Supplement to
"Spectral-based noncentral F mixed effect models, with application to otoacoustic emissions"

Lai Wei[1], Peter F. Craigmile[2,*], Wayne M. King[3]

[1] Biostatistics Center, The Ohio State University, Columbus. OH 43210.

[2] Dept. of Statistics, The Ohio State University, Columbus. OH 43210.

[3] Signal Processing and Communications, The Mathworks, Inc., 3 Apple Hill, Natick, MA 01760

*email: pfc@stat.osu.edu

The MCMC algorithm used to fit the Bayesian DPOAE noncentral F mixed model

We use $\boldsymbol{\theta} \setminus \{\kappa\}$ to denote the collection of parameters contained in $\boldsymbol{\theta}$, but excluding the parameter κ. The MCMC algorithm below assumes that $K_j = K$ for all subjects j. (More carefully subsetting and indexing is needed in the unbalanced case.)

Update $\{\log r_{j,k}\}$: Fix a subject $j = 1, \ldots, J$, and let $\mathbf{log}\, \boldsymbol{r}_{j,\bullet} = (\log r_{j,1}, \ldots, \log r_{j,K})^T$. Then

$$\pi(\mathbf{log}\, \boldsymbol{r}_{j,\bullet} | \boldsymbol{z}, \boldsymbol{\theta} \setminus \{\mathbf{log}\, \boldsymbol{r}_{j,\bullet}\}) \;\propto\; \prod_{k=1}^{K} \left[\prod_{l=1}^{L} f_{Z_{j,k,l}}\big(z_{j,k,l} \mid e^{N \Delta r_{j,k}}\big) \right] n\big(\log r_{j,k} | \alpha_k + G_j \beta_k, \tau_k^2\big) \,.$$

We use a Metropolis-Hastings symmetric random walk update. Supposing we are at $\mathbf{log}\, \boldsymbol{r}_{j,\bullet}$, we propose $\mathbf{log}\, \boldsymbol{r}_{j,\bullet}^{new}$ from a K-variate normal with mean $\mathbf{log}\, \boldsymbol{r}_{j,\bullet}$ and covariance Σ_j, for some $K \times K$ positive definite matrix Σ_j (in practice, we base Σ_j on the estimated covariance matrix of the maximum likelihood estimate of $\mathbf{log}\, \boldsymbol{r}_{j,\bullet}$, calculated using only the data for subject j, scaled to obtain an acceptance probability of around 0.4). Then we accept the new value, $\mathbf{log}\,\boldsymbol{r}_{j,\bullet}^{new}$, with probability $\min(e^q,\, 1)$, where

$$
\begin{aligned}
q \;=\; & \left\{\sum_{k=1}^{K}\sum_{l=1}^{L}\log f_{Z_{j,k,l}}\big(z_{j,k,l} \mid e^{N\Delta r_{j,k}^{new}}\big) + \sum_{k=1}^{K}\log n\big(\log r_{j,k}^{new}|\alpha_k + G_j\beta_k, \tau_k^2\big)\right\} - \\
& \left\{\sum_{k=1}^{K}\sum_{l=1}^{L}\log f_{Z_{j,k,l}}\big(z_{j,k,l} \mid e^{N\Delta r_{j,k}}\big) + \sum_{k=1}^{K}\log n\big(\log r_{j,k}|\alpha_k + G_j\beta_k, \tau_k^2\big)\right\};
\end{aligned}
$$

otherwise we remain at $\mathbf{log}\,\boldsymbol{r}_{j,\bullet}$.

Update $\{\alpha_k\}$ and $\{\beta_k\}$: Fixing a $k = 1, \ldots, K$ we sample α_k and β_k jointly, conditional on the data and other parameters. First note that

$$
\pi(\alpha_k, \beta_k|\boldsymbol{z}, \boldsymbol{\theta} \setminus \{\alpha_k, \beta_k\}) \;\propto\; \left\{\prod_{j=1}^{J} n\big(\log r_{j,k}|\alpha_k + G_j\beta_k, \tau_k^2\big)\right\} n\big(\alpha_k|\mu_\alpha, \sigma_\alpha^2\big)\, n\big(\beta_k|\mu_\beta, \sigma_\beta^2\big)\,.
$$

Let X be a $J \times 2$ design matrix with first column all ones, and second column G_j ($j = 1, \ldots, J$). Then $\boldsymbol{\beta}_{\bullet,k} = (\beta_{1,k}, \ldots, \log r_{j,k})^T$, conditional on α_k, β_k and τ_k^2, is J-variate normal with mean $X(\alpha_k, \beta_k)^T$ and covariance $\tau_k^2 I_J$, where I_J is the $J \times J$ identity matrix. This is a Bayesian regression model. Hence, letting

$$
V \;=\; \begin{bmatrix} \sigma_\alpha^2 & 0 \\ 0 & \sigma_\beta^2 \end{bmatrix},
$$

our sample from $(\alpha_k, \beta_k)^T$, conditional on the data and other parameters, is a bivariate normal draw with a mean $\Sigma^{-1}\boldsymbol{c}$ and covariance Σ^{-1} where $\Sigma = X^TX/\tau_k^2 + V^{-1}$ and $\boldsymbol{c} = X^T\boldsymbol{\beta}_{\bullet,k}/\tau_k^2 + V^{-1}(\mu_\alpha, \mu_\beta)^T$.

Update $\{\tau_k^2\}$: For each $k = 1, \ldots, K$ we have that

$$\pi(\tau_k^2|\boldsymbol{z}, \boldsymbol{\theta} \setminus \{\tau_k^2\}) \;\propto\; \left\{\prod_{j=1}^{J} n\big(\log r_{j,k}|\alpha_k + G_j\beta_k, \tau_k^2\big)\right\} ig\big(\tau_k^2|s_\tau, r_\tau\big)\,,$$

leading us to sample τ_k^2, conditional on the data and other parameters, from an inverse gamma distribution with shape s and rate r, where

$$s \;=\; s_{\sigma,\alpha} + J/2, \quad \text{and} \quad r \;=\; r_{\sigma,\alpha} + \frac{1}{2}\sum_{j=1}^{J}(\log r_{j,k} - \alpha_k - G_j\beta_k)^2.$$

Update μ_α: (The update for μ_β is similar.)

$$\pi(\mu_\alpha|\boldsymbol{z}, \boldsymbol{\theta} \setminus \{\mu_\alpha\}) \;\propto\; \left\{\prod_{k=1}^{K} n\big(\alpha_k|\mu_\alpha, \sigma_\alpha^2\big)\right\} n(\mu_\alpha|m_\alpha, v_\alpha)\,,$$

and hence we sample μ_α, conditional on the data and other parameters, from a normal distribution with mean m/p and variance $1/p$ where

$$m \;=\; \sum_{k=1}^{K} \alpha_k/\sigma_\alpha^2 + m_\alpha/v_\alpha, \quad \text{and} \quad p \;=\; K/\sigma_\alpha^2 + 1/v_\alpha.$$

Update σ_α^2: (The update for σ_β^2 is similar.)

$$\pi(\sigma_\alpha^2|\boldsymbol{z}, \boldsymbol{\theta} \setminus \{\sigma_\alpha^2\}) \;\propto\; \left\{\prod_{k=1}^{K} n\big(\alpha_k|\mu_\alpha, \sigma_\alpha^2\big)\right\} ig\big(\sigma_\alpha^2|s_{\sigma,\alpha}, r_{\sigma,\alpha}\big)\,,$$

and hence we sample σ_α^2, conditional on the data and other parameters, from an inverse gamma distribution with shape s and rate r, where

$$s \;=\; s_\tau + K/2, \quad \text{and} \quad r \;=\; r_\tau + \frac{1}{2}\sum_{k=1}^{K}(\alpha_k - \mu_\alpha)^2.$$