Evolutionary State-Space Model and Its Application to Time-Frequency Analysis of Local Field Potentials

Xu Gao[1], Weining Shen[1], Babak Shahbaba[1], Norbert J. Fortin[2], Hernando Ombao[1,3,4]

[1] *Department of Statistics, University of California, Irvine, California, U.S.A.*

[2] *Department of Neurobiology and Behavior, University of California Irvine, Irvine, California, U.S.A.*

[3] *Department of Cognitive Sciences, University of California, Irvine, California, U.S.A.*

[4] *Statistics Program, King Abdullah University of Science and Technology, Saudi Arabia.*

Supplementary Material

In this Supplementary File, we provide the proof of theorems in the main paper, additional simulation results and LFP data analysis.

S1. Proof of Theorem 1

Our proof is based on the MLE consistency results for the general hidden Markov models in Douc et al. (2011), hereafter D11. More specifically, Theorem 1 and Section 3.1 in D11 are directly applicable to our model. Therefore it is good enough to verify Conditions (L1)–(L4) on Page 12 in D11. Based on D11's notation, their model is formulated as

$$X_t = A_\theta X_{t-1} + R_\theta U_{t-1},$$

$$Y_t = B_\theta X_t + S_\theta V_t,$$

with U_t and V_t being i.i.d. Gaussian vectors with zero means and identity covariance matrices. Under our framework,

$$A_\theta = \begin{pmatrix} \Phi_1^{(1)} & \Phi_2^{(1)} & & & \\ I_q & \mathbf{0} & & & \\ & & \ddots & & \\ & & & \Phi_1^{(R)} & \Phi_2^{(R)} \\ & & & I_q & \mathbf{0} \end{pmatrix}, \quad B_\theta = \begin{pmatrix} M & \mathbf{0} \end{pmatrix}, \quad R_\theta = I, \quad S_\theta = I.$$

Therefore the conditions (L3) and (L4) on Page 12 in D11 trivially hold. Moreover, since both $\Phi_1^{(k)}$ and $\Phi_2^{(k)}$ are diagonal matrices with different AR(2) coefficients and $\boldsymbol{M}$ is assumed to be of full column rank, Condition (L1) in D11 is also satisfied. To verify Condition (L2), we need to show that the eigenvalues of A_θ all lie in the open unit circle. This can done by considering the non-zero eigenvalue λ and eigenvector γ of A_θ such that $A_\theta\gamma = \lambda\gamma$. Without loss of generality, we can let $R = 1$ since A_θ is block-diagonal, and partition $\gamma^T = (\gamma_1^T, \gamma_2^T)$, we have $\gamma_1 = \lambda\gamma_2$, which leads to a quadratic equation $\lambda^2 = \lambda\Phi_1^{(1)} + \Phi_2^{(2)}$. It can be seen that the solution of λ to equation $\lambda^2 - \lambda\Phi_1^{(1)} - \Phi_2^{(1)} = 0$ corresponds to the reciprocal of the roots for the AR(2) process defined by the coefficients of $\Phi_1^{(1)}$ and $\Phi_2^{(2)}$. By the assumption that these AR(2) processes are causal, their roots lie outside the unit circle, which implies that the solutions of λ are inside the unit circle. Therefore the conditions on Page 12 in D11 are all satisfied, which concludes the proof.

S2. Proof of Theorem 2

We first present a lemma that gives an explicit form of the autocovariance function of an AR(2) process. Such results will be helpful for proving the main theorem.

Lemma 1. *Given a (weakly) stationary zero mean AR(2) process S_t, the autocovariance function $\gamma_S(h)$ takes the form*

$$\gamma_S(h) = A_1(\rho e^{\psi i})^{-h} + A_2(\rho e^{-\psi i})^{-h}, \tag{S2.1}$$

where A_1, A_2 can be determined by solving the linear equation $A_1+A_2 = \frac{(1-\phi_2)\sigma_w^2}{(1+\phi_2)(1-\phi_1-\phi_2)(1+\phi_1-\phi_2)}$ and $A_1(\rho e^{\psi i})^{-1} + A_2(\rho e^{-\psi i})^{-1} = \frac{\phi_1\sigma_w^2}{(1+\phi_2)(1-\phi_1-\phi_2)(1+\phi_1-\phi_2)}$.

The proof is due to the fact that $\gamma_S(h) = \phi_1\gamma_S(h-1) + \phi_2\gamma_S(h-2)$.

To prove Theorem 1, we first show that for any fixed J, $\{f_{S^{(j)}}(\omega)\}_{j=1}^J$ are linearly independent. In fact, suppose there exists some constants $b_1, \cdots, b_J$ such that $\sum_{j=1}^J b_j f_{S^{(j)}}(\omega) = 0$, then we must have $\sum_{j=1}^J b_j \sum_{h=-\infty}^{\infty} \gamma_{S^{(j)}}(h)e^{2\pi i\omega h} = \sum_{h=-\infty}^{\infty}\sum_{j=1}^J b_j\gamma_{S^{(j)}}(h)e^{2\pi i\omega h} = 0$. As a direct result from Fourier theorem, we have $\sum_{j=1}^J b_j\gamma_{S^{(j)}}(h) = 0$ for any h. Thus for any positive integer H, $b_1, \cdots, b_J$ are solutions of the linear equation

$$\Gamma\mathbf{b} = 0, \tag{S2.2}$$

where $\Gamma = \begin{bmatrix} \gamma_{S^{(1)}}(0) & \gamma_{S^{(2)}}(0) & \dots & \gamma_{S^{(J)}}(0) \\ \gamma_{S^{(1)}}(1) & \gamma_{S^{(2)}}(1) & \dots & \gamma_{S^{(J)}}(1) \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{S^{(1)}}(H) & \gamma_{S^{(2)}}(H) & \dots & \gamma_{S^{(J)}}(H) \end{bmatrix}_{(H+1)*J}$ and $\mathbf{b} = (b_1, \cdots, b_J)'$. From Lemma 1, it is easy to show that $\gamma_{S^{(j)}}(h) = (\rho^{(j)})^{-h}(A_1^{(j)} + A_2^{(j)})\cos(h\psi^{(j)})$. Note that due to the condition that $\max\{|\omega_1 - \omega_0|, \cdots, |\omega_J - \omega_{J-1}|\} \to 0$ and $A_1^{(j)}, A_2^{(j)}$ are nonlinear functions of j, we have $rank(\Gamma) = min\{H+1, J\}$. It implies $\mathbf{b} = \mathbf{0}$ and $\{f_{S^{(j)}}(\omega)\}_{j=1}^M$ are linearly independent. Then we can implement the Gram-Schmidt process on the family of functions $\{f_{S^{(j)}}(\omega)\}_{j=1}^{\infty}$ to obtain a family of orthonormal functions $\{\tilde{f}_{S^{(j)}}(\omega)\}_{j=1}^{\infty}$ in $L^2(0, \frac{1}{2})$. It follows that for any nonnegative coefficients $a_1, \cdots, a_J$, there exist $\tilde{a}_1, \cdots, \tilde{a}_J$ such that $||f_Y(\omega) - \sum_{j=1}^J a_j^2 f_{S^{(j)}}(\omega)||_2 =$

$||f_Y(\omega) - \sum_{i=1}^{J} \tilde{a}_j^2 \tilde{f}_{S^{(j)}}(\omega)||_2$. If we can show $\{\tilde{f}_{S^{(j)}}(\omega)\}_{j=1}^{\infty}$ is also complete in $L^2(0, \frac{1}{2})$, by Parseval equality, we can obtain that $||f_Y(\omega) - \sum_{j=1}^{J} \tilde{a}_j^2 \tilde{f}_{S^{(j)}}(\omega)||_2 \to 0$ as $J \to \infty$ and equivalently, $||f_Y(\omega) - f_{\hat{Q}_{t,J}}(\omega)||_2 \to 0$ as $J \to \infty$.

To show that $\{\tilde{f}_{S^{(j)}}(\omega)\}_{j=1}^{\infty}$ is complete in $L^2(0, \frac{1}{2})$, it suffices to show $\{f_{S^{(j)}}(\omega)\}_{j=1}^{\infty}$ is complete. Let us define $\mathbb{B} = \{f_{S^{(j)}}(\omega)\}_{j=1}^{\infty}$. For any function $g(\omega)$ in $L^2(0, \frac{1}{2})$, if $g(\omega) \perp \mathbb{B}$, then we have $\int_0^{\frac{1}{2}} g(\omega) f_{S^{(j)}}(\omega) d\omega = 0$ for any j. It is equivalent to $\sum_{h=\infty}^{\infty} \int_0^{\frac{1}{2}} g(\omega) \gamma_{S^{(j)}}(h) e^{2\pi i \omega h} d\omega = \sum_{h=-\infty}^{\infty} \gamma_g(h) \gamma_{S^{(j)}}(h) = 0$ for any j. It boils down to the problem of solving for the linear equation $\Gamma' \gamma = 0$, where Γ is defined in Equation (S2.2) and $\gamma = (\gamma_g(0), \cdots, \gamma_g(H))$ for any J and H. We have proved that Γ is of full row rank and thus $\gamma_g(h) = 0$ for any h. Thus $\{f_{S^{(j)}}(\omega)\}_{j=1}^{\infty}$ is complete in $L^2(0, \frac{1}{2})$.

For the proof of the second part of the theorem, note that the basis of $\mathcal{F}_J$ is equivalent with the Fouriers series basis $\{e^{2\pi i \omega_k h}\}_{k=1}^{J}$. Therefore by the Jackson inequality, the result follows.

S3. Computation algorithms: E-SSM estimation for a single epoch

Here we present the algorithms for E-SSM model parameter estimation when there is only a single epoch, as described in Section 3.1 of the main paper. We propose an iterative algorithm that comprises of Kalman filter and least squares for parameter estimation purpose. We start with initial values $\widetilde{M} = \widetilde{M}_0$, $\boldsymbol{X}_0^0$ and P_0^0. The estimation procedure takes iterations between Algorithms 1 and 2 (shown below) until convergence. In this study, since we are interested in the power of particular frequency bands, we will introduce box constraints to the modulus $\rho_1, \cdots, \rho_q$ to control the spread of the spectra curves. Hence in A.2 of Algorithm 1, we implement an optimization approach with box constraints on modulus $\rho_1, \cdots, \rho_q$ and no constraints on σ^2, τ^2.

When choosing the initial values for $\widetilde{M}_0$, $\boldsymbol{X_0^0}$, P_0^0 in Algorithm 1, we follow the same

Algorithm 1 Kalman Filter and Maximum Likelihood

procedure GIVEN $\widetilde{M}, \boldsymbol{X}_0^0, P_0^0$, ESTIMATE $\boldsymbol{\rho}, \sigma^2, \tau^2$ BY KALMAN FILTER AND MAXIMUM LIKELIHOOD OF INNOVATIONS $\boldsymbol{\epsilon_t}$

A.1 Kalman filter and Kalman gain step

$\Phi_1 \leftarrow diag(2\rho_1^{-1}cos(\psi_1), \cdots, 2\rho_q^{-1}cos(\psi_q))$

$\Phi_2 \leftarrow diag(-\rho_1^{-2}, \cdots, -\rho_q^{-2})$

$\widetilde{\Phi} \leftarrow \begin{bmatrix} \Phi_1 & \Phi_2 \\ \boldsymbol{I_q} & \boldsymbol{0} \end{bmatrix}$

for $t = 0, \ldots, T$ **do**

$\boldsymbol{X}_t^{t-1} \leftarrow \widetilde{\Phi}\boldsymbol{X}_{t-1}^{t-1}$

$P_t^{t-1} \leftarrow \widetilde{\Phi}P_{t-1}^{t-1}\widetilde{\Phi}' + \sigma^2 \begin{bmatrix} \boldsymbol{I_q} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$

$K_t \leftarrow P_t^{t-1}\widetilde{M}'[\widetilde{M}P_t^{t-1}\widetilde{M}' + \tau^2\boldsymbol{I_p}]^{-1}$ ▷ The Kalman gain

$\boldsymbol{X}_t^t \leftarrow \boldsymbol{X}_t^{t-1} + K_t(\boldsymbol{Y}_t - \widetilde{M}\boldsymbol{X}_t^{t-1})$

$P_t^t \leftarrow (\boldsymbol{I_{2q}} - K_t\widetilde{M})P_t^{t-1}$

A.2 Maximum likelihood estimation

for $t = 0, \ldots, T$ **do**

$\boldsymbol{\epsilon_t} \leftarrow \boldsymbol{Y}_t - \widetilde{M}\boldsymbol{X}_t^{t-1}$

$\Sigma_t \leftarrow \widetilde{M}P_t^{t-1}\widetilde{M}' + \tau^2\boldsymbol{I_p}$

$l_Y(\boldsymbol{\rho}, \sigma^2, \tau^2) \leftarrow \frac{1}{2}\sum_{t=1}^{T}\log|\Sigma_t| + \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{\epsilon_t}'\Sigma_t^{-1}\boldsymbol{\epsilon_t}$ ▷ The negative loglikelihood

$(\hat{\boldsymbol{\rho}}, \hat{\sigma}^2, \hat{\tau}^2) \leftarrow \underset{(\boldsymbol{\rho}, \sigma^2, \tau^2)}{\operatorname{argmin}} \; l_Y(\boldsymbol{\rho}, \sigma^2, \tau^2)$ ▷ Maximizing the likelihood of innovations

return $\hat{\boldsymbol{\rho}}, \hat{\sigma}^2, \hat{\tau}^2$

strategy in Harvey (1990). We let $\boldsymbol{X_0^0} = \mathbf{0}$ and $P_0^0 = \kappa \boldsymbol{I}$, where $\boldsymbol{I}$ is the identify matrix and κ is large such that $(\kappa \boldsymbol{I})^{-1} \approx \mathbf{0}$. We also allow $\widetilde{M_0} = \mathbf{0}$ or take values from a uniform distribution $U(0, 1)$.

Algorithm 2 Kalman Filter and Least Squares Estimation

procedure GIVEN THE CURRENT ESTIMATES OF $\boldsymbol{\rho}, \sigma^2, \tau^2$, WE CAN OBTAIN THE ESTIMATES OF $\widetilde{M}$ BY KALMAN FILTER AND LEAST SQUARES ESTIMATION.

B.1 Kalman filter and Kalman gain step

$\Phi_1 \leftarrow diag(2\rho_1^{-1} cos(\psi_1), \cdots, 2\rho_q^{-1} cos(\psi_q))$

$\Phi_2 \leftarrow diag(-\rho_1^{-2}, \cdots, -\rho_q^{-2})$

$\widetilde{\Phi} \leftarrow \begin{bmatrix} \Phi_1 & \Phi_2 \\ \boldsymbol{I_q} & \mathbf{0} \end{bmatrix}$

for $t = 0, \ldots, T$ **do**

$\boldsymbol{X}_t^{t-1} \leftarrow \widetilde{\Phi} \boldsymbol{X}_{t-1}^{t-1}$

$P_t^{t-1} \leftarrow \widetilde{\Phi} P_{t-1}^{t-1} \widetilde{\Phi}' + \sigma^2 \begin{bmatrix} \boldsymbol{I_q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$

$K_t \leftarrow P_t^{t-1} \widetilde{M}' [\widetilde{M} P_t^{t-1} \widetilde{M}' + \tau^2 \boldsymbol{I_p}]^{-1}$ ▷ The Kalman gain

$\boldsymbol{X}_t^t \leftarrow \boldsymbol{X}_t^{t-1} + K_t(\boldsymbol{Y}_t - \widetilde{M} \boldsymbol{X}_t^{t-1})$

$\boldsymbol{X}_t^t \leftarrow \boldsymbol{X}_t^t / \text{sd}(\boldsymbol{X}_t^t)$ ▷ $\text{sd}(\boldsymbol{X}_t^t)$ denotes the standard deviation of $\boldsymbol{X}_t^t$

//Remark: We scale $\boldsymbol{X}_t^t$ to unit variance for identifiability issues discussed before.

$P_t^t \leftarrow (\boldsymbol{I_{2q}} - K_t \widetilde{M}) P_t^{t-1}$

B.2 Least square estimation from Equation (**??**)

$\boldsymbol{Y} \leftarrow (\boldsymbol{Y_1}, \cdots, \boldsymbol{Y_T})$ ▷ $\boldsymbol{Y} \in \mathbb{R}^{p \times T}$

$\boldsymbol{X} \leftarrow (\boldsymbol{X_1^1}, \cdots, \boldsymbol{X_T^T})$ ▷ $\boldsymbol{X} \in \mathbb{R}^{q \times T}$

for $w = 1, \ldots, p$ **do**

$\widetilde{M_w} \leftarrow (\boldsymbol{X} * \boldsymbol{X}')^{-1} * \boldsymbol{X} * \boldsymbol{Y}'_{(w)}$ ▷ $\boldsymbol{Y}_{(w)}$ denotes the wth row of $\boldsymbol{Y}$

$\widetilde{M} \leftarrow (\widetilde{M_1}, \cdots, \widetilde{M_w})'$

return $\widetilde{M}$

S4. Sensitivity analysis

We also conduct sensitivity analysis for the proposed E-SSM via simulation studies under multiple epochs setting.

In Scenario 1, we assumed that the underlying latent sources followed autoregressive moving average (ARMA) processes with the same AR coefficients in Section 6.2 of the main paper. In terms of the MA parameters, we allowed the order (mv_{order}) and coefficients (mv_{coef}) vary to some extent. We then applied the proposed E-SSM method and calculated the sum of square errors (SSE) of the "mixing" matrix. As a benchmark, we also applied the proposed E-SSM to the "true" AR(2) generated sources and obtained an SSE of 0.068. From the results shown in Table 1, the SSE increases as we increase the order of mv and the value of mv coefficients. However, we can still observe reasonable SSE when the mv order and coefficients go below 5.

In Scenario 2, we added random noise into the underlying AR(2) processes. Similar to Scenario 1, we fixed the AR coefficients at the same values with those in Section 6.2 of the main paper. The noise was generated from uniform and Gaussian distribution separately. The uniform distribution was supported from $-u$ to u, where u took values in $\{.01, .1, 1, 5, 10\}$. For the Gaussian distribution $N(0, \sigma_n^2)$, we let σ_n take values in $\{.01, .1, 1, 5, 10\}$. We summarized the SSE in Table 2.

In Scenario 3, we increased the order of underlying AR processes. To be specific, we generated AR(10) processes with FIVE peaks at 2, 12, 22, 32, 42 Hertz. The modulus were similar to those in Section 6.2 of the main paper. Web Figure 18 shows the true and the estimated "mixing" matrices. It can be easily found that the proposed E-SSM captures the structure directly even though the sources deviate from AR(2) processes. To further investigate the performance of E-SSM, we "reconstructed" the observed signals using the estimated "mixing" matrix and "uncovered" latent sources. Web Figure 19 presents the periodograms of true and estimated

signals from channel 16. It shows that the estimated signals mainly recover the true processes with only some slight discrepancies. The dynamics across epochs are Allen et al. (2016) perfectly captured.

In Scenario 4, we generated 5 latent independent AR(2) signals corresponding to delta (δ: 0 - 4 Hertz), theta (θ: 4 - 8 Hertz), alpha (α: 8 - 12 Hertz), lower beta (β: 12 - 18 Hertz) and gamma (γ: $>$ 32 Hertz). To generate the observed signals, we *only* choose 3 latent independent AR(2) processes (delta, theta and lower beta bands) and 20 electrodes. SWe allow the modulus $(\rho_1^{(r)}, \rho_2^{(r)}, \rho_3^{(r)})$ to increase from $(1.001, 1.001, 1.001)$ with an increment of 0.00005 as the epoch r propagates. All the remaining parameters are the same as in Section 6.1 of the main paper. To evaluate the robustness of the proposed method, we also fit *FIVE* frequency bands into the observed signals. Web Figure 4 shows the periodogram of the generated signals from electrode 1. We fit the proposed model with *FIVE* frequency bands. Web Figure 5 shows the true mixing matrix (left) and its estimation (right). From the true matrix, we can observe zero columns corresponding to "alpha" and "gamma" bands that indicate the observed signals are generated only by the three remaining bands (delta, theta and lower beta bands). From the estimation result, it is clear that the columns of "alpha" and "gamma" bands are roughly zero, which shows the proposed E-SSM successfully capture the three latent sources (delta, theta and lower beta bands) while neglecting the impacts from alpha and gamma bands. Web Figure 6 shows the periodograms of the true and estimated signals from the three underlying AR(2) processes. Similar to the results in Section 6.2 of the main paper, we can see the pattern of the periodograms from the reconstructed AR(2) process is consistent with that of the true AR(2) process.

S5. Figures

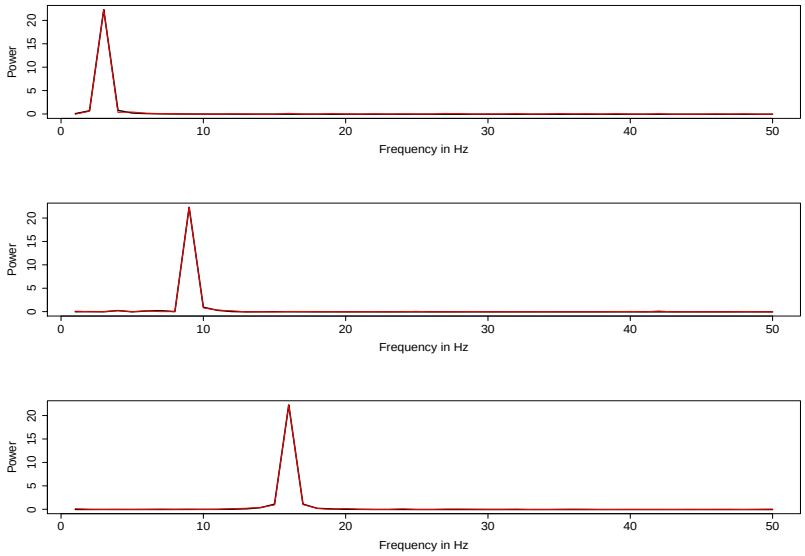

Figure 1: The periodograms of the true (black) and estimated (red) latent processes.

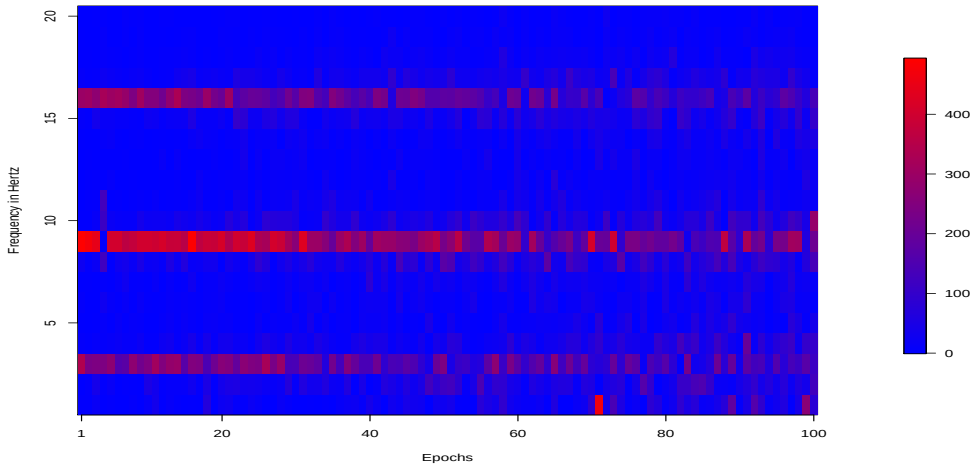

Figure 2: The periodogram of generated signals from electrode 1 computed over all 100 epochs. From the heat map, we are observing the powers are evolving across epochs. At early stage, three dominating frequency bands can be identified clearly. As epoch evolves, such pattern is getting less clear.

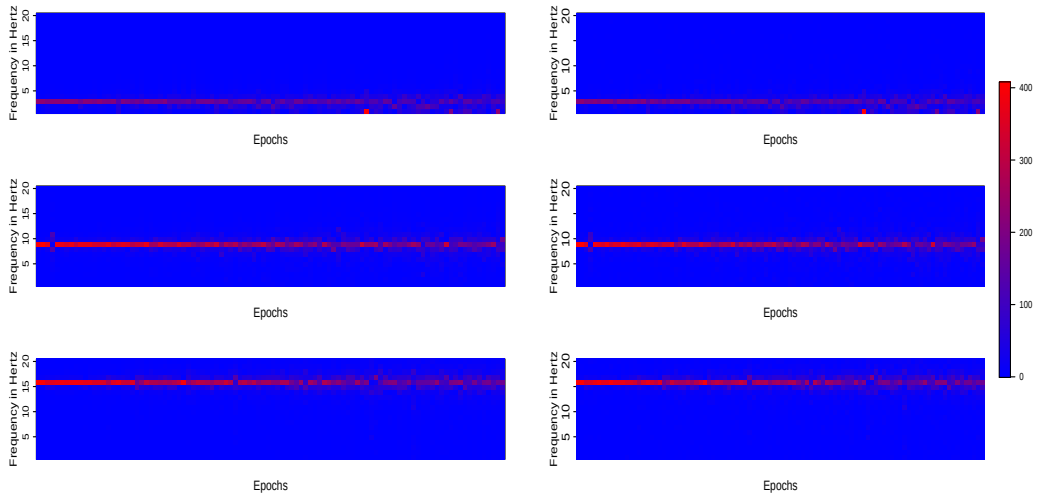

Figure 3: The periodograms of the true (left) and estimated (right) latent AR(2) processes corresponding to delta (top), alpha (middle) and beta (bottom) frequency band.

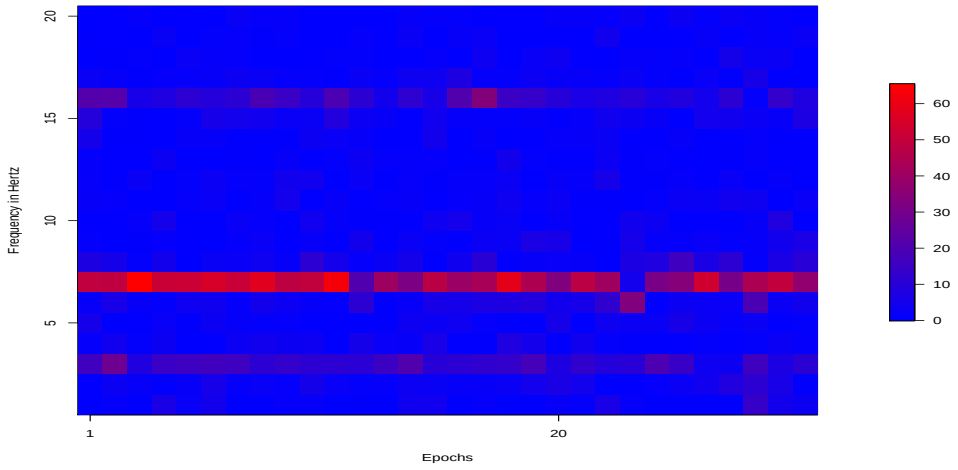

Figure 4: The periodogram of generated signals from electrode 1 computed over all 30 epochs.

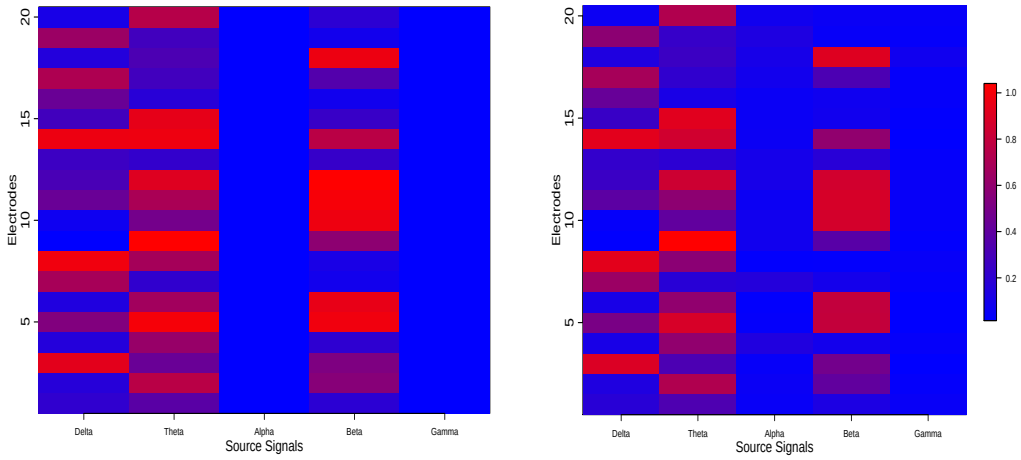

Figure 5: The true mixing matrix (left) and estimated mixing matrix (right). Darker color indicates heavier weight given by the corresponding latent processes. Columns corresponding to "alpha" and "gamma" bands are zero in the true mixing matrix (left). In the estimated mixing matrix (right), those two columns are also close to zero.

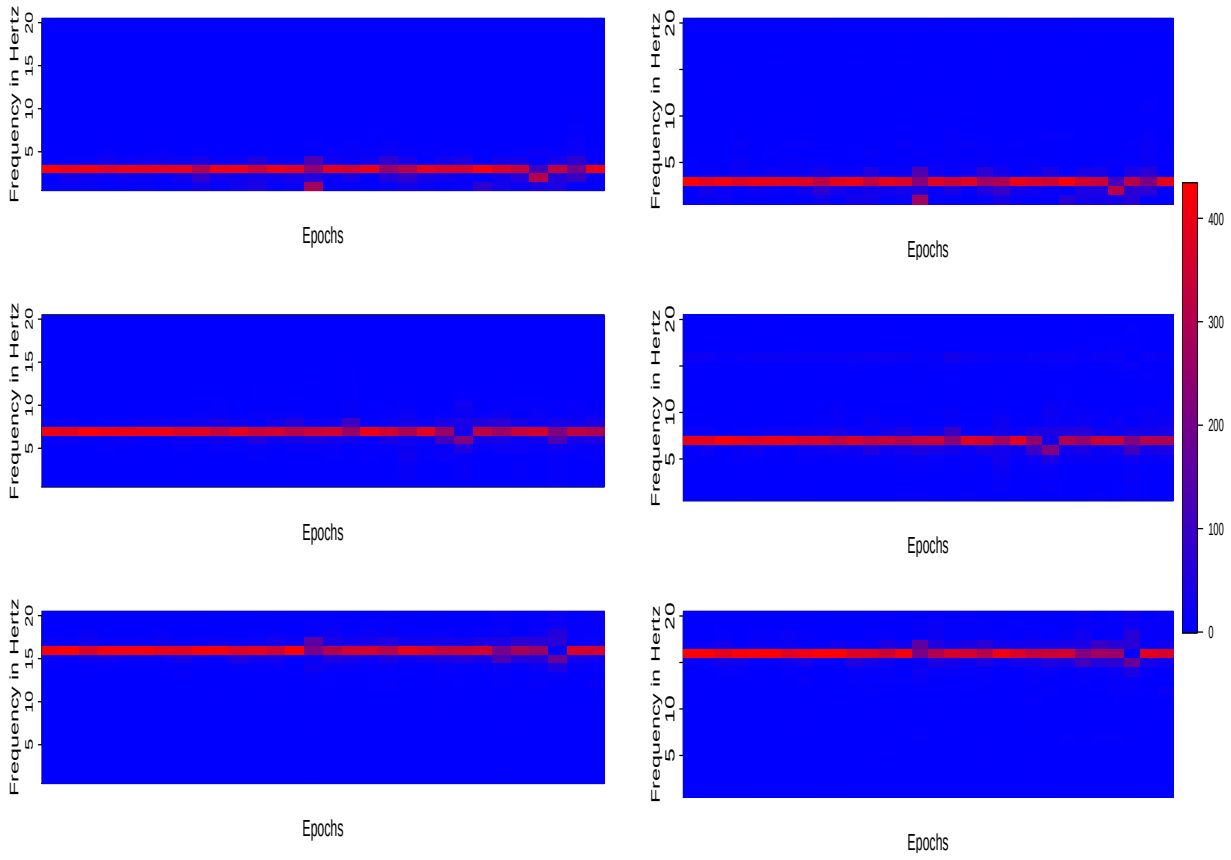

Figure 6: The periodograms of the true (left) and estimated (right) latent AR(2) processes corresponding to delta (top), theta (middle) and beta (bottom) frequency band.

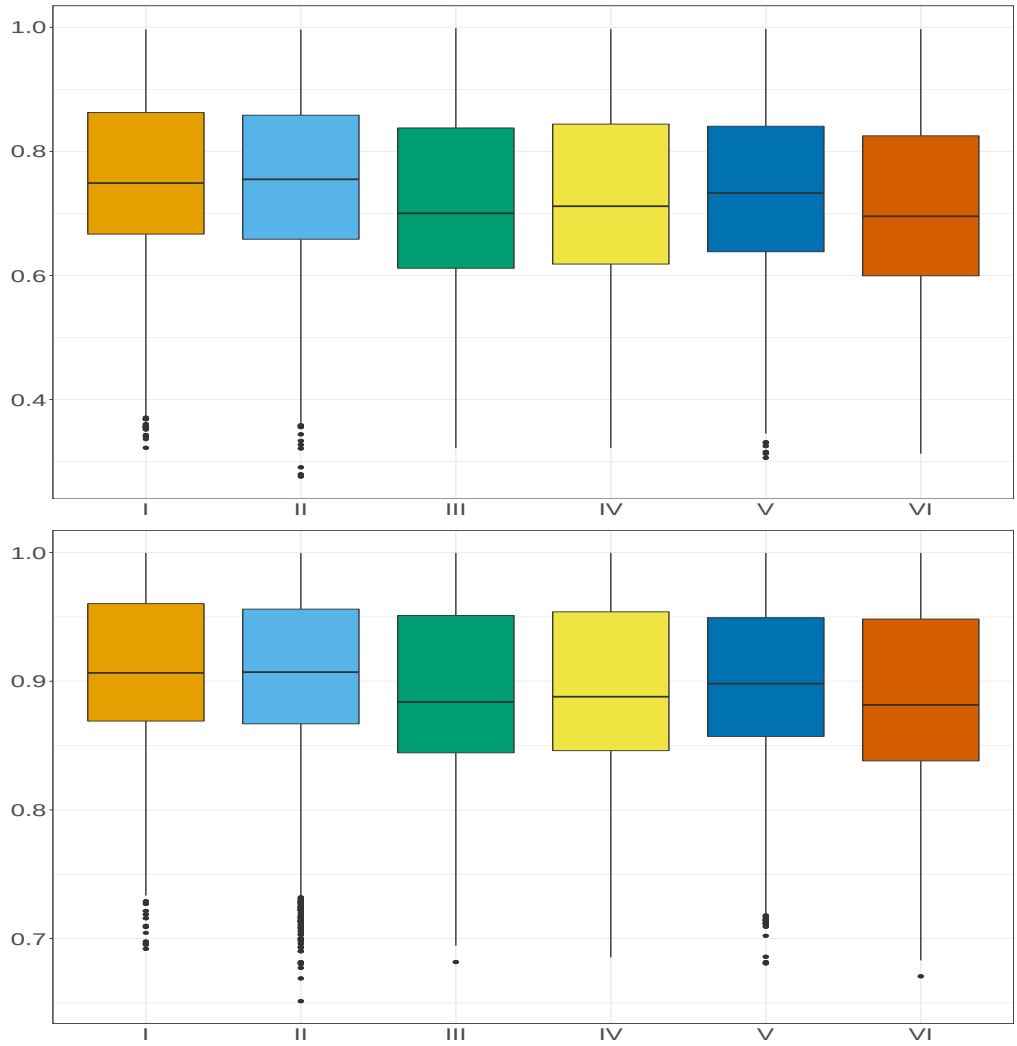

Figure 7: The boxplots of variance accounted by different components across different stages during the experiment. The results were obtained by conducting principal component analysis on frequency domain. Epochs in the entire experiment have been classified as 6 stages with each consisting of 40 epochs (Stage I: 1-40, II: 41-80, III: 81-120, IV: 121-160, V: 161-200, VI: 201-247). The first component is shown on top and the first three cumulative components is at the bottom. We could observe that about 90% of variance can be explained by three components.

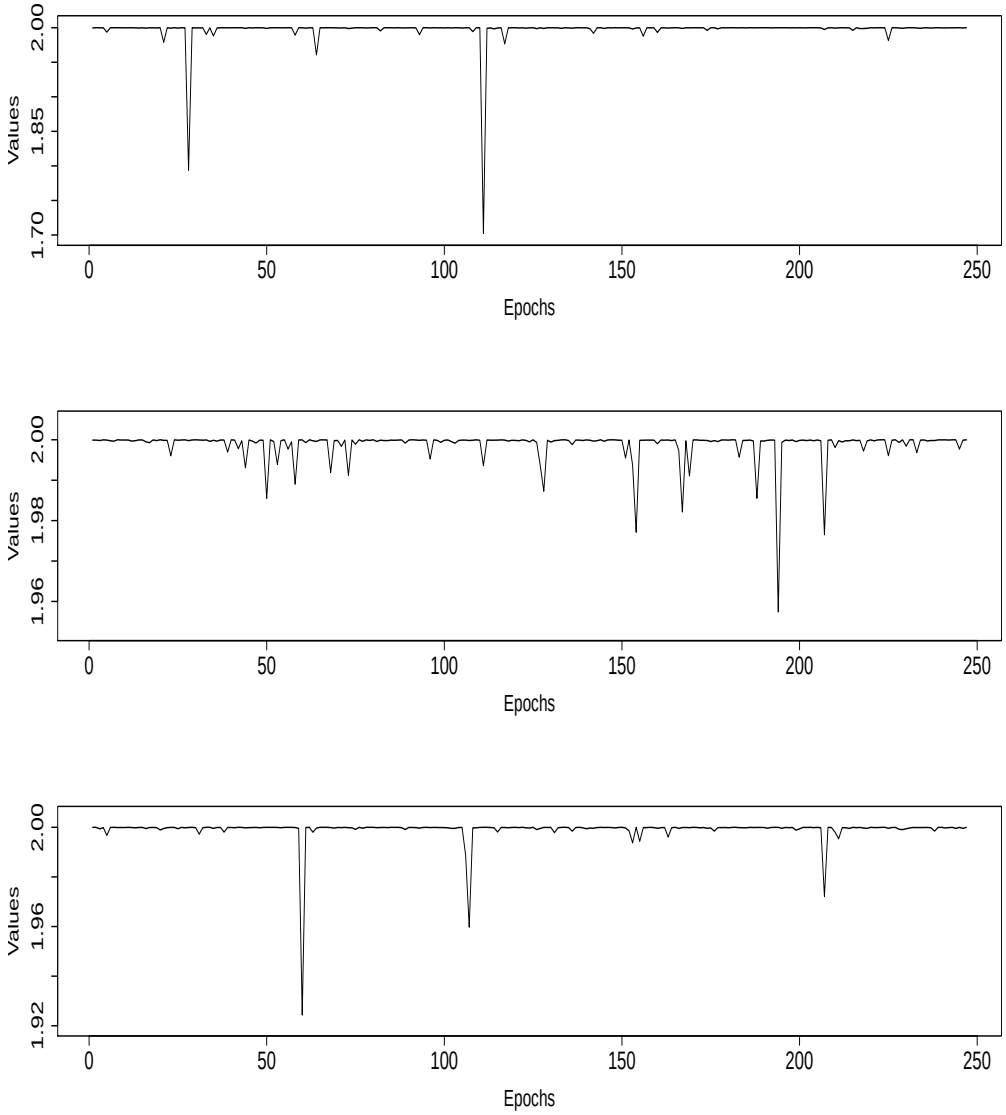

Figure 8: The time series plots of modulus corresponding to delta (above), alpha (middle) and gamma (bottom) frequency bands.

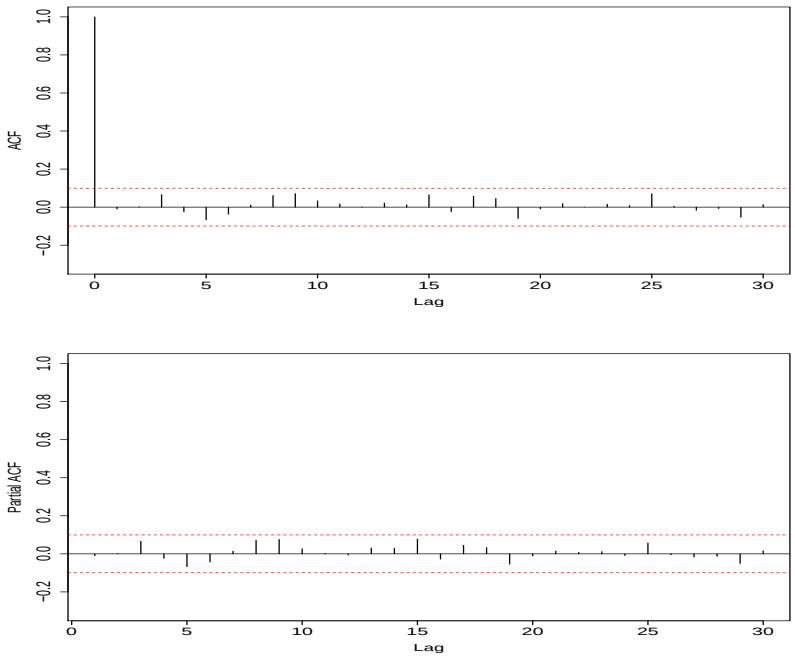

Figure 9: Top: Auto-correlation function (ACF) of the residual plots from electrode 1. Bottom: Partial auto-correlation function (PACF) of the residual plots from electrode 1. The dashed lines indicate the threshold for non-zero correlation. These plots, along with the Ljung-Box test for white noise ($p - value \approx 0.75$) suggest that the residuals are white noise and hence the E-SSM model fits the data well. These same plots were observed in all the other electrodes but we do not report them here due to space constraints.

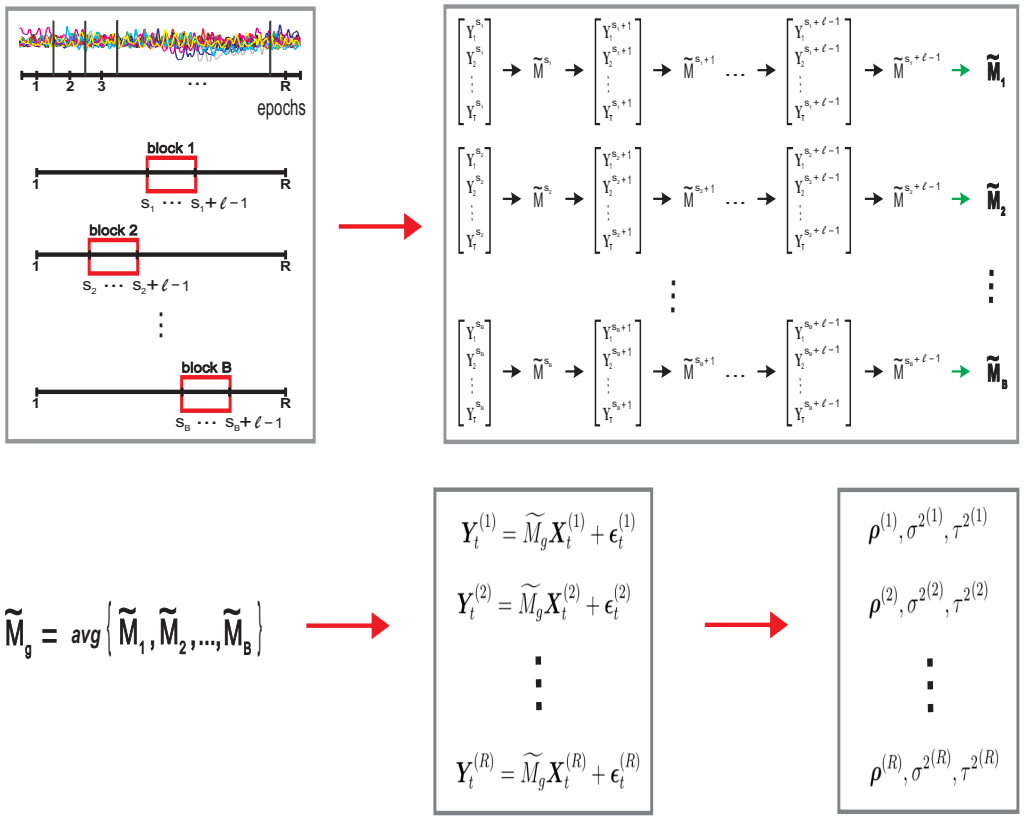

Figure 10: Schematic illustration of the estimation methods that summarize II.A, II.B and II.C in Section 3.

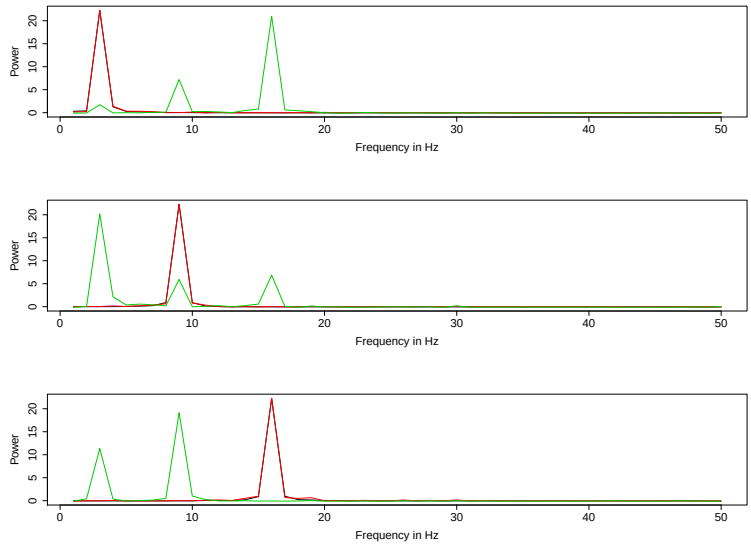

Figure 11: The periodograms of the latent processes obtained from the true signals (in black), the estimated signals from E-SSM (in red) and the ICA (in green).

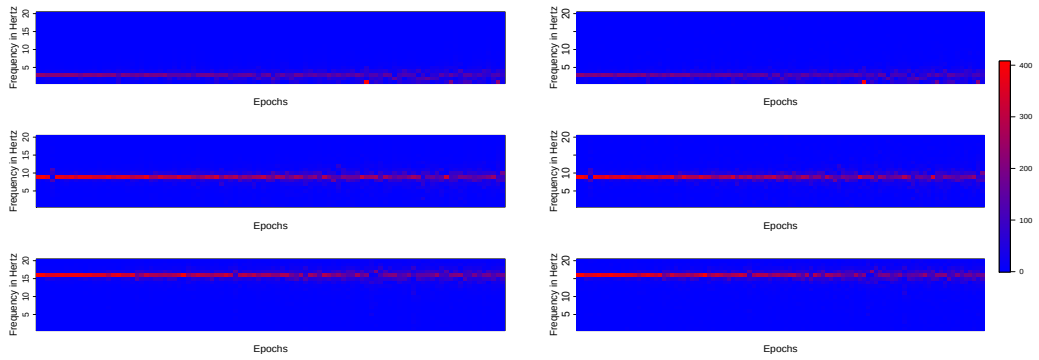

Figure 12: The periodograms of the true (left) and estimated (right) latent AR(2) processes obtained from E-SSM corresponding to delta (top), alpha (middle) and beta (bottom) frequency band.

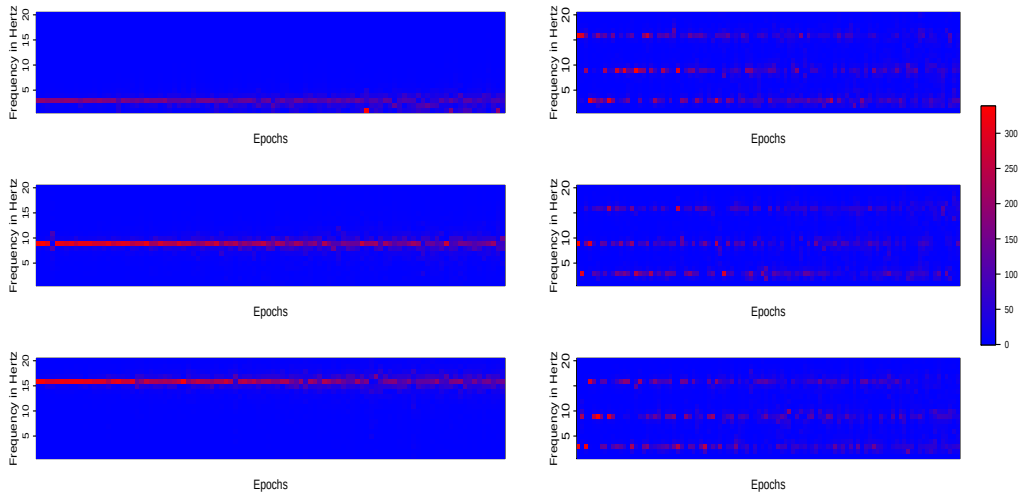

Figure 13: The periodograms of the true (left) processes corresponding to delta (top), alpha (middle) and beta (bottom) frequency band vs the "uncovered" sources obtained from ICA (right).

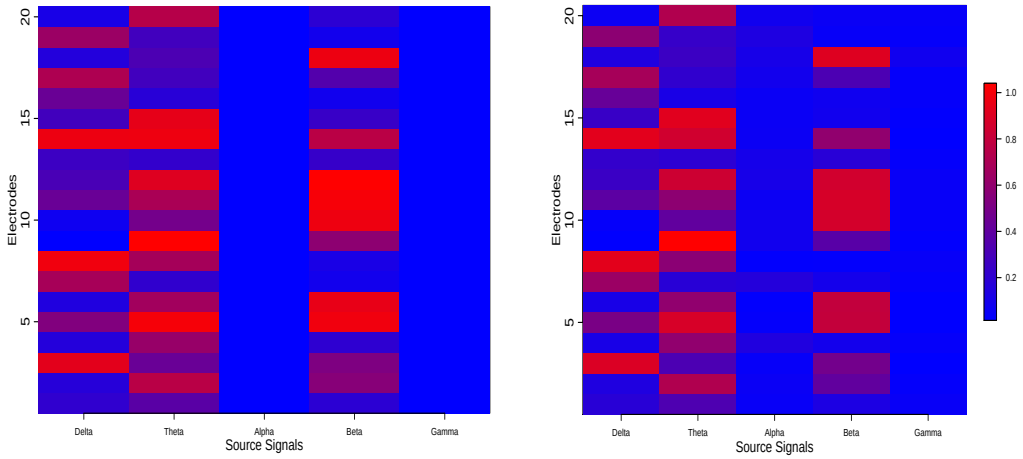

Figure 14: The true mixing matrix (left) and estimated mixing matrix (right) obtained from the proposed E-SSM. Darker color indicates heavier weight given by the corresponding latent processes. Columns corresponding to "alpha" and "gamma" bands are zero in the true mixing matrix (left). In the estimated mixing matrix (right), those two columns are also close to zero.

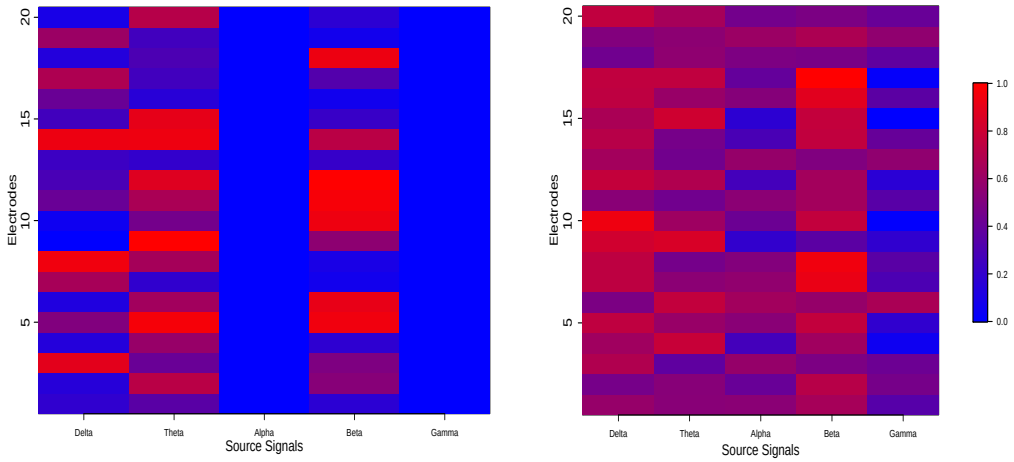

Figure 15: The true mixing matrix (left) and estimated mixing matrix (right) obtained from ICA. The weights corresponding to "alpha" and "gamma" bands are zero while ICA does not provide similar results.

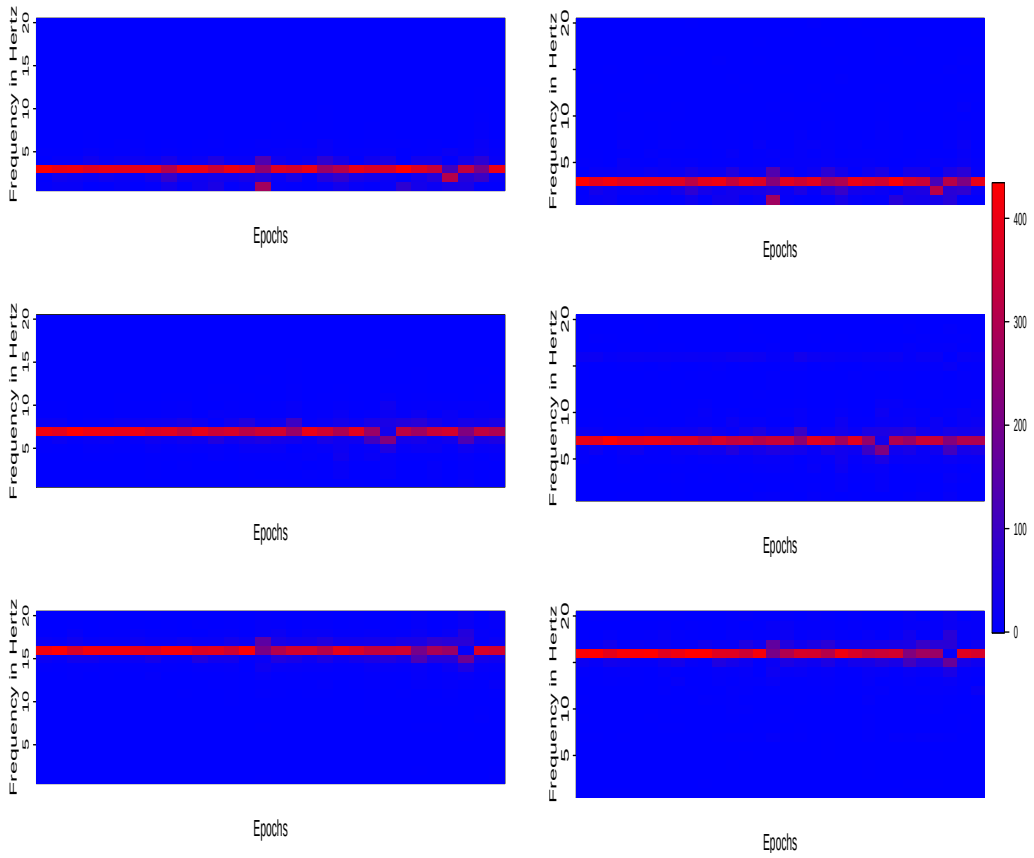

Figure 16: The periodograms of the true (left) and estimated (right) latent AR(2) processes corresponding to delta (top), theta (middle) and beta (bottom) frequency band obtained from E-SSM.

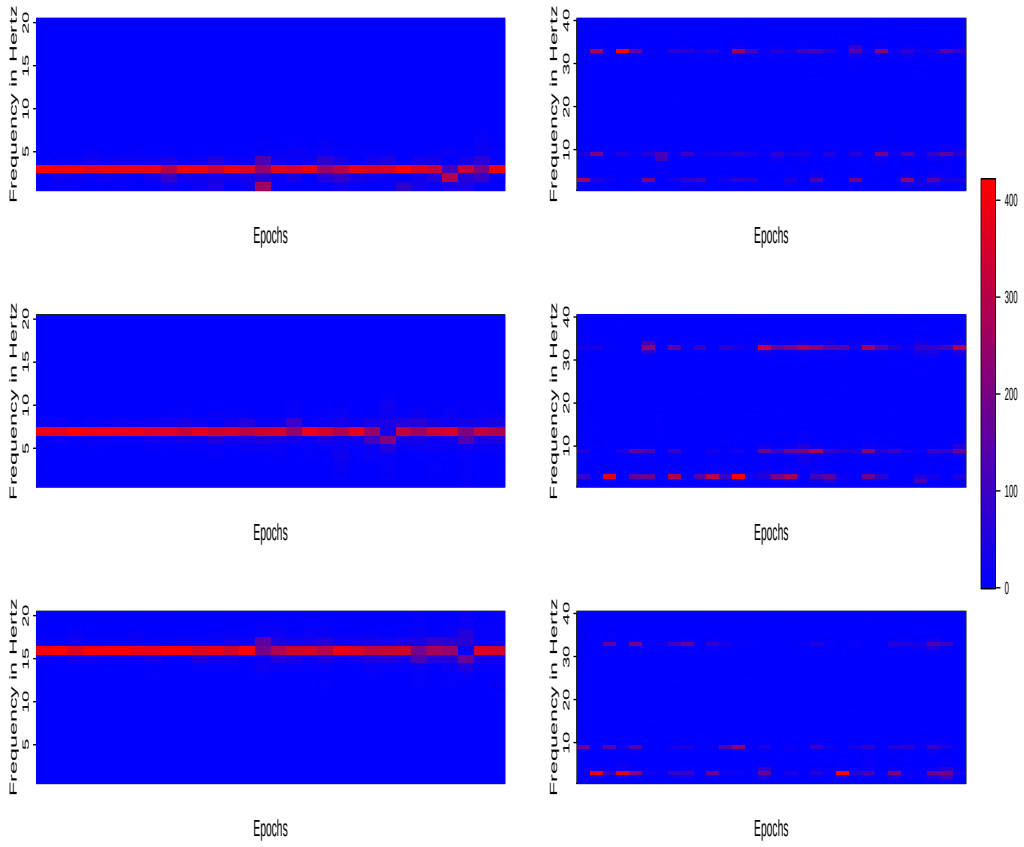

Figure 17: The periodograms of the true (left) process corresponding to delta (top), theta (middle) and beta (bottom) frequency band vs the "uncovered" sources obtained from ICA.

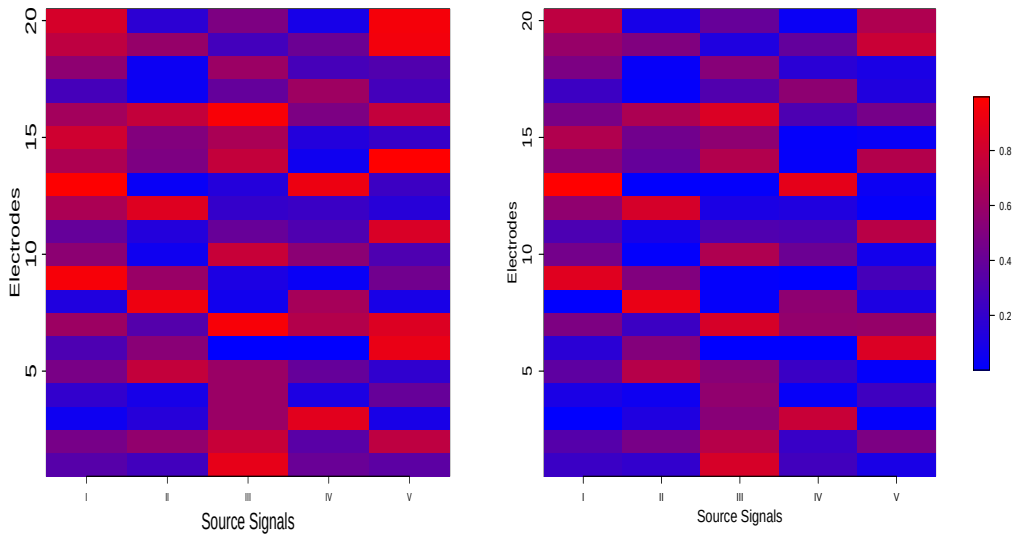

Figure 18: Sensitivity analysis Scenario 3: the true mixing matrix (left) and estimated mixing matrix (right) obtained from E-SSM.

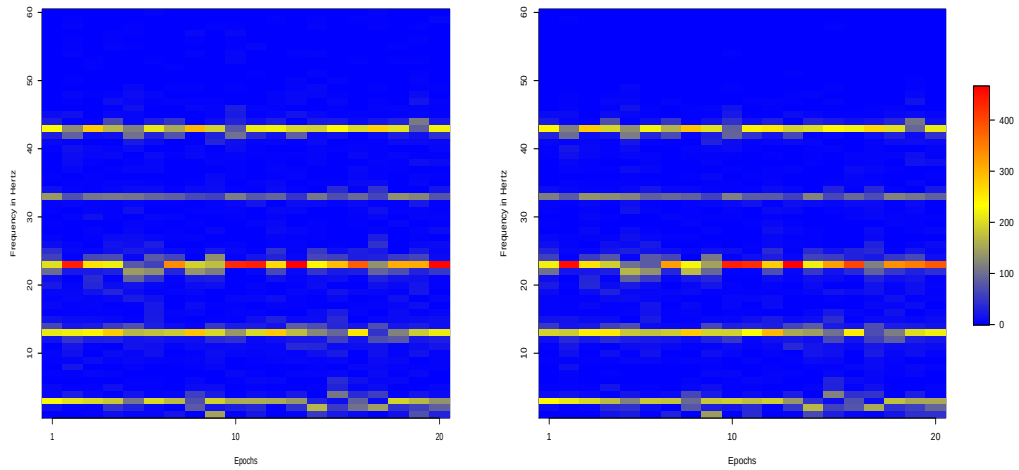

Figure 19: The periodograms of true (left) and estimated (right) signals from channel 16. Note that five peaks are present and the modulus are progressing across epochs.

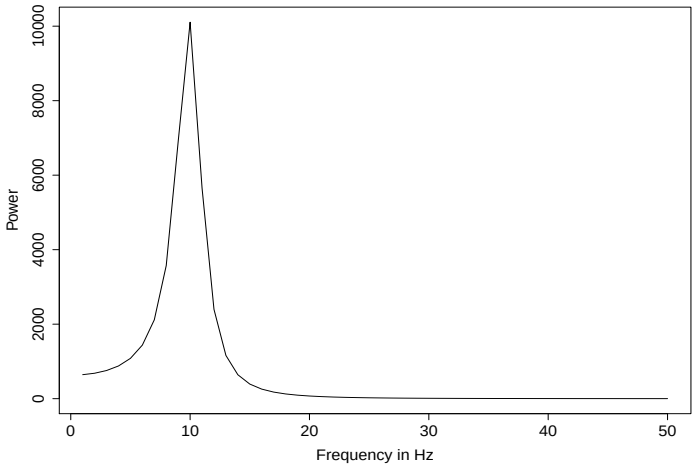

Figure 20: The theoretical spectra of an AR(2) process with power concentrated at the alpha band: $\phi_1 = 1.976, \phi_2 = -0.980, \sigma_w = 0.1$

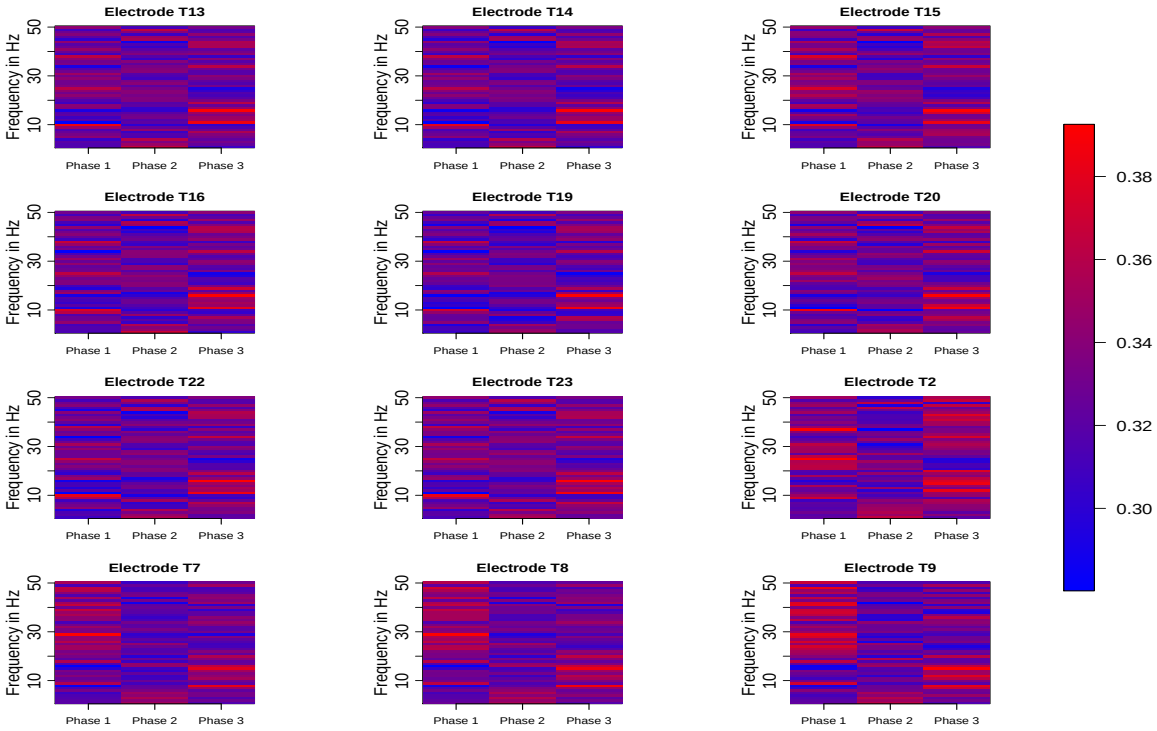

Figure 21: The evolution of the relative periodogram (summing up to 1 for each frequency) across the duration of experiment. Each plot displays the estimated power spectrum during the 3 phases: Phase 1 (epoch 1 - 80), Phase 2 (epoch 81 - 160) and Phase 3 (epoch 161 - 247). Frequency bands around particular hertz are present, which can be modeled by an AR(2).

S6. Tables

Table 1: Sensitivity analysis Scenario 1 – SSE (compared at 0.068).

	$mv_{order} = 1$	$mv_{order} = 3$	$mv_{order} = 5$	$mv_{order} = 10$	$mv_{order} = 50$	$mv_{order} = 100$
$mv_{coef} = 1$	0.103	0.122	0.121	0.137	0.689	0.757
$mv_{coef} = 5$	0.115	0.128	0.199	0.377	0.581	0.808
$mv_{coef} = 10$	0.123	0.179	0.374	0.489	0.493	0.943

Table 2: Sensitivity analysis Scenario 2 – SSE (compared at 0.068).

Uniform	$u = 0.01$	$u = 0.1$	$u = 1$	$u = 5$	$u = 10$
	0.267	0.405	0.436	0.442	0.682
Gaussian	$\sigma_n = 0.01$	$\sigma_n = 0.1$	$\sigma_n = 1$	$\sigma_n = 5$	$\sigma_n = 10$
	0.761	0.797	0.831	0.898	1.867

References

Allen, T. A., D. M. Salz, S. McKenzie, and N. J. Fortin (2016). Nonspatial sequence coding in ca1 neurons. *The Journal of Neuroscience 36*(5), 1547–1563.

Douc, R., E. Moulines, J. Olsson, and R. van Handel (2011). Consistency of the maximum likelihood estimator for general hidden markov models. *The Annals of Statistics 39*(1), 474–513.

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge university press.