Supplementary Materials

Compressed Sensing for Characterization of Tinnitus

Alec Hoyland, Nelson Barnett, Benjamin W. Roop, Adam C. Lammert

This document contains a description of additional experiments to find a performant stimulus generation method and a discussion of sparsity in tinnitus signals. The code for all experiments is freely available at https://github.com/alec-hoyland/tinnitus-project and the data are available upon request.

I. Stimulus Generation

In the context of this paper, a stimulus generation method is a process that generates a random waveform that is:

1) auditorally-distinguishable
2) statistically uncorrelated, and
3) similar to tinnitus percepts.

Additionally, compressed sensing requires that the matrix of stimuli should satisfy the restricted isometry property, which many random matrices do with high probability (*e.g.* Gaussian random matrices) [1], [2].

A. Auditorally-Distinguishable Stimuli

We used mel-frequency binning to ensure that our stimuli were auditorally-distinguishable. The mel scale is a perceptual scale of pitches judged by listeners to be equal in distance from one to another (Fig. 1) [3].

The formula

$$m = 2595 \log_1 0 \left(1 + \frac{f}{700}\right) \tag{1}$$

converts f Hz to m mels.

Furthermore, to reduce the system complexity by more than 80x, we implement tonotopic binning, where the frequency scale is binned along 100 equally mel-spaced bins (Fig. 2).

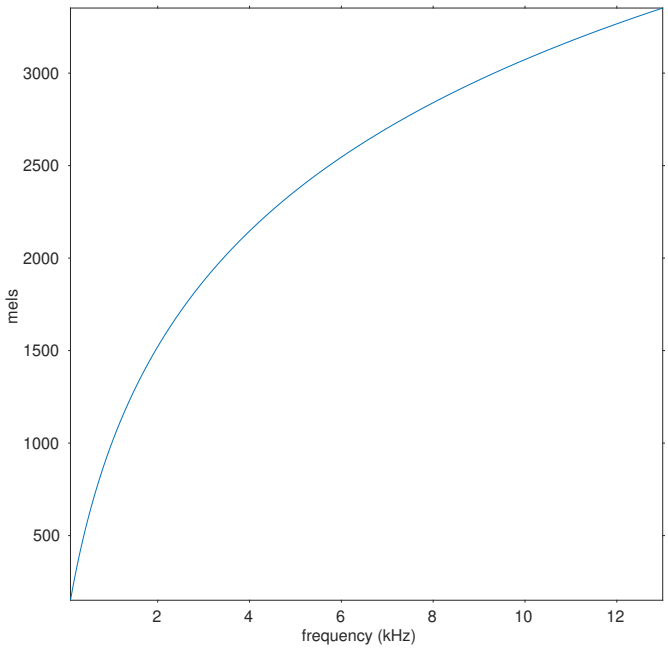

Fig. 1. The relationship between Hz and mels is logarithmic.

II. Hyperparameter Sweep

To determine a suitable stimulus generation method, we performed a hyperparameter sweep over nine different stimulus generation methods. We used an *in-silico* model of the experiment, with an artificial subject, but matched against American Tinnitus Association (ATA) tinnitus examples, as in the human experiment.

The methods are as follows:

1) *Bernoulli*: A binned method in which each tonotopic bin has a probability *bin_prob* of being set to 0 dB otherwise it is set to -20 dB.
2) *Brimijoin*: A binned method inspired by [4]. Each bin is assigned an amplitude value chosen from a uniform distribution of discrete values on the interval $[-20, 0]$. We used 6 steps, which is consistent with the original paper.
3) *Gaussian Noise No Bins*: The amplitude of each frequency was determined by Gaussian noise with known mean and variance, which are hyperparameters *amplitude_mean* and *amplitude_var*.
4) *Gaussian Noise*: A binned method where the bin amplitudes were determined by a Gaussian random variable with known mean and variance, *e.g.* $\mathcal{N}(amplitude\_mean$ and $amplitude\_var)$.
5) *Gaussian Prior*: A binned method where the number of filled bins was set by round$(\mathcal{N}(n_bins_filled_mean, n_bins_filled_var))$, and that many bins were filled randomly at 0 dB (unfilled bins were set to -20 dB).
6) *Power Distribution*: A binned method where the amplitude of each bin is drawn from a distribution matching the histogram of amplitudes of the American Tinnitus Association tinnitus examples from their website.
7) *Uniform Noise No Bins*: The amplitude associated with each frequency is determined by a uniform random variable on the interval $[-20, 0]$ dB.
8) *Uniform Noise*: The amplitude associated with each tonotopic bin is determined by a uniform random variable on the interval $[-20, 0]$ dB.
9) *Uniform Prior*: The number of filled bins is determined by drawing an integer from the discrete interval

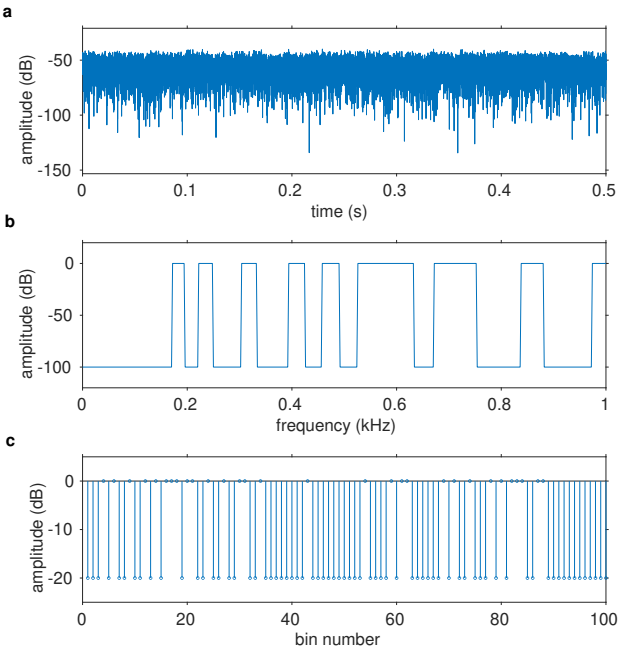

Fig. 2. Example stimulus. **(a)** shows the waveform of the stimulus, **(b)** the frequency spectrum from which the waveform was generated, and **(c)** the 100-dimensional bin representation.

$[min\_bins, max\_bins] \in \mathbb{Z}^+$. Then, that many bins are set to 0 dB, and all other bins are set to -20 dB. *min_bins* and *max_bins* are hyperparameters.

9 methods and 8 hyperparameters were varied in the hyperparameter sweep grid search (Table I). Methods were evaluated *in-silico* using the model:

$$y = \mathrm{sgn}(\Psi \bar{x}) \tag{2}$$

where $y \in \mathbb{R}^n$ is the binary response vector, $\Psi \in \mathbb{R}^{n \times d}$ is the stimulus matrix, and $\bar{x} \in \mathbb{R}^d$ is the true target signal (the spectrum of the ATA tinnitus example).

Each stimulus generation method, with different hyperparameters, was evaluated for target signals: buzzing, electric, roaring, screeching, static, and tea kettle.

HYPERPARAMETERS

Name	Values	Unit
n_bins	10, 30, 100, 200, 300	n.d.
bin_prob	0.1, 0.3, 0.5, 0.8	n.d.
amplitude_mean	-35	dB
amplitude_var	5, 10, 20	$\sqrt{\mathrm{dB}}$
n_bins_filled _mean	1, 3, 10, 20, 30	n.d.
n_bins_filled_var	0.01, 1, 3, 10	n.d.
min_bins	1, 3, 10, 20, 30	n.d.
max_bins	10, 20, 30, 50	n.d.

TABLE I. Hyperparameter values tested in the parameter sweep. Not all methods use all hyperparameters, and some combinations of hyperparameters are invalid (*e.g.* when *min_bins* > *max_bins*). n.d. means "non-dimensional" and refers to a unitless number.

Reconstructions were computed using linear regression and compressed sensing (*cf.* the main paper) and reconstruction accuracy was measured using Pearson's r^2 vs. the target signal.

Fine-tuning of parameters proceeded iteratively, bearing in mind both the performance of the stimulus generation methods in the *in-silico* experiment as well as computational efficiency, qualitative distinguishability of stimuli from each other by a human listener, and viability of compressed sensing as an efficiency enhancement. This process resulted in *n_bins* being fixed to 100 and the number of trials in the *in-silico* experiment being set to 1000.

HYPERPARAMETER SWEEP RESULTS

Stimulus Type	*n_bins_filled_mean*	*n_bins_filled_var*	*min_bins*	*max_bins*	*CS_r²*	*LR_r²*
Gaussian Prior	20	1			0.83	0.86
Uniform Prior			30	30	0.82	0.86
Uniform Prior			20	20	0.82	0.86
Gaussian Prior	20	1			0.82	0.86
Gaussian Prior	12	0.01			0.81	0.83

TABLE II. Top-5 stimulus generation methods by r^2 score averaged over buzzing, electric, roaring, screeching, and static tinnitus examples, with $n\_bins = 100$ and $n\_trials = 1000$.

Table II shows the top 5 best-performing methods in the *in-silico* experiment. Gaussian Prior and Uniform Prior outperform all other methods, with 10-30% of bins being filled. The variance of the number of filled bins for top-performing methods is extremely low. Due to rounding to the nearest integer, a Gaussian Prior method with very low variance converges to a Uniform Prior method with equal min and max bin parameters. In the main paper, we used the Uniform Prior method with 30 filled bins ($min\_bins = max\_bins = 30$).

III. SPARSITY OF ATA TINNITUS EXAMPLES

A necessary condition for compressed sensing to produce reconstructions results beyond the Shannon-Nyquist limit is that the signal to be reconstructed is sparse in some basis [5]. A signal is considered to be at least k sparse when it has at most k nonzero elements. A signal is considered k-sparse under some basis, when a basis-transformed version of that signal is k-sparse. Here we show that the 8,193-dimensional ATA example tinnitus spectra are sparse under the discrete cosine basis. The signal representation in the basis is produced by computing its discrete cosine transform (DCT). We computed the DCT of each ATA example spectrum and set all but the top 32 DCT-frequencies to zero before taking the inverse transform back to the Fourier domain. Figure 3 shows that while some high-frequency information was lost, that the 32-sparse signal preserves the salient information of the signal. The tonotopic bin representation is also sparse under the discrete cosine basis (Fig. 4). The sparsity value $k = 32$ was chosen by examining the decay of the sorted power spectrum of the tinnitus Fourier spectra in the discrete cosine basis (Fig. 5). We see that the power of the signal rapidly decays after only a few DCT-frequencies.

REFERENCES

[1] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, May 2008.

[2] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

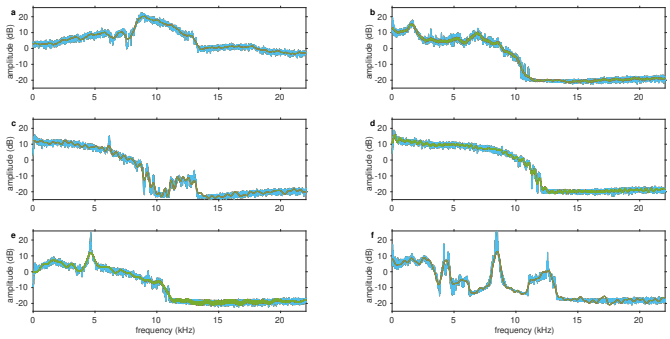

Fig. 3. Tinnitus signals are approximately sparse. Each subfigure plots the true spectrum of an ATA tinnitus example in blue (8,193 nonzero dimensions) and a 32-sparse (32 non-zero dimensions) approximation in green. The sparse signal is a low-pass filtered version of the original signal that retains the salient dynamics. **(a)** through **(f)** correspond to buzzing, electric, roaring, static, tea kettle, and screeching ATA examples respectively.

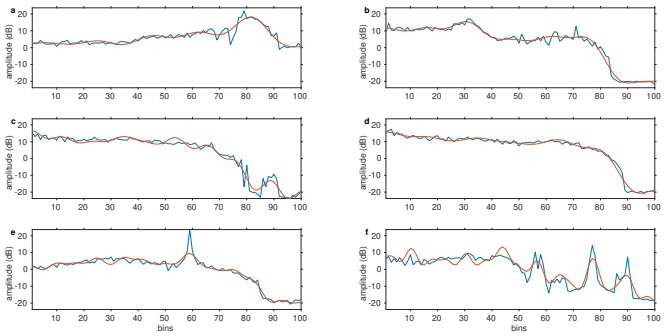

Fig. 4. Tonotopic bin representations of tinnitus signals are approximately sparse in the DCT basis. Each subfigure plots the tonotopic bin representation of an ATA tinnitus spectrum in blue (100 nonzero dimensions) and the bin representation of a 10-sparse approximation in red (10 nonzero dimensions). The sparse signal is a low-pass filtered version of the original signal that retains the salient dynamics. **(a)** through **(f)** correspond to buzzing, electric, roaring, static, tea kettle, and screeching ATA examples respectively.

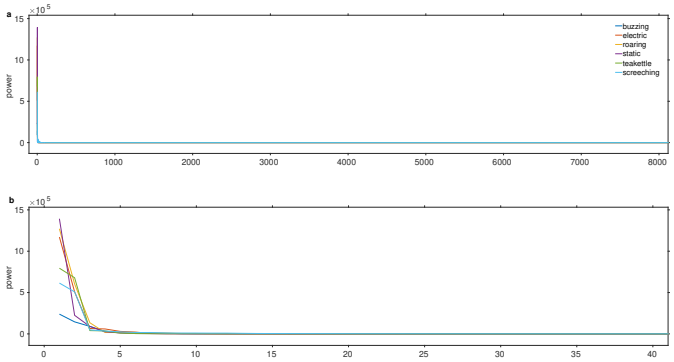

Fig. 5. DCT-transformed tinnitus spectra have high power in a few frequencies and low power elsewhere. **a** shows the power spectrum of DCT-transformed ATA tinnitus example Fourier spectra. **b** shows the highest 40 DCT frequencies in ranked order, illustrating a sharp decay to zero, indicating sparsity of tinnitus spectra in the DCT basis.

[3] L. S. Hamilton, Y. Oganian, and E. F. Chang, "Topography of speech-related acoustic and phonological feature encoding throughout the human core and parabelt auditory cortex," p. 2020.06.08.121624, Jun. 2020.

[4] 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.

[5] M. Taghouti, "Compressed Sensing," in *Computing in Communication Networks*, F. H. P. Fitzek, F. Granelli, and P. Seeling, Eds. Academic Press, Jan. 2020, pp. 197–215.