Estimation of State Space Models with Gaussian Mixture Process Noise

Sina Miran[1,4], Jonathan Z. Simon[1,2,4], Michael C. Fu[3,4], Steven I. Marcus[1,4], Behtash Babadi[1,4]

Electrical & Computer Engineering[1], *Biology*[2], *Business School*[3], and *Institute for Systems Research*[4], University of Maryland College Park

I. Introduction and Motivation

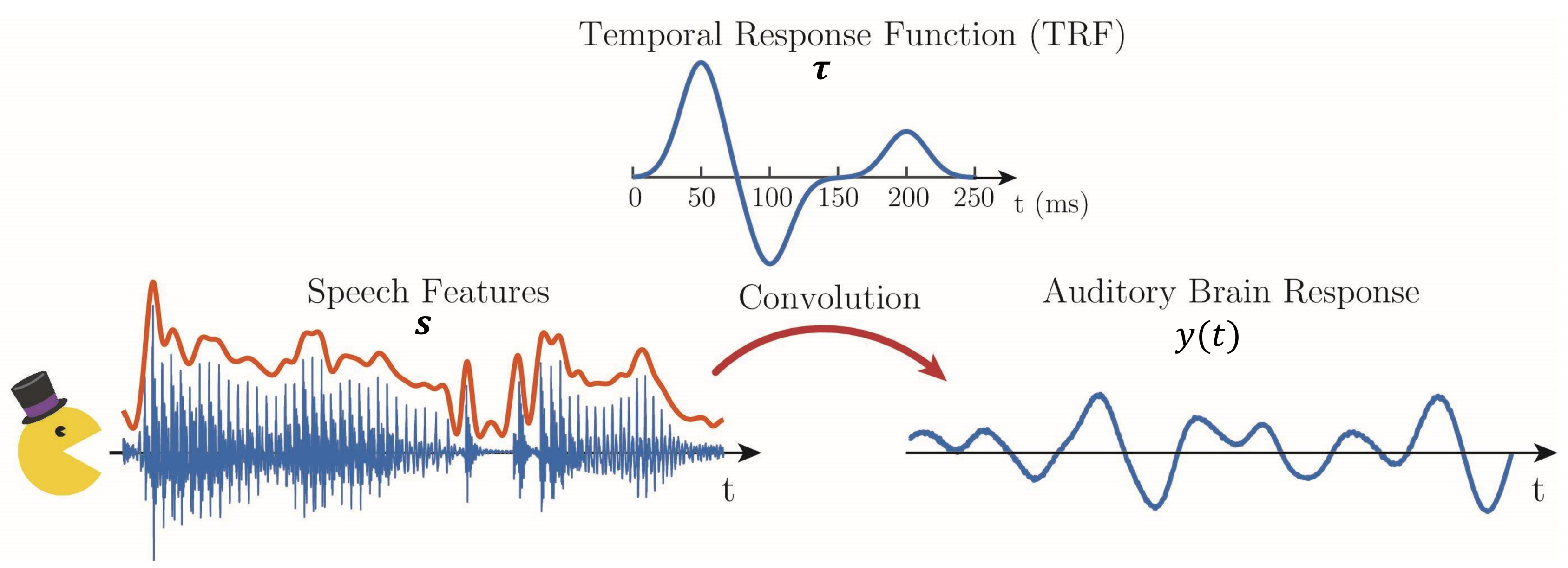
Consider the general state-space model with additive noise:

$$\begin{cases} \boldsymbol{x}_t = f(\boldsymbol{x}_t) + \boldsymbol{w}_t \,, \\ \boldsymbol{y}_t = g(\boldsymbol{x}_t) + \boldsymbol{v}_t \,, \end{cases} \qquad \boxed{\boldsymbol{w}_t \sim N(\boldsymbol{0}, \boldsymbol{Q}) \;\; \text{i. i. d}} \rightarrow \boxed{\text{common assumption}} \qquad \boldsymbol{v}_t \sim N(\boldsymbol{0}, \boldsymbol{R}) \;\; \text{i. i. d}$$

- Functional forms of $f(.)$ and $g(.)$ determined by expert domain knowledge
- A Gaussian assumption on observation noise $\boldsymbol{v}_t$ is mostly consistent with empirical histograms of measurement uncertainty.
- **However,** process noise $\boldsymbol{w}_t$ heavily depends on **how the latent process evolves over time** in the specific task/experiment.

↳ **i.i.d** and **Gaussian assumptions** for process noise $\boldsymbol{w}_t$ **are violated** in real world applications!

Application in Auditory Neuroscience:
estimating the **Temporal Response Function (TRF)** as a *time-varying filter*



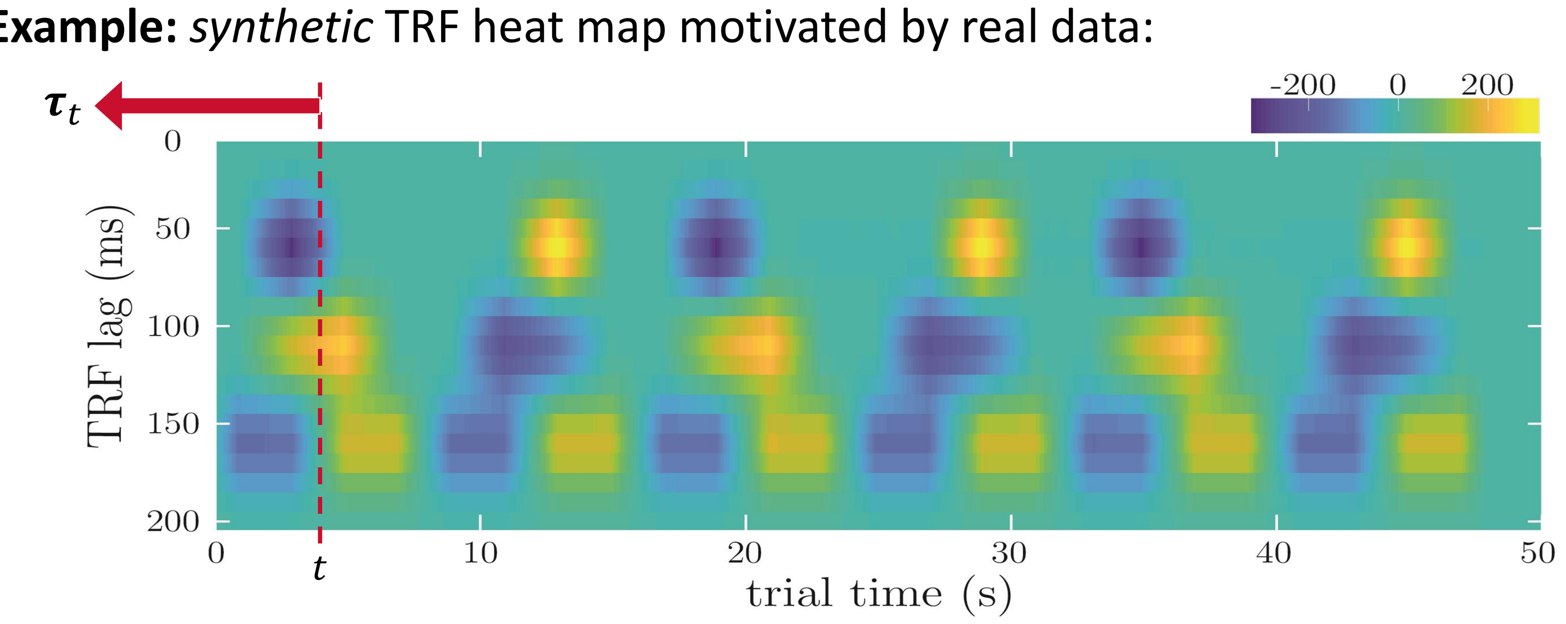
State-space model for estimating the TRF:

$$\begin{cases} \boldsymbol{x}_t = \alpha \boldsymbol{x}_{t-1} + \boldsymbol{w}_t \,, \\ \boldsymbol{\tau}_t = \boldsymbol{G}\boldsymbol{x}_t \,, \\ y_t = \boldsymbol{s}_t^{\mathrm{T}} \boldsymbol{\tau}_t + \boldsymbol{v}_t \,, \end{cases}$$

α : scalar close to unity	$\boldsymbol{G}$: a fixed dictionary matrix
$\boldsymbol{\tau}_t$: TRF at time t	$\boldsymbol{s}_t$: past envelope lags at time t
y_t: extracted auditory response (1-dim)	

Goal: estimate $\boldsymbol{x}_{1:T}$ and covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$

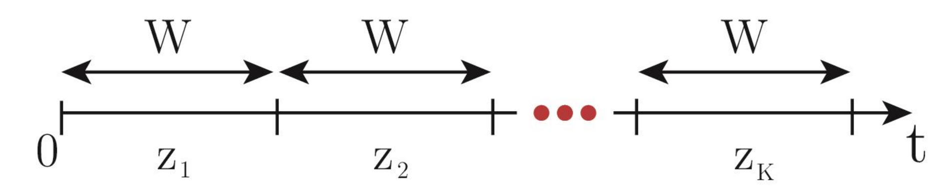
Example: *synthetic* TRF heat map motivated by real data:



- TRF components can be associated with attentive behavior and determine how speech features are processed in the brain.
- The components mostly exhibit a *repetitive behavior* including periods of *increasing*, remaining *constant*, and *decreasing*.
- This behavior can be best explained by a *multi-modal density* for the process noise $\boldsymbol{w}_t$ rather than a Gaussian density!

III. Proposed Model and Parameter Estimation

- To approximate an overall multi-modal density for $\boldsymbol{w}_t$, consider a Gaussian mixture with M components and parameter set $\Theta = \{p_{1:M}, \boldsymbol{\mu}_{1:M}, \boldsymbol{\Sigma}_{1:\mathrm{M}}\}$
- Consider K non-overlapping windows of length W samples ($T = KW$), and define a switching Gaussian process for $\boldsymbol{w}_t$:



- $z_k \in \{1, \dots, M\}$ determines the Gaussian component deriving the process noise in window k for $1 \le k \le K$

$$\begin{cases} \mathrm{P}(z_k = m) = p_m \,, & 1 \le k \le K, 1 \le m \le M \\ \boldsymbol{w}_t \sim N(\boldsymbol{\mu}_{z_k}, \boldsymbol{\Sigma}_{z_k}) \,, & (k-1)W + 1 \le t \le kW \end{cases}$$

- State estimation in state-space models under Gaussian mixture densities has been studied using *Gaussian sum filters/smoothers* and *particle methods*.
- **However**, estimation of mixture parameters has not been well-established!

Goal: estimate Θ from observations $\boldsymbol{y}_{1:T}$ ⟶ $\widehat{\Theta} = \underset{\Theta}{\operatorname{argmax}}\, \mathrm{P}(\boldsymbol{y}_{1:T} | \Theta)$

Estimation: Use **Expectation Maximization (EM)** with latent variables $\boldsymbol{x}_{1:T}, z_{1:K}$

$$Q(\Theta | \widehat{\Theta}^{(\ell)}) = \mathrm{E}_{x,z}\{\log \mathrm{P}(\boldsymbol{y}_{1:T}, \boldsymbol{x}_{1:T}, z_{1:K} | \Theta)\} = \sum_{k=1}^{K} \sum_{m=1}^{M} \mathrm{E}_x \left\{ \hat{\epsilon}_{k,m}^{(\ell)} (\log p_m + \log \pi_{k,m}) \right\} + c_0$$

$$\pi_{k,m} = \mathrm{P}(\boldsymbol{x}_{(k-1)W+1:kW} \,|\, \boldsymbol{x}_{(k-1)W}, z_k = m, \Theta), \qquad \hat{\epsilon}_{k,m}^{(\ell)} = \frac{\hat{p}_m^{(\ell)} \hat{\pi}_{k,m}^{(\ell)}}{\sum_{m'=1}^{M} \hat{p}_{m'}^{(\ell)} \hat{\pi}_{k,m'}^{(\ell)}}$$

- $\hat{\epsilon}_{k,m}^{(\ell)}$ is constant w.r.t. Θ and $\hat{\pi}_{k,m}^{(\ell)}$ defined Similar to $\pi_{k,m}$ but for $\widehat{\Theta}^{(\ell)}$
- The expectation w.r.t. $\boldsymbol{x}$ does not have a closed form!

E Step: perform *particle smoothing* to obtain N_s sample paths for the densities $\boldsymbol{x}_{(k-1)W:kW} | \boldsymbol{y}_{1:T}, \widehat{\Theta}^{(\ell)}$ each with probability $\lambda_k^{(n)}$

M Step: use the particle representations to approximate the expectations and update the parameter estimates as below

$$\hat{p}_m^{(\ell+1)} = \frac{1}{K} \sum_k \sum_n \lambda_k^{(n)} \hat{\epsilon}_{k,m}^{(\ell,n)} \,,$$

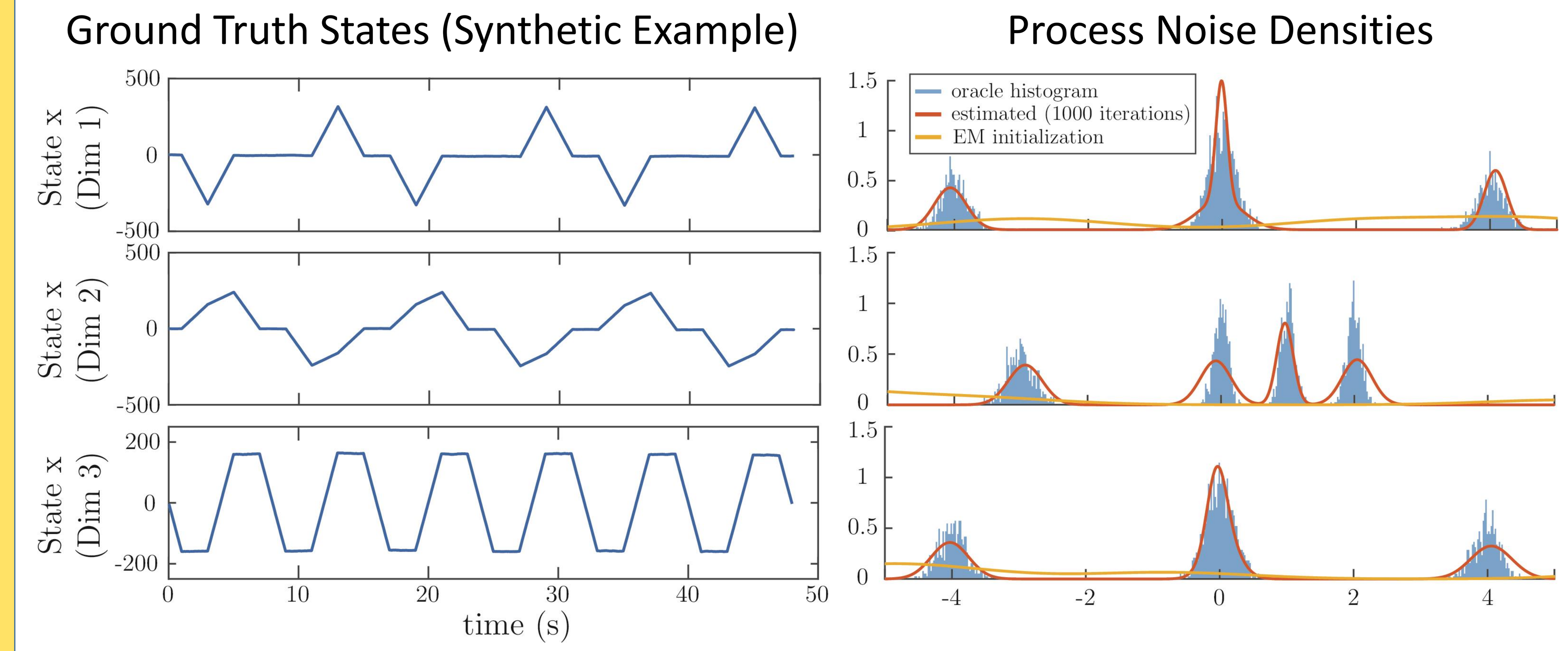
$$\hat{\mu}_m^{(\ell+1)} = \frac{1}{T \hat{p}_m^{(\ell+1)}} \sum_k \sum_n \lambda_k^{(n)} \hat{\epsilon}_{k,m}^{(\ell,n)} \sum_j \boldsymbol{u}_{k,j}^{(n)}$$

$$\hat{\Sigma}_m^{(\ell+1)} = \frac{1}{T \hat{p}_m^{(\ell+1)}} \sum_k \sum_n \lambda_k^{(n)} \hat{\epsilon}_{k,m}^{(\ell,n)} \sum_j \left(\boldsymbol{u}_{k,j}^{(n)} - \hat{\mu}_m^{(\ell+1)} \right) \left(\boldsymbol{u}_{k,j}^{(n)} - \hat{\mu}_m^{(\ell+1)} \right)^{\mathrm{T}}$$

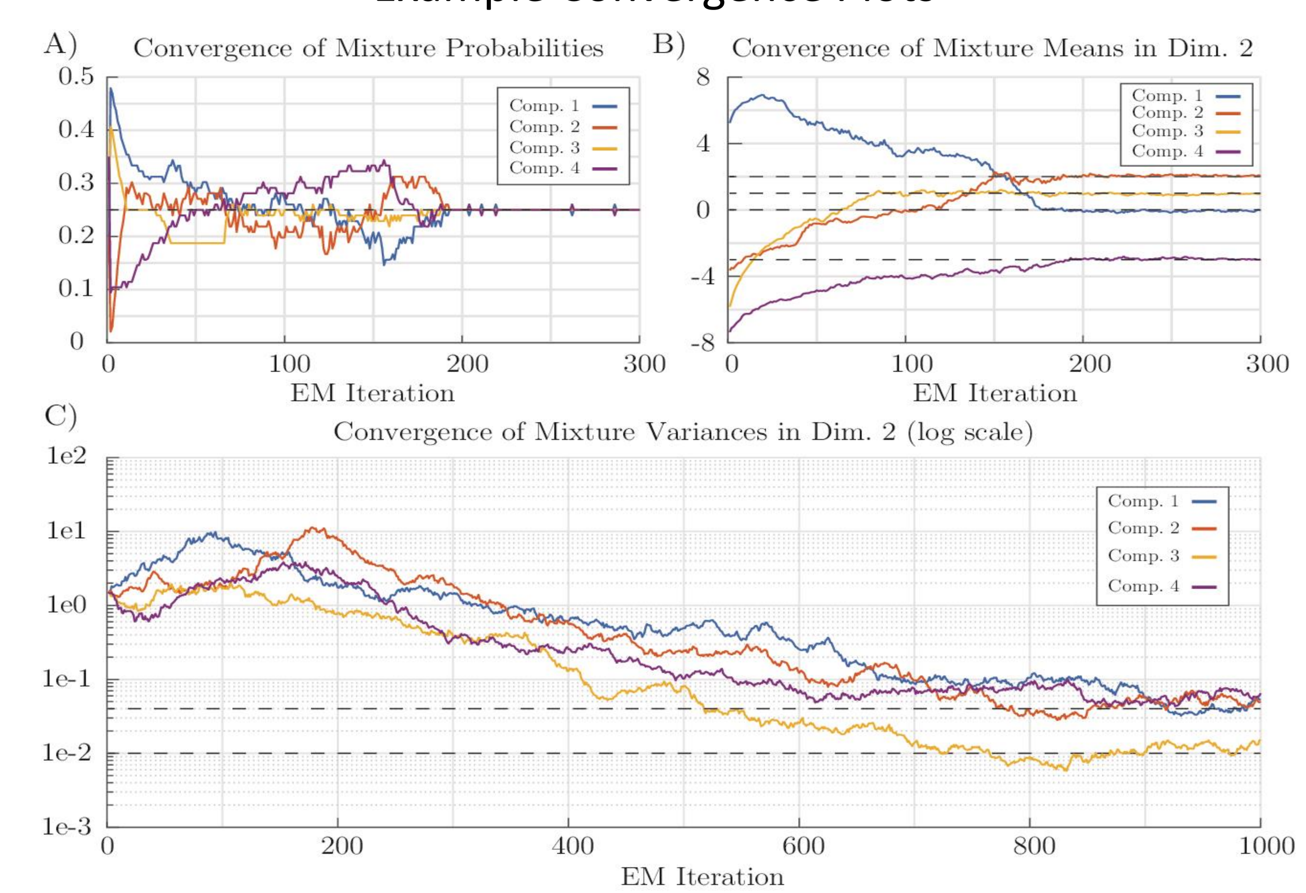
where $\boldsymbol{u}_{k,j}^{(n)} = \boldsymbol{x}_{(k-1)W+j}^{(n)} - f\left(\boldsymbol{x}_{(k-1)W+j-1}^{(n)}\right)$

The steps are repeated until convergence.

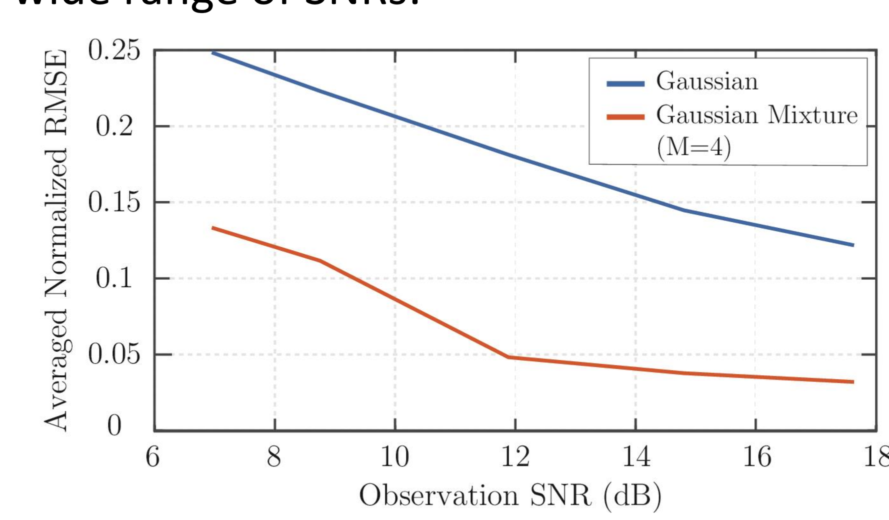
IV. Simulation Results



Example Convergence Plots



State RMSE improvement over Gaussian model due to a richer representation for process noise for a wide range of SNRs:



V. Summary and Future Work

Provided a framework for estimating a multi-modal process noise density in state-space model, which results in a more robust state inference. Future work includes:

- In large state dimensions, particle smoothing is computationally intensive
  ↳ need scalable alternatives
- Application to recorded MEG data in auditory experiments

Research was supported by Defense Advanced Research Projects Agency (DARPA) under cooperative agreement #N660011824024

Poster PDF available at: http://ter.ps/simonpubs, Email: smiran@umd.edu