

## Supplemental Material S1. Detailed description of the statistical models.

The pure tone threshold measured on the  $i^{th}$  ear at audiometric frequency  $f_{PTT}$  is denoted by  $y_i^{f_{PTT}}$ , where  $f_{PTT} = \{0.25, 0.5, 1, 2, 3, 4, 6, 8, 9, 10, 11.2, 12.5, 14, 16 \text{ kHz}\}$ . These pure tone thresholds at individual frequencies can be combined to make up an audiogram, defined as  $\mathbf{y}_i$ , which is a vector with 14 elements corresponding to the pure tone thresholds at each frequency. Bold-faced symbols indicate vectors or matrices. The audiograms are assumed to be multivariate normal random variables such that

$$\mathbf{y}_i \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}) \quad (1)$$

where  $\boldsymbol{\Sigma}$  is a first-order auto-regressive covariance matrix with heterogeneous diagonal elements. Defining  $\boldsymbol{\Sigma}$  in this manner allows the residual variability to differ according to audiometric frequency. The parameter  $\boldsymbol{\mu}_i$  represents the predicted audiogram for the  $i^{th}$  ear.

The effects of the DP-gram on  $\boldsymbol{\mu}_i$  are modeled as a functional predictor. The  $i^{th}$  ear's DP-gram, denoted by  $\mathbf{x}_i$ , is a vector of 11 elements composed of DPOAE levels measured at 11  $f_2$  primary frequencies  $f_{DPOAE}$ , where  $f_{DPOAE} = \{1, 1.2, 1.5, 2, 2.5, 3, 4, 5, 6, 8, 10 \text{ kHz}\}$ . Conceptually, the goal is to predict the  $i^{th}$  ear's audiogram  $\mathbf{y}_i$  with the same ear's DP-gram  $\mathbf{x}_i$ . This is accomplished by identifying frequency-specific weight functions that transform the DP-gram into a predicted pure tone threshold for each audiometric frequency (see **Figure 1**). The statistical problem is that of estimating the weight functions given the available data.

The simplest way to predict the audiogram is to use a model where  $\boldsymbol{\mu}_i$  is equal to the mean pure tone threshold at each audiometric frequency, or rather, more precisely, a vector of independent model intercepts at each frequency. Because the predicted audiogram is simply the sample mean audiogram, the weight function for this model is equal to zero across all  $f_{DPOAE}$  and  $f_{PTT}$  and the predicted audiogram will be identical for all ears regardless of their DPOAE levels. This simple model was fit to the data to serve as a baseline for comparison with a more elaborate model that includes DPOAEs and is referred to as the “mean audio model.”

Another, intuitively appealing model to consider has independent regression coefficients for each DPOAE level on each  $f_{PTT}$ . This results in a model with 154 independent regression coefficients (11  $f_{DPOAE}$  effects  $\times$  14  $f_{PTT} = 154$ ). Although this model should predict the sample data quite well, it will generalize poorly to other samples because it will tend to overfit the data due to high correlations between DPOAE levels across  $f_{DPOAE}$ . The result of this type of model is a set of weight functions, one per audiometric frequency, that appears jagged across  $f_{DPOAE}$  with a high degree of uncertainty.

An alternative approach is to treat both the audiogram and the DP-gram as continuous functions sampled at discrete frequencies. This is sensible because the audiometric test frequencies and the DPOAE  $f_2$  primaries were chosen by convention. It is possible to obtain data from additional test frequencies if time and equipment allows. With the  $i^{th}$  ear's DP-gram now written as a function of  $f_{DPOAE}$ , i.e.  $x_i(f_{DPOAE})$ , the predicted pure tone threshold  $\mu_i^{f_{PTT}}$  at audiometric frequency  $f_{PTT}$  can be represented as

$$\mu_i^{f_{PTT}} = \int x_i(f_{DPOAE}) \cdot \beta(f_{PTT}, f_{DPOAE}) df_{DPOAE} \quad (2)$$

This equation indicates that integration occurs over the  $f_2$  frequencies of the DP-gram, weighted according to the regression coefficient function  $\beta$ . These integrals comprise the mean vector  $\mu_i$  in equation (1). Low-rank smoothing begins by representing each frequency in the audiogram or DP-gram as a collection of nine basis functions, rather than as 14 or 11 discrete frequencies, respectively. This approach reduces the number of unknown coefficients to 81 compared to the 154 for the independent coefficients model described in the previous paragraph and by imposing a “roughness” penalty on the coefficient surface, it induces smoothing across frequency for the regression coefficients. The notation for this model is quite complex, so please refer to Scheipl et al. (2015, section 1.1) and Wood et al. (2012) for more detail. This model is referred to as the “full model.”