A Sate-Space Model for Decoding Auditory Attentional Modulation from MEG in a Competing-Speaker Environment

Sahar Akram[1,2], Jonathan Z. Simon[1,2,3], Shihab Shamma[1,2], Behtash Babadi[1,2]
[1]Department of Electrical and Computer Engineering, [2] Institute for Systems Reasearch,
[3] Department of Biology, University of Maryland, College Park, MD
sakram@umd.edu, jzsimon@umd.edu, sas@umd.edu, behtash@umd.edu

THIRTY YEARS OF SYSTEMS RESEARCH EXCELLENCE

INSTITUTE FOR SYSTEMS RESEARCH
A. JAMES CLARK SCHOOL OF ENGINEERING

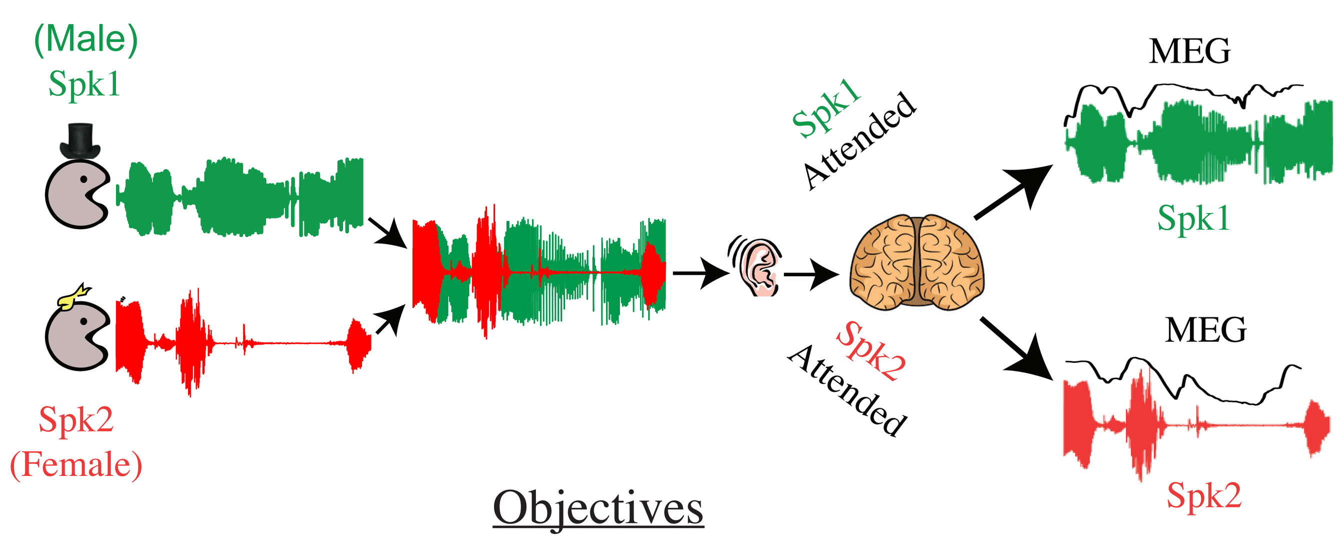
The Cocktail Party Problem

Speech Segregation:

Identifying and tracking a target speaker, corrupted by acoustic interference [1].

Neural Activity at the Cortical level:

Strongly modulated by low-frequency temporal modulations (envelope) of the attended target speaker [2].



Objectives

Existing Techniques:

1) Full spectrotemporal features as covariates
2) Low temporal Resolution ~ Minutes

Our Contribution:

1) Parsimonious use of covariates
2) High temporal resolution ~ Seconds
3) Scalability

The Proposed Model

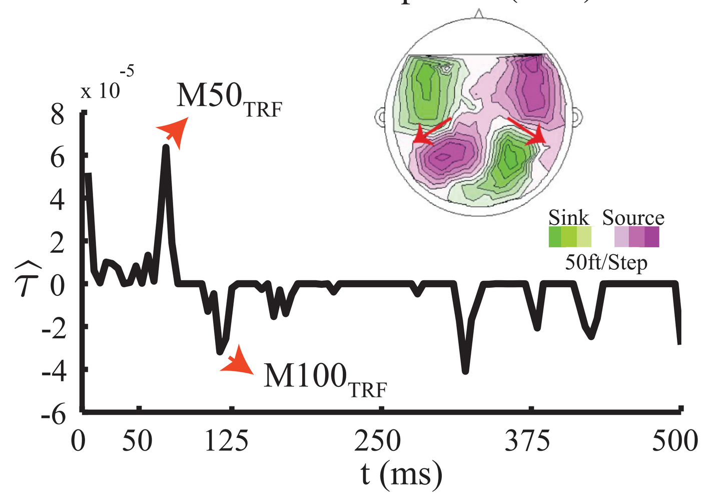
Forward Model:

$$y_t = \tau_t * E_t + v_t$$

Auditory MEG (y_t), Speech envelope (E_t), Noise (v_t), Temporal receptive field (τ_t)

$$\widehat{\boldsymbol{\tau}} = \underset{\boldsymbol{\tau}}{\operatorname{argmin}} \sum_{t=1}^{T} \|y_t - \boldsymbol{\tau}' \boldsymbol{E}_t\|_2^2 + \gamma \|\boldsymbol{\tau}\|_1$$

Magnetic field map of the auditory MEG component (DSS) [3]



Attention Model:

$$n_k = \begin{cases} 0 & \text{attending to Spk2} \\ 1 & \text{attending to Spk1} \end{cases} \qquad n_k \sim \text{Bernoulli}(q_k)$$

$$q_k = \text{logit}^{-1}(z_k) := \frac{\exp(z_k)}{1+\exp(z_k)}$$

$$z_k = z_{k-1} + w_k \qquad w_k \sim \mathcal{N}(0, \eta_k) \qquad \eta_k \sim \text{Gamma}(\alpha, \beta)$$

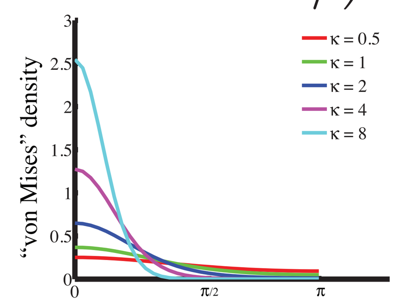
Inverse Model:

$$n_k \xrightarrow{1} \theta_{1,k} := \arccos\left(\left\langle \frac{\mathbf{y}_k}{\|\mathbf{y}_k\|_2}, \frac{\boldsymbol{\tau}^{a'}\mathbf{E}_{1,t} + \boldsymbol{\tau}^{u'}\mathbf{E}_{2,t}}{\|\boldsymbol{\tau}^{a'}\mathbf{E}_{1,t} + \boldsymbol{\tau}^{u'}\mathbf{E}_{2,t}\|_2} \right\rangle\right)$$

$$n_k \xrightarrow{0} \theta_{2,k} := \arccos\left(\left\langle \frac{\mathbf{y}_k}{\|\mathbf{y}_k\|_2}, \frac{\boldsymbol{\tau}^{a'}\mathbf{E}_{2,t} + \boldsymbol{\tau}^{u'}\mathbf{E}_{1,t}}{\|\boldsymbol{\tau}^{a'}\mathbf{E}_{2,t} + \boldsymbol{\tau}^{u'}\mathbf{E}_{1,t}\|_2} \right\rangle\right)$$

$$\theta_{i,k} \sim \mathscr{VM}(\kappa_i) := \frac{e^{\kappa_i \cos(\theta_{i,k})}}{\pi I_0(\kappa_i)} \qquad \kappa_i \sim \frac{e^{\kappa_i c}}{I_0(\kappa_i)^d}, \quad i = 1,2$$

von Mises Distribution



The Inverse Solution

The Inverse Problem: Estimating $\Omega := \{\kappa_1, \kappa_2, \{z_k\}_{k=1}^{K}, \{\eta_k\}_{k=1}^{K}\}$, given the observed data $\{\theta_{i,k,r}\}_{i,k,r=1}^{2,T,R}$ from R trials.

MAP Estimate: Maximize

$$\log p\left(\Omega \middle| \{\theta_{i,k,r}\}_{i,k,r=1}^{2,K,R}\right) = \sum_{r,k=1}^{R,K} \log\left[\frac{q_k}{\pi I_0(\kappa_1)} \exp(\kappa_1 \cos(\theta_{1,k,r})) + \frac{1-q_k}{\pi I_0(\kappa_2)} \exp(\kappa_2 \cos(\theta_{2,k,r}))\right]$$
$$+ \left[(\kappa_1+\kappa_2)c_0 d - d\left(\log I_0(\kappa_1) + \log I_0(\kappa_2)\right)\right]$$
$$- \sum_{r,k=1}^{R,K} \left\{\frac{1}{2\eta_k}(z_k - z_{k-1})^2 + \frac{1}{2}\log \eta_k + (\alpha+1)\log \eta_k + \frac{\beta}{\eta_k}\right\} + \text{cst.}$$

Convex, but highly non-linear and coupled in time.

Efficient solution: two nested EM algorithms

» Outer EM iteration ℓ:
  E-Step: Compute $\mathbb{E}^{(\ell)}\{n_{k,r}\} := \mathbb{E}\left\{n_{k,r} \middle| \{\theta_{i,k,r}\}_{i,k,r=1}^{2,K,R}, \Omega^{(\ell)}\right\}$ *
  M-Step: Update $\kappa_1^{(\ell+1)}$ and $\kappa_2^{(\ell+1)}$ **
  » Inner EM iteration m:
    E-Step: Compute $\bar{z}_{k|K}^{(\ell+1,m)} := \mathbb{E}\left\{z_k \middle| \mathbb{E}^{(\ell)}\{n_{k,r}\}_{k,r=1}^{K,R}\right\}$ ***
    M-Step: Update $\eta_k^{(\ell+1,m+1)}$ ****
  » end
» end

Decoupled in time with tractable non-linear operations

*

$$\mathbb{E}^{(\ell)}\{n_{k,r}\} = \frac{\frac{q_k^{(\ell)}}{\pi I_0\left(\kappa_1^{(\ell)}\right)} \exp\left(\kappa_1^{(\ell)} \cos(\theta_{1,k,r})\right)}{\frac{q_k^{(\ell)}}{\pi I_0\left(\kappa_1^{(\ell)}\right)} \exp\left(\kappa_1^{(\ell)} \cos(\theta_{1,k,r})\right) + \frac{1-q_k^{(\ell)}}{\pi I_0\left(\kappa_2^{(\ell)}\right)} \exp\left(\kappa_2^{(\ell)} \cos(\theta_{2,k,r})\right)}$$

**

$$\kappa_i^{(\ell+1)} = A^{-1}\left(\frac{c_0 d + \sum_{r,k=1}^{R,K} \varepsilon_{i,k,r}^{(\ell)} \cos(\theta_{i,k,r})}{d + \sum_{r,k=1}^{R,K} \varepsilon_{i,k,r}^{(\ell)}}\right), \quad \varepsilon_{i,k,r}^{(\ell)} = \begin{cases} \mathbb{E}^{(\ell)}\{n_{k,r}\} & i=1 \\ 1 - \mathbb{E}^{(\ell)}\{n_{k,r}\} & i=2 \end{cases}, \quad A(x) := I_1(x)/I_0(x)$$

***

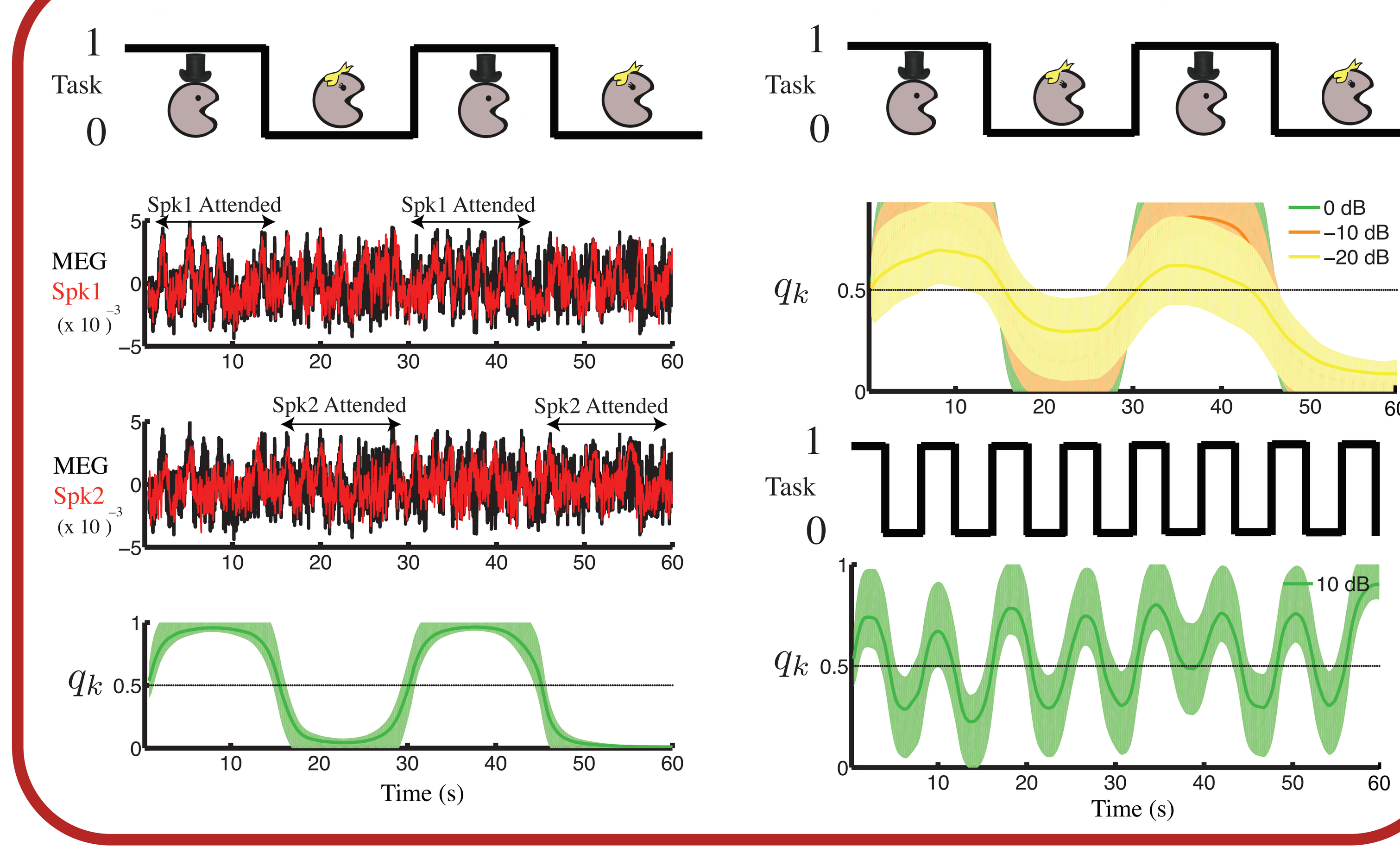
$$\begin{cases} \bar{z}_{k|k-1}^{(\ell+1,m)} = \bar{z}_{k-1|k-1}^{(\ell+1,m)} \\ \sigma_{k|k-1}^{(\ell+1,m)} = \sigma_{k-1|k-1}^{(\ell+1,m)} + \frac{\eta_k^{(\ell+1,m)}}{R} \\ \bar{z}_{k|k}^{(\ell+1,m)} = \bar{z}_{k|k-1}^{(\ell+1,m)} + \sigma_{k|k-1}^{(\ell+1,m)} \left[\sum_{r=1}^{R} \mathbb{E}^{(\ell)}\{n_{k,r}\} - R\frac{\exp\left(\bar{z}_{k|k}^{(\ell+1,m)}\right)}{1+\exp\left(\bar{z}_{k|k}^{(\ell+1,m)}\right)}\right] \\ \sigma_{k|k}^{(\ell+1,m)} = \left[\frac{1}{\sigma_{k|k-1}^{(\ell+1,m)}} + R\frac{\exp\left(\bar{z}_{k|k}^{(\ell+1,m)}\right)}{\left(1+\exp\left(\bar{z}_{k|k}^{(\ell+1,m)}\right)\right)^2}\right]^{-1} \end{cases}$$

$$\begin{cases} s_k^{(\ell+1,m)} = \sigma_{k|k}^{(\ell+1,m)} / \sigma_{k+1|k}^{(\ell+1,m)} \\ \bar{z}_{k|K}^{(\ell+1,m)} = \bar{z}_{k|k}^{(\ell+1,m)} + s_k^{(\ell+1,m)} \left(\bar{z}_{k+1|K}^{(\ell+1,m)} - \bar{z}_{k+1|k}^{(\ell+1,m)}\right) \\ \sigma_{k|K}^{(\ell+1,m)} = \sigma_{k|k}^{(\ell+1,m)} + s_k^{(\ell+1,m)} \left(\sigma_{k+1|K}^{(\ell+1,m)} - \sigma_{k+1|k}^{(\ell+1,m)}\right) s_k^{(\ell+1,m)} \end{cases}$$

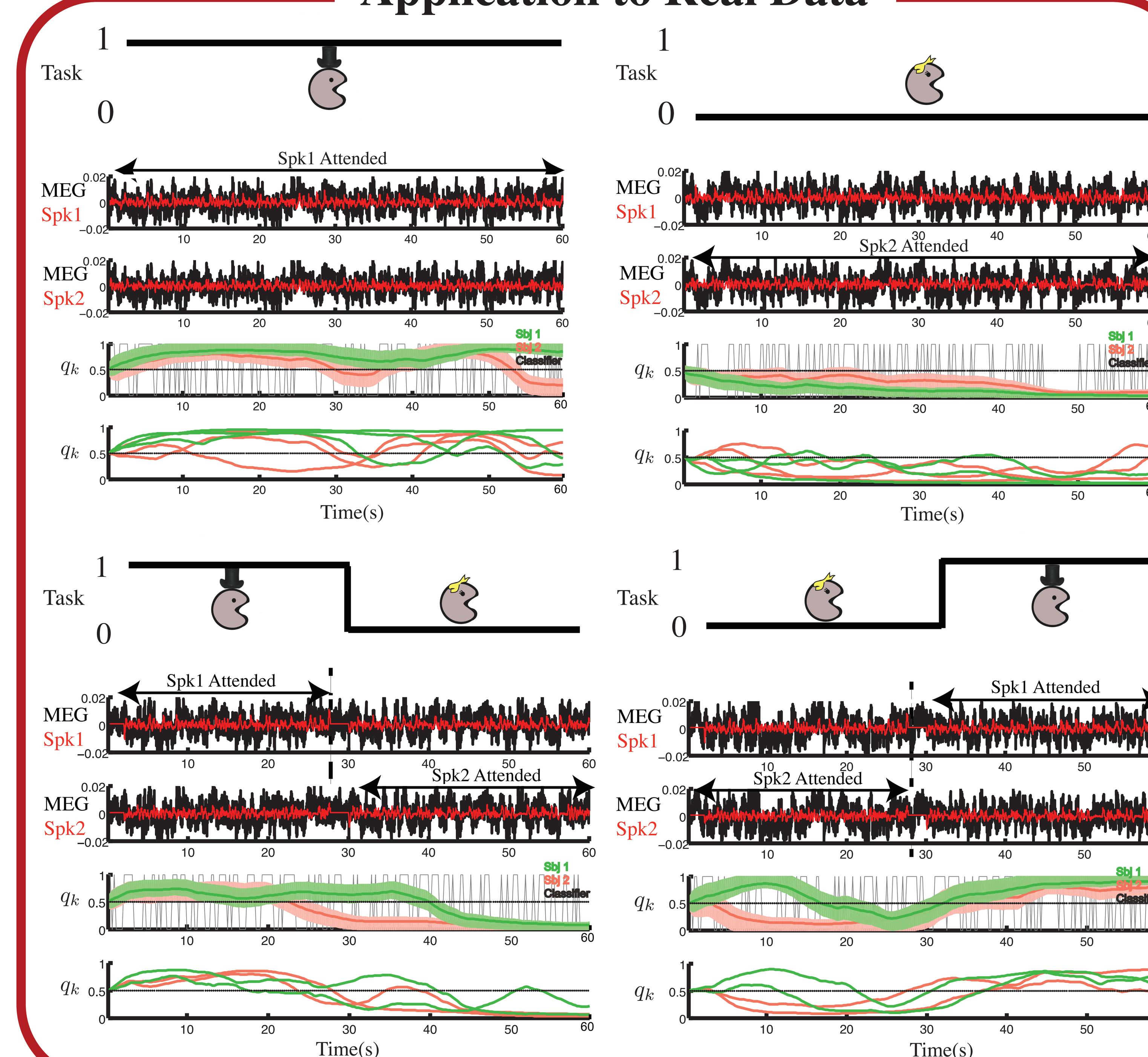
****

$$\eta_k^{(\ell+1,m+1)} = \frac{\left(\bar{z}_{k|K}^{(\ell+1,m)} - \bar{z}_{k-1|K}^{(\ell+1,m)}\right)^2 + \sigma_{k|K}^{(\ell+1,m)} + \sigma_{k-1|K}^{(\ell+1,m)} - 2\sigma_{k|K}^{(\ell+1,m)} s_{k-1}^{(\ell+1,m)} + 2\beta}{1 + 2(\alpha+1)}$$

Simulation Resaults



Application to Real Data



References

[1] Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. The Journal of the aoustical society of America, 25(5), 975-979.

[2] Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. Proceedings of the National Academy of Sciences, 109(29), 11854-11859.

[3] de Cheveigne, A., & Simon, J. Z. (2008). Denoising based on spatial filtering. Journal of neuroscience methods, 171(2), 331-339.

Acknowledgement

The authors thank Krishna Puvvada for interesting discussions. This project was supported by NIH grant 1R01AG036424.