Supplementary: State Space Methods for Phase Amplitude Coupling Analysis

Supplementary Section .1 Derivations

Power Spectral Density for the State Space Oscillator Process

In this section, we derive the parametric expression for the power spectral density (PSD) of an oscillation $x_j$ by building an autoregressive moving average process (ARMA) with the same spectral content. For convenience, we will drop the index $j$ in what follows. First, we note that an oscillation is asymptotically second order stationary. Let us compute its autocovariance sequence. Since $\mathcal{R}$ is a rotation matrix $\mathcal{R}(k\omega) = \mathcal{R}(k\mathcal{R})$ and $\mathcal{R}\mathcal{R}^T = I$. Therefore, from equation (1), it comes:

$$
\mathbb{E}(x_t x_t^\top) = a \mathcal{R}(\omega) \mathbb{E}(x_{t-1} x_{t-1}^\top) (a \mathcal{R}(\omega))^\top + Q
$$

$$
= Q \sum_{t=0}^{N} a^{2t}
$$

$$
= Q / (1 - a^2) + o(a^{2N})
$$

and:

$$
\mathbb{E}(x_{t+k} x_t^\top) = a \mathcal{R}(k\omega) \mathbb{E}(x_{t+k-1} x_t)
$$

$$
= a^k \mathcal{R}(k\omega) Q / (1 - a^2) + o(a^{2N+k})
$$

(34)

We can hence write, for $t = 1..N$, $k = 0..N - t$, $s_{t+k} = \mathbb{E}(x_{t+k} x_t^\top) = s_k$. As a consequence, an oscillation can be approximated by a second order stationary process, and in virtue of the Wiener-Khinchin theorem, its theoretical power spectral density is:

$$
\mathcal{S}(f) = \lim_{N \to \infty} \frac{1}{N} \left( s_0 + 2 \sum_{t=1}^{N} s_t e^{-2\pi f / N} \right)
$$

(35)

We now consider the ARMA(2,1):

$$
\bar{x}_t = \phi_1 \bar{x}_{t-1} + \phi_2 \bar{x}_{t-2} + \bar{u}_t + \psi_1 \bar{u}_{t-1}, \bar{u}_t \sim N(0, \bar{\sigma}^2)
$$

(37)

to which we impose, for $t = 1..N$, $k = 0..N - t$: $\mathbb{E}(\bar{x}_t \bar{x}_{t+k}^\top) = s_k$. It follows that:

$$
s_k = \phi_1 s_{k-1} + \phi_2 s_{k-2}, k > 2
$$

$$
s_1 = \frac{\phi_1 s_0 + \psi_1 \bar{\sigma}^2}{1 - \phi_2}
$$

$$
s_0 = \phi_1 s_1 + \phi_2 s_2 + \bar{\sigma}^2 (1 + \phi_1 \psi_1 + \psi_1^2)
$$

(38)

Taking:

$$
\phi_1 = 2a \cos(\omega)
$$

$$
\phi_2 = -a^2
$$

satisfies the first equality of equation (38). The remaining conditions can then be rewritten:

$$
\psi_1 \bar{\sigma}^2 = -a \sigma^2 \cos(\omega)
$$

$$
0 = \psi_1^2 + \frac{1 + a^2}{a \cos(\omega)} \psi_1 + 1
$$

(40)
We use the Kalman filter to estimate the hidden oscillations given the observations and the model parameters. They first predict

\[ r_{t} \text{ At iteration } t \]

An Expectation-Maximization (EM) Algorithm for Independent and Harmonic Oscillation Decompositions

Since all noise terms are assumed to be additive Gaussian, the complete data log likelihood for one time window of length \( N \) is:

\[
\log L = \log p(x_1, \ldots, x_N, y_1, \ldots, y_N) \\
\quad = -\frac{N}{2} \log |Q| - \frac{1}{2} \sum_{i=1}^{N} (x_i - \Phi x_{i-1})^T Q^{-1} (x_i - \Phi x_{i-1}) \\
\quad - \frac{N}{2} \log |R| - \frac{1}{2} \sum_{i=1}^{N} (y_i - M x_i)^T R^{-1} (y_i - M x_i) \\
\]  

Finally, the PSD of an oscillation centered in \( f_0 \) is:

\[
\mathcal{F}[x_1](f) = \hat{\sigma}^2 \frac{1 + \psi_t e^{-2\pi f/F_s}}{1 - \psi_t e^{-2\pi f/F_s} - \phi_2 e^{-4\pi f/F_s}} 
\]

An Expectation-Maximization (EM) Algorithm for Independent and Harmonic Oscillation Decompositions

Since all noise terms are assumed to be additive Gaussian, the complete data log likelihood for one time window of length \( N \) is:

\[
\log L = \log p(x_1, \ldots, x_N, y_1, \ldots, y_N) \\
\quad = -\frac{N}{2} \log |Q| - \frac{1}{2} \sum_{i=1}^{N} (x_i - \Phi x_{i-1})^T Q^{-1} (x_i - \Phi x_{i-1}) \\
\quad - \frac{N}{2} \log |R| - \frac{1}{2} \sum_{i=1}^{N} (y_i - M x_i)^T R^{-1} (y_i - M x_i) \\
\]  

We wish to maximize \( \log L \) with respect to \( \Theta = (\Phi, Q, R) \) but we do not have access to the hidden oscillations \( x_t \). We use an expectation maximization algorithm to alternatively and iteratively estimate (E-Step) and maximize (M-Step) the log likelihood. At iteration \( r \), we use the Kalman filter to estimate \( x_t \) given a set \( \Theta_r \) which gives us access to:

\[
G_r(\Theta) = \mathbb{E}_r(\log L | \{y_t\}_{t=1}^{N}) 
\]

Then, we deduce \( \Theta_{r+1} \):

\[
\Theta_{r+1} = \arg\max_{\Theta} G_r(\Theta) 
\]

Kalman Filter and Fixed Interval Smoother Estimates

We use the Kalman filter to estimate the hidden oscillations given the observations and the model parameters. They first predict the state at the next time point, then compare that prediction to the observation, and finally produce an updated estimate based on the most recently observed data. Given the full observation time series, we can apply backward smoothing to refine the update to account for the full observation series (i.e., fixed interval smoothing).

For \( t, t_1, t_2 = 1 \ldots N \), we note:

\[
x_t^N = \mathbb{E}_r(x_t | \{y_t\}_{t=1}^{N}), P_{t_1}^{N} = \text{cov}_r(x_{t_1}, x_{t_2} | \{y_t\}_{t=1}^{N}) \quad \text{and} \quad P_t^N = P_{t,t}^N 
\]

(48)
We note that we consider the oscillation decomposition with harmonic components. An oscillation is defined by a rotation matrix \( R \), a scaling parameter \( a \), and a harmonic component \( \Omega \). For \( j = 1, d \), we consider \( a \) independent oscillations, which are associated with \( h = 1, h_j \) harmonics. For \( j = 1, d \), an oscillation with fundamental frequency \( \omega_j \) is the sum of \( h = 1, h_j \) harmonics respectively defined by \( \mathcal{R}(h\omega_j) \), \( a_{j,h} \) and \( Q_{j,h} = \sigma^2_{j,h}I_{2 \times 2} \). We denote \( D = \sum_{j=1}^{d} h_j \) the total number of oscillatory components.

For \( V \in \mathbb{R}^{2D \times 2D} \), \( j = 1, d \) and \( h = 1, h_j \), we denote \( V_{j,h} \) the 2 by 2 diagonal block associated with the \( h_{j}^{th} \) harmonic of oscillation \( j \). \( \Phi \) and \( \Theta \) are block diagonal matrices whose diagonal blocks are \( a_{j,h}\mathcal{R}(h_{j,h}) \) and \( Q_{j,h} \):

\[
\Phi = \text{diag} \left( a_{1,1}\mathcal{R}(\omega_1), ... a_{1,h_1}\mathcal{R}(h_1\omega_1), ... , a_{d,1}\mathcal{R}(\omega_d), ... , a_{d,h_d}\mathcal{R}(h_d\omega_d) \right)
\]

\[
\Theta = \text{diag} \left( Q_{1,1}, Q_{1,h_1}, ... , Q_{d,1}, Q_{d,h_d} \right)
\]

Additionally, we will use \( M = [1 \ 0 \ 1 \ 0 \ ... \ 0] \in \mathbb{R}^{2D} \) and for \( U \in \mathbb{R}^{2 \times 2} \), we note: \( r(U) := U_{21} - U_{12} \) and \( \text{tr}(U) = U_{11} + U_{22} \).

Taking the conditional expectation of the log likelihood log \( L \) at iteration \( r \) for a fixed set of parameter \( \Theta_r = (\Phi, \Theta, R)_r \), we obtain:

\[
G(\Phi, \Theta, R)_r \approx -\frac{N}{2} \log |\Theta| - \frac{1}{2} \text{tr} \left( Q^{-1} \left( C - B\Phi^T - \Phi B^T + \Phi A \Phi^T \right) \right)
\]

\[
= -\frac{N}{2} \log |\Theta| - \frac{1}{2} \text{tr} \left( Q^{-1} \left( \sum_{i=1}^{N} (y_i - Mx_i) (y_i - Mx_i)^T + MP^N_i M^T \right) \right)
\]

\[
= G(\Phi, \Theta) + G(R)_r
\]

where:

\[
A = \sum_{i=1}^{N} (P_{i-1}^N + x_{i-1}^N (x_{i-1})^T)
\]

\[
B = \sum_{i=1}^{N} (P_{i-1}^N + x_{i-1}^N (x_{i-1})^T)
\]

\[
C = \sum_{i=1}^{N} (P_{i-1}^N + x_{i-1}^N (x_{i-1})^T)
\]
\section*{M-Step}

We can maximize $G_r$ with respect to $R$ and $(\Phi, Q)$ independently. We have:

\[ \frac{\partial G_r}{\partial R}(R_{r+1}) = 0 \iff R_{r+1} = \frac{1}{N} \sum_{t=1}^{N} \left( (y_t - MX_t^N)^2 + MP_t^N M^T \right) \]  

(55)

Since $Q$ is (block) diagonal, we can write:

\[ \text{tr}(Q^{-1}V) = \sum_{j=1}^{d} \sum_{h=1}^{h_j} \text{tr}(Q^{-1}_{j,h}V_{j,h}) = \sum_{j=1}^{d} \sum_{h=1}^{h_j} \frac{1}{\sigma_{j,h}^2} \text{tr}(V_{j,h}) \]  

(56)

$A$ is symmetric and $\Phi$ is a block diagonal matrix whose element are $2 \times 2$ rotation matrices, we develop equation (53) and obtain:

\[ G_r(\Phi, Q) \triangleq -N \sum_{j=1}^{d} \sum_{h=1}^{h_j} \log \sigma_{j,h}^2 \]

\[ - \sum_{j=1}^{d} \sum_{h=1}^{h_j} \frac{1}{2\sigma_{j,h}^2} \left[ \text{tr}(C_{j,h}) - 2a_{j,h}(\text{tr}(B_{j,h}) \cos(h\omega_j) + \text{tr}(B_{j,h}) \sin(h\omega_j)) + a_{j,h}^2 \text{tr}(A_{j,h}) \right] \]

(57)

Differentiating with respect to process noises covariances $\sigma_{j,h}^2$, scaling parameter $a_{j,h}$ and fundamental frequencies $\omega_j$ yields:

\[
\begin{cases}
\frac{\partial G_r}{\partial a_{j,h}}(\Phi, Q) = 0 \iff a_{j,h} \text{tr}(A_{j,h}) = \text{tr}(B_{j,h}) \cos(h\omega_j) + \text{tr}(B_{j,h}) \sin(h\omega_j) \\
\frac{\partial G_r}{\partial \sigma_{j,h}^2}(\Phi, Q) = 0 \iff \sigma_{j,h}^2 = \frac{1}{2N} \left( \text{tr}(C_{j,h}) - a_{j,h}^2 \text{tr}(A_{j,h}) \right) \\
\frac{\partial G_r}{\partial \omega_j}(\Phi, Q) = 0 \iff \sum_{h=1}^{h_j} \frac{h}{\sigma_{j,h}^2} \left[ a_{j,h}(\text{tr}(B_{j,h}) \sin(h\omega_j) - \text{tr}(B_{j,h}) \cos(h\omega_j)) \right] = 0
\end{cases}

(58)

We inject the upper equations of (58) into the third one and we note:

\[ \tilde{\omega}_{j,h} = \frac{1}{h} \tan^{-1} \left( \frac{\text{tr}(B_{j,h})}{\text{tr}(B_{j,h})} \right) \quad \text{and} \quad S_{j,h} = \frac{2\text{tr}(A_{j,h})\text{tr}(C_{j,h})}{\text{tr}(B_{j,h})^2 + \text{tr}(B_{j,h})^2} - 1 \]

(59)

Using trigonometric identities, equations (58) can finally be rewritten:

\[
\begin{cases}
\frac{\partial G_r}{\partial a_{j,h}}(\Phi, Q) = 0 \iff a_{j,h} \text{tr}(A_{j,h}) = \text{tr}(B_{j,h}) \cos(h\omega_j) + \text{tr}(B_{j,h}) \sin(h\omega_j) \\
\frac{\partial G_r}{\partial \sigma_{j,h}^2}(\Phi, Q) = 0 \iff \sigma_{j,h}^2 = \frac{1}{2N} \left( \text{tr}(C_{j,h}) - a_{j,h}^2 \text{tr}(A_{j,h}) \right) \\
\frac{\partial G_r}{\partial \omega_j}(\Phi, Q) = 0 \iff \sum_{h=1}^{h_j} \frac{h \sin(2h(\omega_j - \tilde{\omega}_{j,h}))}{S_{j,h} - \cos(2h(\omega_j - \tilde{\omega}_{j,h}))} = 0
\end{cases}

(60)

Overall, for $j = 1..d$, if $h_j > 1$, we numerically solve for $\omega_j$ using equation (60) and deduce $a_{j,h}$ and $\sigma_{j,h}^2$ for $h = 1..h_j$.

If $h_j = 1$, we immediately have:

\[ w_j = \tan^{-1}(\text{rt}(B_j)/\text{tr}(B_j)) \quad a_j = \sqrt{\text{tr}(B_j)^2 + \text{tr}(B_j)^2} \quad \sigma_j^2 = \frac{1}{2N} \left( \text{tr}(C_j) - a_j^2 \text{tr}(A_j) \right) \]

(61)
Linear Regression: priors, hyperparameters, and normalizing constants

As in⁴, we use $\tau_\beta = \sigma_\hat{\beta}^{-2}$ and we assume that the likelihood of the model defined in equation (21) is:

$$p(A^t, \phi^t|\beta, \tau_\beta) \propto \tau_\beta^{-N/2} \exp \left( -\frac{\tau_\beta}{2} (A^t - X(\phi^t)\hat{\beta})^T (A^t - X(\phi^t)\hat{\beta}) \right) 1_{(\beta \in W(\chi))}$$

(62)

where $\{A^t\}_t = A^t$ and $\{\phi^t\}_t = \phi^t$. The conjugate prior is:

$$p(\beta, \tau_\beta) \propto \tau_\beta^{\frac{\nu+1}{2}-1} \exp \left( -\frac{\tau_\beta}{2} \left[ \hat{\beta} \nu + (\beta - \beta)\nu (\beta - \beta) \right] \right) 1_{(\beta \in W(\chi))}$$

(63)

We choose prior hyperparameters $\nu, \tilde{b}, \tilde{v}$ and $\bar{\beta} = [\bar{\beta}_0, \bar{\beta}_1, \bar{\beta}_2]^{\top}$ to convey as little information as possible on the phase and the strength of the modulation. Marginalizing equation (63) over $\tau_\beta$, yields a truncated multivariate t-distribution:

$$p(\beta) \propto \left( 1 + \frac{1}{\nu} (\beta - \bar{\beta})^{\top} (\nu/\tilde{b}) (\beta - \bar{\beta}) \right)^{-\frac{\nu+1}{2}} 1_{(\beta \in W(\chi))}$$

(64)

$\nu \geq 3$ insures that the multivariate-t variance is defined. It is: $\nu(\nu - 2)^{-1} \nu^{-1}$. Then, we consider the independent random variables $\check{A}, \check{K}$ and $\check{\phi}$ such that:

$$\check{A} \sim \Gamma(A_0/c, c)$$

$$\check{K} \sim \text{Uniform}(0, 1)$$

$$\check{\phi} \sim \text{Uniform}(-\pi, \pi)$$

(65)

We note $\gamma = [\check{A} \quad \check{A}\check{K}\cos(\check{\phi}) \quad \check{A}\check{K}\sin(\check{\phi})]^{\top}$, use $\bar{b} = 1$, $\bar{v} = 3$ and we define $\check{\beta}$, and $\check{\nu}$ such that:

$$\check{\beta} = \mathbb{E}(\gamma) \text{ and } \check{\nu}^{-1} = \text{Cov}(\gamma)(\bar{v} - 2)/\check{\nu}$$

(66)

Additionally, we notice that if $A^t_i \approx A_0[1 + K^{\text{mod}} \cos(\phi^t + \phi^{\text{mod}})]$ and since $\mathbb{E}_s(\cos(\phi^t + \phi^{\text{mod}})) = 0$ (where $\mathbb{E}_s$ represents an average over trials or windows and $\langle \cdot \rangle_{t}$ is a temporal average across a given window), $\mathbb{E}_s(\langle A^t \rangle_{t}) = A_0$ and $\langle A^t \rangle_{t} \in [0, 2A_0]$. Therefore, it is reasonable to use $A_0 = \langle A^t \rangle_{t}$ and $c = 1$. Overall:

$$\check{\nu} = 3, \check{\nu} = \langle A^t \rangle^{-1}_{t} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 12 \end{pmatrix}$$

$$\check{\beta} = \begin{pmatrix} \langle A^t \rangle_{t} \\ 0 \\ 0 \end{pmatrix} \text{ and } \bar{b} = 1$$

(67)

Posterior parameters are then given by:

$$\nu = \check{\nu} + N, \quad V = \check{\nu} + X(\phi^{\text{mod}})^{T} X(\phi^{\text{mod}})$$

$$\bar{\beta} = V^{-1} \left( \check{\nu} \check{\beta} + X(\phi^{\text{mod}}) A^t \right), \quad b = (\check{\nu} \bar{b} + H)/\nu$$

(68)

Where:

$$\hat{\beta}_{\text{OLS}} = (X(\phi^{\text{mod}})^{T} X(\phi^{\text{mod}}))^{-1} X(\phi^{\text{mod}})^{T} A^t \quad \text{and:}$$

$$H = (A^t - X(\phi^{\text{mod}}) \hat{\beta}_{\text{OLS}}) (A^t - X(\phi^{\text{mod}}) \hat{\beta}_{\text{OLS}})$$

$$+ (\hat{\beta}_{\text{OLS}} - \bar{\beta})^{T} X(\phi^{\text{mod}}) X(\phi^{\text{mod}}) (\hat{\beta}_{\text{OLS}} - \bar{\beta})$$

$$+ (\bar{\beta} - \bar{\beta})^{T} \check{\nu} (\bar{\beta} - \bar{\beta})$$

(69)
We derive the posterior distribution:

\[ p(\beta, \tau_\beta | (A_t^\tau, \phi_t^\tau)_t) = \frac{1}{Z} \tau_\beta^{-\frac{\nu + 3}{2}} \exp \left( -\frac{\tau_\beta}{2} \left[ b \nu + (\beta - \bar{\beta})^T V (\beta - \bar{\beta}) \right] \right) 1(\nu > 0) \]  

(70)

The normalizing constant \( Z \) is obtained by integration and is:

\[ Z = \frac{\Gamma(\nu/2)(2\pi)^{3/2}}{|V|^{1/2}(\nu b/2)^{\nu/2}} P(\beta \in W(\mathcal{R})) \]  

(71)

where \( P(\beta \in W(\mathcal{R})) \) is computed using the multivariate t-distribution of parameters \( \nu, \bar{\beta} \) and \( b^{-1}V \).

Finally, we deduce equation (24) by marginalizing equation (70) over \( \tau_\beta \) and have:

\[ Z = \frac{\Gamma(\nu/2)(2\pi)^{3/2}}{\Gamma(1/2 + 3/2 + \nu/2)} \times P(\beta \in W(\mathcal{R})) \]  

(72)

Note that for large samples it might be useful to use:

\[ \frac{\Gamma((\nu + 3)/2)}{\Gamma(\nu/2)} = \left( \frac{\nu + 1}{2} \right) \left( \frac{\nu}{2} \right)^{1/2} \left( 1 - \frac{1}{4\nu} + \frac{1}{32\nu^2} + \frac{1}{128\nu^3} + O\left( \frac{1}{\nu^4} \right) \right) \]  

(73)

**Second State-Space**

In this section, for a given AR order \( p \), we estimate the parameters \( R_\beta, Q_\beta \) and \( \{h_k\}_{k=1}^p \) defined equation (6). Let \( C_m = \mathbb{E}(\beta_t \beta_{t-m}^\top) \) be the autocovariance sequence of the modulation vectors estimated with equation (25). We have:

\[ C_m = \sum_{k=1}^p h_k C_{m-k} + R_\beta (\delta_{0,m} - h_m) + Q \delta_{0,m} \]  

(74)

where \( \delta_{i,j} \) is the Kronecker delta. Equation (74) can be rewritten:

\[
\begin{pmatrix}
C_0 \\
C_1 \\
\vdots \\
C_p
\end{pmatrix} = \begin{pmatrix}
\sum_{k=1}^p h_k C_k + Q \\
C_0 - R_\beta & C_1 & \ldots & C_{p-1} \\
\vdots & \vdots & \ddots & \vdots \\
C_{p-1} & C_1 & \ldots & C_0 - R_\beta
\end{pmatrix} \begin{pmatrix}
h_1 \\
h_2 \\
\vdots \\
h_p
\end{pmatrix}
\]  

(75)

For an observation noise candidate \( R_\beta \), if we can invert equation (75), we immediately access \( Q_\beta \) and \( \{h_k\}_{k=1}^p \). Using the Kalman Filter, we hence deduce the likelihood of the candidate model as in 22.

Therefore, we note \( R_\beta \) the smallest eigenvalue of the Toeplitz matrix \( C = (C_{i-j})_{i,j=0..p} \), and, numerically optimize the model likelihood with respect to \( R_\beta \) in \((0,R_\beta^{\text{m}})\), where we know that \( (C - IR_\beta) \) is full rank.

From \( R_\beta \) we get \( Q_\beta \) and \( \{h_k\}_{k=1}^p \) then, once again, we use the Kalman Filter to estimate \( \beta^{\text{esp}}_r \).

**Modulogram Equivalence**

We derive an approximated parametric modulogram for a window of length \( \delta t = N/F_s \) centered in \( \tau \). We will use:
Additionally, we assume that:

- For clarity we will use \( t \) without distinction and we remind that:

\[
A_k^f = A_0\left[1 + K_{\tau}^{\text{mod}} \cos(\phi_k^\tau - \phi_{\tau}^{\text{mod}})\right] + \varepsilon_k, \ v_k \sim \mathcal{N}(0, \sigma_k^2) \tag{77}
\]

\[
P_{\text{A}}(\tau, \psi) = \frac{\int_{r^\tau+\delta t/2}^{t^\tau-\delta t/2} \int_{\psi^\tau-\delta \psi/2}^{\psi^\tau+\delta \psi/2} A_k^f \delta(\phi_{\tau}^{\text{mod}} - \psi') \, dt \, d\psi'}{2\pi \int_{r^\tau+\delta t/2}^{t^\tau-\delta t/2} A_k^f \, dt} = \frac{P_1}{P_2} \tag{78}
\]

Additionally, we assume that:

- for \( k \in \Omega_{\tau} \), \( \phi_k^\tau = \frac{\theta_k}{F_s} + \eta_k \), where \( E(\eta_k) = 0 \), \( \text{Var}(\eta_k) = \sigma^2_\eta \) and \( \sigma^2_\eta << \frac{\theta_k}{F_s} < 1 \)
- and, for simplicity, for \( \psi \in [-\pi, \pi] \), for all \( h : \mathbb{R}^+ \rightarrow \mathbb{R} \) smooth, \( \sum_{k \in \Omega_{\tau, \psi}} h(k) \approx \sum_{k \in \Omega_{\tau, \psi}} h(k) \)

From the central limit theorem, \( \sum_{k \in \Omega_{\tau}} \varepsilon_k = \mathcal{O}_p(\sqrt{N}) \) and \( \sum_{k \in \Omega_{\tau}} \eta_k = \mathcal{O}_p(\sqrt{N}) \).

We hence have:

\[
P_1 = \frac{A_0}{F_s} \sum_{k \in \Omega_{\tau}} \left(1 + K_{\tau}^{\text{mod}} \cos \left(\frac{\omega_k}{F_s} - \phi_{\tau}^{\text{mod}} + \eta_k\right)\right) + \frac{1}{F_s} \sum_{k \in \Omega_{\tau}} \varepsilon_k
\]

\[
= \frac{A_0}{F_s} \sum_{k \in \Omega_{\tau}} \left(1 + K_{\tau}^{\text{mod}} \cos \left(\frac{\omega_k}{F_s} - \phi_{\tau}^{\text{mod}}\right)\right) + \frac{1}{F_s} \sum_{k \in \Omega_{\tau}} \left(\varepsilon_k - \frac{A_0}{F_s} K_{\tau}^{\text{mod}} \sin \left(\frac{\omega_k}{F_s} - \phi_{\tau}^{\text{mod}}\right) \eta_k + \mathcal{O}(\eta_k)\right) \tag{79}
\]

But \( \sum_{k \in \Omega_{\tau}} \cos \left(\frac{\omega_k}{F_s} - \phi_{\tau}^{\text{mod}}\right) = \cos \left(\frac{\omega_k}{F_s} (N - 1) - \phi_{\tau}^{\text{mod}}\right) \frac{\sin(Nw_k/(2F_s))}{\sin(\omega_k/(2F_s))} \leq \frac{2F_s}{\omega_k} \). Therefore:

\[
P_1 = \frac{A_0}{F_s} \frac{N}{\sqrt{N}} + \mathcal{O}_p(\sqrt{N}) \tag{80}
\]

On the other hand:

\[
P_2 = \frac{1}{F_s} \sum_{k \in \Omega_{\tau, \psi}} A_k^f
\]

\[
= \frac{A_0}{F_s} \sum_{k \in \Omega_{\tau, \psi}} \left(1 + K_{\tau}^{\text{mod}} \cos \left(\frac{\omega_k}{F_s} - \phi_{\tau}^{\text{mod}} + \eta_k\right)\right) + \frac{1}{F_s} \sum_{k \in \Omega_{\tau, \psi}} \varepsilon_k
\]

But \( \sum_{k \in \Omega_{\tau, \psi}} \varepsilon_k = \mathcal{O}_p(\delta \psi \sqrt{l}) \) and \( l \propto N \) so:

\[
P_2 = \frac{A_0}{F_s} \sum_{k \in \Omega_{\tau, \psi}} \left(1 + K_{\tau}^{\text{mod}} \cos \left(\frac{\omega_k}{F_s} - \phi_{\tau}^{\text{mod}} + \eta_k\right)\right) + \mathcal{O}_p(\sqrt{N}) \tag{82}
\]
Using the same argument as the one detailed above we get:

\[ P_2 = \frac{A_0}{F_s} \sum_{k \in \tilde{\Omega}_{\tau, \psi}} \left( 1 + K_\tau^{\text{mod}} \cos \left( \frac{\omega_k}{F_s} \phi_{\tau}^{\text{mod}} \right) \right) + O_p(\sqrt{N}) \]  

(83)

Additionally:

\[ \frac{A_0}{F_s} \sum_{k \in \tilde{\Omega}_{\tau, \psi}} \left( 1 + K_\tau^{\text{mod}} \cos \left( \frac{\omega_k}{F_s} - \phi_{\tau}^{\text{mod}} \right) \right) \]

\[ = \frac{A_0}{\delta \psi} \int_{\tau - \frac{\phi}{2}}^{\tau + \frac{\phi}{2}} \int_{\psi - \frac{\delta \psi}{2}}^{\psi + \frac{\delta \psi}{2}} (1 + K_\tau^{\text{mod}} \cos \left( \omega_k t - \phi_{\tau}^{\text{mod}} \right) \right) \delta_{\tilde{\Omega}_{\tau, \psi}} dtd\psi + \gamma(\omega_s / F_s) \]

(84)

Where \( \gamma \) is a function such that \( \gamma(x) \rightarrow 0 \). Since:

\[ \int_{\tau - \frac{\phi}{2}}^{\tau + \frac{\phi}{2}} \int_{\psi - \frac{\delta \psi}{2}}^{\psi + \frac{\delta \psi}{2}} (1 + K_\tau^{\text{mod}} \cos \left( \omega_k t - \phi_{\tau}^{\text{mod}} \right) \right) \delta_{\tilde{\Omega}_{\tau, \psi}} dtd\psi + \gamma(\omega_s / F_s) \]

\[ = \frac{L}{\omega_s} \times \int_{\psi - \frac{\delta \psi}{2}}^{\psi + \frac{\delta \psi}{2}} (1 + K_\tau^{\text{mod}} \cos \left( \phi' - \phi_{\tau}^{\text{mod}} \right) \right) d\phi' + O(1) \]

\[ = \frac{L \delta \psi}{\omega_s} \left( 1 + \frac{\sin \left( \delta \psi / 2 \right)}{\delta \psi / 2} K_\tau^{\text{mod}} \cos \left( \psi - \phi_{\tau}^{\text{mod}} \right) \right) + O(1) \]  

(85)

For \( \frac{\omega_s}{F_s} \) "sufficiently small", we can write:

\[ P_2 = \frac{A_0}{\omega_s} \left( 1 + \frac{\sin \left( \delta \psi / 2 \right)}{\delta \psi / 2} K_\tau^{\text{mod}} \cos \left( \psi - \phi_{\tau}^{\text{mod}} \right) \right) + O_p(\sqrt{N}) \]  

(86)

Finally:

\[ \text{PAM}(\tau, \psi) = \frac{1}{2\pi} \left( 1 + \frac{\sin \left( \delta \psi / 2 \right)}{\delta \psi / 2} K_\tau^{\text{mod}} \cos \left( \psi - \phi_{\tau}^{\text{mod}} \right) + O_p(1/\sqrt{N}) \right) \]

\[ = \frac{1}{2\pi} \left( 1 + \frac{\sin \left( \delta \psi / 2 \right)}{\delta \psi / 2} K_\tau^{\text{mod}} \cos \left( \psi - \phi_{\tau}^{\text{mod}} \right) \right) + O_p(1/\sqrt{N}) \]  

(87)
Supplementary Section .2 Additional Results

In this section we present additional results to support the validation and comparison of our algorithm:

Fig. S.1 is a simulated signal composed of three sinusoidal oscillations (A) where the amplitude of the fastest oscillation (C) was modulated by the phase of the middle one (E). The modulation peaked at an angle of $\pi/3$. Both information criterion (AIC and BIC) select the correct number of component (D) and our model correctly infers the latent components (B-C-E) and the phase amplitude coupling (F).

![Figure S.1](image1.png)

**Figure S.1.** Model selection and Phase Amplitude Coupling (PAC) in the case of multiple oscillations. The full Signal (A) is decomposed using 3 oscillations (B-C-E) according to model selection criteria (AIC or BIC) (D). Then PAC (F) is estimated by looking at pairs of oscillations. The arrow indicates the true coupling phase between oscillation 2 and 3, which is correctly inferred by our algorithm. Credible intervals are estimated using 100 samples.

Fig. S.2 compares the two methods used in this work in order to estimate spectrograms: (A) multitaper-spectrograms are directly computed from data (Figs. 4, 8) while (B) parametric-spectrograms (Fig. 2) are analytically derived from the model see equation (44).

![Figure S.2](image2.png)

**Figure S.2.** Comparison between Multitaper (A) and Parametric (B) spectrograms using data from Fig. 2
Figure S.3. Phase Amplitude Coupling profile of another subject infused with increasing target effect site concentration of propofol. Left: response probability curves (a) aligned with modulograms (c) (distribution of alpha amplitude with respect to slow phase) computed with standard (top) and parametric (bottom) techniques. Right: propofol infusion target concentration (b) aligned with corresponding Modulation Indices (d). Standard technique significance was assessed using 200 random permutations and CI where estimated using 200 × 200 samples.
Fig. S.3 is the PAC profile of another subject infused with increasing target effect site concentration of propofol.

Fig. S.4 is a comparison of the modulation dynamic estimation between standard analysis and our dSSP.

Fig. S.5 are typical 6s signal traces of signal generated with equation (12).

Fig. S.6, S.7 and S.8 Are the modulation phase ($\phi^{\text{mod}}$) and frequency ($f_s, f_f$) recovery estimation and comparison between standard methods, DAR and SSP.

Figure S.4. Comparison of the modulation estimates using standard methods and our new dSSP method. Slow and fast oscillations were generated by filtering white noise around $f_s = 1\text{Hz}$ and $f_f = 10\text{Hz}$ with $\Delta f_{\text{gen}} = 1\text{Hz}$ and normalized to standard deviation $\sigma_s = 1$ and $\sigma_f = 0.8$. The time scale over which $K^{\text{mod}}$ and $\phi^{\text{mod}}$ changed varied between 20 minutes to 5 and 2 minutes. See Fig. S.5 for typical signal traces.
Figure S.5. Typical 6s signal traces generated with equation (12) for 4 different conditions. The raw signal (a) is the sum of the observation noise (b), slow, and alpha oscillations (c). We show slow frequency multitaper PSD (TW=4, K=5 tapers) for different values of $\Delta f_{gen}^s$ (d).

Figure S.6. Modulation phase $\phi_{mod}$ estimation and comparison between standard methods (black), DAR (pink) and SSP (blue). 400 windows of 2 seconds were generated with: a slow oscillation (filtered from white noise around $f_s = 3\text{Hz}$ with $\Delta f_{gen}^s$ and normalized to standard deviation $\sigma_s$) and a modulated fast oscillation ($\phi_{mod} = -\pi/3$, modeled with a sinusoid $f_s = 50\text{Hz}$ and normalized to $\sigma_f$). We added unit normalized Gaussian noise and we used 3 Signal To Noise Ratio (SNR) conditions ($(\sigma_s, \sigma_f) = (2, 1.5), (1, 0.6)$ and $(0.7, 0.3)$). We show typical signal traces for these different conditions in Fig S.8.
Figure S.7. Slow and Fast oscillation recovery using different algorithms: standard methods (black), DAR (pink) and SSP (blue). 400 windows of 6 seconds were generated with: a slow oscillation (filtered from white noise around $f_s = 1$Hz with bandwidth $\Delta f^{gen}$ and normalized to standard deviation $\sigma_s$) and a modulated fast oscillation ($\phi^{mod} = -\pi/3$, modeled with a sinusoid $f_s = 10$Hz and normalized to $\sigma_f$). We added unit normalized Gaussian noise and we used 3 Signal To Noise Ratio (SNR) conditions (($\sigma_s, \sigma_f$) = (2, 1.5), (1, 0.6) and (0.7, 0.3)). We report typical signal traces for the different conditions (top), slow oscillation recovery alongside the true slow frequency PSD (middle), and fast frequency recovery (bottom). The red arrow indicates the true multitaper PSD (TW=4, K=5 tapers) peak for each oscillation.

Figure S.8. Slow and Fast oscillation recovery using different algorithm: standard methods (black), DAR (pink) and SSP (blue). 400 windows of 2 seconds were generated with: a slow oscillation (filtered from white noise around $f_s = 3$Hz with bandwidth $\Delta f^{gen}$ and normalized to standard deviation $\sigma_s$) and a modulated fast oscillation ($\phi^{mod} = -\pi/3$, modeled with a sinusoid $f_s = 50$Hz and normalized to $\sigma_f$). We added unit normalized Gaussian noise and we used 3 Signal To Noise Ratio (SNR) conditions (($\sigma_s, \sigma_f$) = (2, 1.5), (1, 0.6) and (0.7, 0.3)). We report typical signal traces for the different conditions (top), slow oscillation recovery alongside the true slow frequency PSD (middle), and fast frequency recovery (bottom). The red arrow indicates the true multitaper PSD (TW=4, K=5 tapers) peak for each oscillation.