol. Subjects listen to a series of random stimuli and respond either "yes" or "no" depending the perceived similarity to their tinnitus. The recorded stimulus-response pairs are used to form an estimate of the subject's tinnitus percept.

C. Experiment

Subjects completed the experiment in an IAC 120-A double wall audiometric booth, listening over Sennheiser HD 560S headphones (Sennheiser electronic GmbH & Co. KG, Wedermark, Germany) at a self-determined comfortable level. Presentation level was not recorded. Subjects were instructed to answer "yes" to stimuli that either "completely sounded like" or "had elements that resembled" their tinnitus (*cf*. Fig. 1). Both subjects completed 10 blocks of 80 trials ($p = 800$) with self-controlled breaks between each block. All stimuli were 500 ms long. Procedures were approved by the UMass IRB.

D. Reconstruction

RC was implemented to reconstruct the internal tinnitus representation given a stimulus matrix $\Psi \in \mathbb{R}^{p \times b}$ and a response vector $y \in \{-1, 1\}^p$, where 1 represents a "yes", -1 a "no", p is the number of trials, and b is the number of bins. RC assumes the subject response model:

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

where $x \in \mathbb{R}^b$ is the internal representation of interest. Assuming uncorrelated stimulus dimensions, inverting this model yields a restricted version of the Normal equation [16]:

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

E. Resynthesis

500 ms resynthesis waveforms were generated from binned reconstructions (2) by rescaling the reconstructions to a 20 dB range, assigning random phase, and applying an inverse Fourier transform.

F. Qualitative Assessment

Subjects were played 500 ms, unlabeled, repeatable sounds of their resynthesis and white noise in a randomized order one time, ranking the statement "This sounds like my tinnitus" on a scale from 1 (strongly disagree) to 5 (strongly agree) after each.

G. Prediction

An established quantitative measure of reconstruction quality is the accuracy of subject response predictions on the basis of those reconstructions [19]. To calculate this, stimuli-response pairs of experiment data were partitioned according to a 5-fold cross-validation scheme. For each fold, a reconstruction, $\hat{x}$, was made by applying (2) to Ψ_{train} and y_{train}. Predicted responses, $\hat{y}$, were generated via (1), where $\Psi = \Psi_{test} - \overline{\Psi}_{test}$, $x = \hat{x} - \overline{\hat{x}}$, and $\overline{\Psi}_{test}$ is the mean of Ψ_{test} in the second dimension.

Balanced accuracy (3) was used as the primary validation metric. Robustness to class imbalance is critical because subjects responded "yes" less often (see Table I).

$$BA = \frac{1}{2}\left(\frac{TP}{P} + \frac{TN}{N}\right), \tag{3}$$

where *TP* and *TN* are the number of true positive and negative predictions, and *P* and *N* are the number of positive and negative responses.

In the absence of repeated exposures to the same stimuli, it can be challenging to establish an upper bound on prediction accuracy, which should be considerably lower than 1 due to subject response error [19]. We therefore contextualize response prediction accuracies by presenting accuracies on the training data, an indication of performance with overfitting, and prediction results from two nonlinear models—locally weighted linear regression (LWLR) and XGBoost [20]—which may indicate an upper bound on performance. LWLR is a natural extension of linear regression that allows for nonlinear modeling, while XGBoost is a highly nonlinear tree-based method. These techniques were chosen as two distinct contrasts to the linearity of RC.

The same cross-validation scheme was run for LWLR. The diagonal weight matrix, W, was found via:

$$W_{i_{j,j}} = \exp\left(-0.5\, u^T H u\right) \tag{4}$$

for $i \in 1, \ldots, p_{test}$ and $j \in 1, \ldots, p_{train}$, where $H = hI_b$ for $h \in \mathbb{R}$ and $u = \Psi_{train_j} - \Psi_{test_i}$. Local regression coefficients were determined using (5) with $\epsilon = 0.01$ and estimations $\hat{z}$ (6) were converted into response predictions via (7).

$$\beta_i = (\Psi_{train}^T W_i \Psi_{train} + \epsilon I_b)^{-1}(\Psi_{train}^T W_i y_{train}) \tag{5}$$

$$\hat{z}_i = \Psi_{test_i} \beta_i \tag{6}$$

$$\hat{y} = sign(\hat{z} + \lambda) \tag{7}$$

Both h and λ were optimized each fold using a development set.

XGBoost using regression, squared error loss, and a gradient boosted tree method was run in Julia over the same cross-validation protocol. Estimations were converted to response predictions using (7) with λ, number of boost rounds, and both l_1 and l_2 regularizers similarly optimized on a development set.
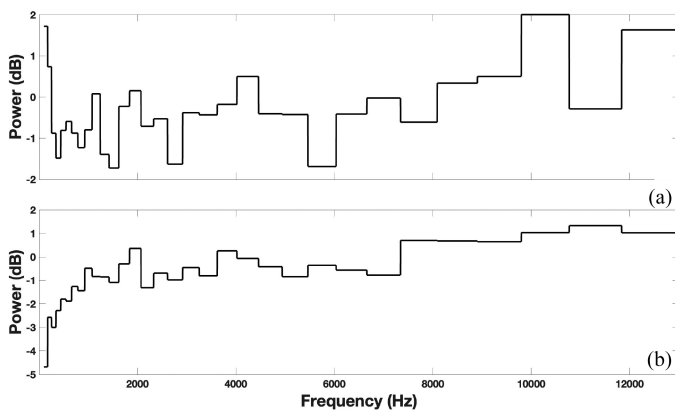
III. RESULTS

Prediction BAs using RC for both subjects were well above 0.5, with RC outperforming both LWLR and XGBoost for subject 2 (Table I). Both RC and LWLR results improve with percent "yes" responses, while XGBoost performs consistently between subjects. Distinct differences are observable between

Subject	% "yes"	Training			Testing		
		RC	LWLR	XGBoost	RC	LWLR	XGBoost
1	13.000	0.6641	0.6961	0.8763	0.5963	0.6080	0.6429
2	39.625	0.7156	0.7646	0.7607	0.6922	0.6752	0.6417

RC is comparable to both LWLR and XGBoost For subject 1 and outperforms both for subject 2. Values generally increase with % "yes" responses, with the exception of XGBoost. Balanced accuracies on the training data are consistently greater than testing but less than 1, indicating minimal Overfitting.

Fig. 2. Reconstructions for (a) subject 1 and (b) subject 2 generated via (2).

the subjects' reconstructions (Fig. 2), indicating subjects were attuned to different elements of the stimuli. Subject 1 rated their resynthesis a 4 and white noise a 2. Subject 2 rated their resynthesis a 1 and white noise a 3 (see Section II-F).

IV. CONCLUSION

This study demonstrates the effectiveness of RC in reconstructing accurate tinnitus spectra from individuals suffering from tinnitus. Reconstruction accuracy is well supported by response prediction analysis (Table I), for which RC performs close to an upper bound estimated by two nonlinear models. Furthermore, BA on training data shows minimal overfitting. Altogether, this indicates that the majority of the performance loss is likely due to irreducible error in the subjects' responses, rather than a lack of appropriately modeling the frequency domain.

Inherent vagueness in how subjective descriptions of tinnitus relate to the frequency spectrum is a fundamental challenge in tinnitus research. The subjective descriptions provided by the subjects in this study were non-overlapping, which corresponds appropriately with the presence of differences in the reconstructed frequency spectra. The reconstructed tinnitus spectra in Fig. 2 represent the first of their kind for individuals experiencing non-tonal, spectrally complex tinnitus, which may ultimately prove useful for understanding individual variation in tinnitus, including the subjective descriptions of patients.

The heterogeneity of tinnitus implies that two subjects cannot adequately represent the range of tinnitus experiences. In a larger sample population, it will be critical to minimize subject response error, and increase the proportion of "yes" responses from subjects, to reliably reconstruct accurate representations of complex tinnitus spectra. Methodological refinement to the stimuli and instructions may significantly improve these aspects, though we note the distribution and consistency of responses reported here is not unlike those seen in other RC experiments [16]. While subjective resynthesis evaluations were mixed, high response prediction BAs point to a sound experimental protocol and reconstruction method. The RC reconstruction technique results in ambiguities in the dynamic range and energy distributions within bins. Resyntheses and subjective ratings may be greatly improved by developing methods to disambiguate these reconstruction features.

REFERENCES

[1] A. McCormack, M. Edmondson-Jones, S. Somerset, and D. Hall, "A systematic review of the reporting of tinnitus prevalence and severity," *Hear. Res.*, vol. 337, pp. 70–79, 2016.

[2] 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 Surg.*, vol. 142, no. 10, pp. 959–965, Oct. 2016.

[3] C. R. Cederroth et al., "Editorial: Towards an understanding of tinnitus heterogeneity," *Front. Aging Neurosci.*, vol. 11, pp. 53–53, 2019.

[4] J. P. Simoes et al., "Multidisciplinary tinnitus research: Challenges and future directions from the perspective of early stage researchers," *Front. Aging Neurosci.*, vol. 13, 2021, Art. no. 647285.

[5] L. E. Roberts, G. Moffat, M. Baumann, L. M. Ward, and D. J. Bosnyak, "Residual inhibition functions overlap tinnitus spectra and the region of auditory threshold shift," *J. Assoc. Res. Otolaryngol.*, vol. 9, no. 4, pp. 417–435, 2008.

[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," *Proc. Nat. Acad. Sci.*, vol. 107, no. 3, pp. 1207–1210, Jan. 2010.

[7] 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," *Hear. Res.*, vol. 269, no. 1, pp. 95–101, Oct. 2010.

[8] A. Stein et al., "Inhibition-induced plasticity in tinnitus patients after repetitive exposure to tailor-made notched music," *Clin. Neurophysiol.*, vol. 126, no. 5, pp. 1007–1015, May 2015.

[9] A. Mohan, S. L. Leong, D. De Ridder, and S. Vanneste, "Symptom dimensions to address heterogeneity in tinnitus," *Neurosci. Biobehavioral Rev.*, vol. 134, 2022, Art. no. 104542.

[10] E. Genitsaridi, D. J. Hoare, T. Kypraios, and D. A. Hall, "A review and a framework of variables for defining and characterizing tinnitus subphenotypes," *Brain Sci.*, vol. 10, no. 12, 2020, Art. no. 938.

[11] T. Pan, R. S. Tyler, H. Ji, C. Coelho, A. K. Gehringer, and S. A. Gogel, "The relationship between tinnitus pitch and the audiogram," *Int. J. Audiol.*, vol. 48, no. 5, pp. 277–294, 2009.

[12] A. Norena, C. Micheyl, S. Chéry-Croze, and L. Collet, "Psychoacoustic characterization of the tinnitus spectrum: Implications for the underlying mechanisms of tinnitus," *Audiol. Neurotol.*, vol. 7, no. 6, pp. 358–369, 2002.

[13] D. Vajsakovic, M. Maslin, and G. D. Searchfield, "Principles and methods for psychoacoustic evaluation of tinnitus," in *The Behavioral Neuroscience of Tinnitus*. Berlin, Germany: Springer, Feb. 2021, pp. 419–451.

[14] J. A. Henry, "'Measurement' of tinnitus," *Otol. Neurotol.*, vol. 37, no. 8, pp. e276–e285, 2016.

[15] A. Ahumada and J. Lovell, "Stimulus features in signal detection," *J. Acoustical Soc. Amer.*, vol. 49, no. 6B, pp. 1751–1756, Jun. 1971.

[16] F. Gosselin and P. G. Schyns, "Superstitious perceptions reveal properties of internal representations," *Psychol. Sci.*, vol. 14, no. 5, pp. 505–509, Sep. 2003.

[17] W. O. Brimijoin, M. A. Akeroyd, E. Tilbury, and B. Porr, "The internal representation of vowel spectra investigated using behavioral response-triggered averaging," *J. Acoustical Soc. Amer.*, vol. 133, no. 2, pp. EL118–EL122, Feb. 2013.

[18] A. Hoyland et al., "Reverse correlation uncovers more complete tinnitus spectra," *IEEE Open J. Eng. Med. Biol.*, vol. 4, pp. 116–118, 2023.

[19] P. Neri and D. M. Levi, "Receptive versus perceptive fields from the reverse-correlation viewpoint," *Vis. Res. (Oxford)*, vol. 46, no. 16, pp. 2465–2474, 2006.

[20] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2016, pp. 785–794.