Derivation of the ML solutions

The NLL cost

The estimated normalized PSD matrix is

$$\hat{\mathbf{\Gamma}}_x = \begin{bmatrix} 1 & \hat{\gamma}_x \\ \hat{\gamma}_x^* & 1 \end{bmatrix} \tag{1}$$

where $\gamma_x(\omega) = \frac{E[X_1(\omega,t)X_2^*(\omega,t)]}{\sqrt{E[|X_1(\omega,t)|^2]E[|X_2(\omega,t)|^2]}}$, and $\hat{\gamma}_x(\omega)$ is an estimate of $\gamma_x(\omega)$ obtained by replacing the expectations with sample averages.

The true PSD matrix is supposed to have the form

$$\mathbf{\Gamma}_x = P_s \begin{bmatrix} 1 & e^{j\phi} \\ e^{-j\phi} & 1 \end{bmatrix} + P_v \begin{bmatrix} 1 & \gamma_v \\ \gamma_v^* & 1 \end{bmatrix} = P_s \mathbf{\Phi} + P_v \mathbf{\Gamma}_v \tag{2}$$

where P_s and P_v are the powers of source signal and noises, respectively, and $\mathbf{\Phi}$ and $\mathbf{\Gamma}_v$ the source signal and noise coherence matrices, respectively. Note that we do *not* assume $P_s + P_v = 1$. The degree of freedoms of this model is three.

Then, the sample size normalized negative logarithm likelihood (NLL) function for the power normalized observations is given by

$$J(\hat{\mathbf{\Gamma}}_x | P_s, P_v, \phi) = \log\det \mathbf{\Gamma}_x + \mathrm{tr}(\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x) + 2\log\pi \tag{3}$$

where det and tr denote the determinant and trace of a square matrix, respectively.

The gradients of NLL cost

To derive the ML solution, we need to solve the system of equations $\partial J/\partial P_s = 0$, $\partial J/\partial P_v = 0$ and $\partial J/\partial \phi = 0$. Note that for any invertible matrix $\boldsymbol{A}$, we have

$$\mathbf{0} = d\boldsymbol{I} = d(\boldsymbol{A}\boldsymbol{A}^{-1}) = d\boldsymbol{A}\boldsymbol{A}^{-1} + \boldsymbol{A}d\boldsymbol{A}^{-1}$$

Thus,

$$d\boldsymbol{A}^{-1} = -\boldsymbol{A}^{-1}d\boldsymbol{A}\boldsymbol{A}^{-1}$$

With this fact, we can easily write down the derivative of $\mathbf{\Gamma}_x^{-1}$ with respect to P_s, P_v and ϕ. For example, we have

$$\frac{\partial \mathbf{\Gamma}_x^{-1}}{\partial \phi} = -\mathbf{\Gamma}_x^{-1}\frac{\partial \mathbf{\Gamma}_x}{\partial \phi}\mathbf{\Gamma}_x^{-1} = -\mathbf{\Gamma}_x^{-1} P_s \frac{\partial \begin{bmatrix} 1 & e^{j\phi} \\ e^{-j\phi} & 1 \end{bmatrix}}{\partial \phi}\mathbf{\Gamma}_x^{-1} = -\mathbf{\Gamma}_x^{-1} P_s \begin{bmatrix} 0 & je^{j\phi} \\ -je^{-j\phi} & 0 \end{bmatrix}\mathbf{\Gamma}_x^{-1}$$

Here, we give the gradients in their compact forms as below,

$$\begin{aligned} \frac{\partial J}{\partial P_s} &= \mathrm{tr}(\mathbf{\Gamma}_x^{-1}\mathbf{\Phi}) - \mathrm{tr}(\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x\mathbf{\Gamma}_x^{-1}\mathbf{\Phi}) \\ \frac{\partial J}{\partial P_v} &= \mathrm{tr}(\mathbf{\Gamma}_x^{-1}\mathbf{\Gamma}_v) - \mathrm{tr}(\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x\mathbf{\Gamma}_x^{-1}\mathbf{\Gamma}_v) \\ \frac{1}{P_s}\frac{\partial J}{\partial \phi} &= \mathrm{tr}(\mathbf{\Gamma}_x^{-1}\frac{d\mathbf{\Phi}}{d\phi}) - \mathrm{tr}(\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x\mathbf{\Gamma}_x^{-1}\frac{d\mathbf{\Phi}}{d\phi}) \end{aligned} \tag{4}$$

where $\frac{d\mathbf{\Phi}}{d\phi} = \begin{bmatrix} 0 & je^{j\phi} \\ -je^{-j\phi} & 0 \end{bmatrix}$.

Solving the set of equations (4)

We drop out the subscript ml to simplify our notations.

Let us rewrite the set of equations $\partial J/\partial P_s = 0$, $\partial J/\partial P_v = 0$ and $\partial J/\partial \phi = 0$ as

$$\mathrm{tr}[(\boldsymbol{I}-\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x)(\mathbf{\Gamma}_x^{-1}\mathbf{\Phi})] = 0,\ \mathrm{tr}[(\boldsymbol{I}-\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x)(\mathbf{\Gamma}_x^{-1}\mathbf{\Gamma}_v)] = 0,\ \mathrm{tr}[(\boldsymbol{I}-\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x)(\mathbf{\Gamma}_x^{-1}\frac{d\mathbf{\Phi}}{d\phi})] = 0$$

We can always rewrite $\mathrm{tr}(\boldsymbol{AB})$ as $[\mathrm{vec}(\boldsymbol{A})]^T\mathrm{vec}(\boldsymbol{B}^T)$. Thus, $\mathrm{vec}(\boldsymbol{I} - \mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x)$ can be solved from the set of equations

$$\begin{bmatrix} [\mathrm{vec}(\mathbf{\Gamma}_x^{-1}\mathbf{\Phi})]^T \\ [\mathrm{vec}(\mathbf{\Gamma}_x^{-1}\mathbf{\Gamma}_v)]^T \\ [\mathrm{vec}(\mathbf{\Gamma}_x^{-1}\frac{d\mathbf{\Phi}}{d\phi})]^T \end{bmatrix} \mathrm{vec}(\boldsymbol{I} - (\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x)^T) = \mathbf{0}$$

It is easy to verify that $\mathrm{vec}(\mathbf{\Phi})$, $\mathrm{vec}(\mathbf{\Gamma}_v)$ and $\mathrm{vec}(\frac{d\mathbf{\Phi}}{d\phi})$ are three independent vectors. So $\mathrm{vec}(\mathbf{\Gamma}_x^{-1}\mathbf{\Phi})$ and $\mathrm{vec}(\mathbf{\Gamma}_x^{-1}\mathbf{\Gamma}_v)$ and $\mathrm{vec}(\mathbf{\Gamma}_x^{-1}\frac{d\mathbf{\Phi}}{d\phi})$ are indepdent as well since $\mathbf{\Gamma}_x^{-1}$ is not singular. Thus, we have three independent equations, and the model just has a degree of freedoms of three. Hence, we must have $\boldsymbol{I} - (\mathbf{\Gamma}_x^{-1}\hat{\mathbf{\Gamma}}_x)^T = \mathbf{0}$, i.e., $\mathbf{\Gamma}_x = \hat{\mathbf{\Gamma}}_x$. Starting from $\mathbf{\Gamma}_x = \hat{\mathbf{\Gamma}}_x$, we can have

$$P_s + P_v = 1,\ P_s e^{j\phi} + P_v\gamma_v = \hat{\gamma}_x$$

Thus,

$$e^{j\phi} = \frac{\hat{\gamma}_x - P_v\gamma_v}{P_s} = \frac{\hat{\gamma}_x - P_v\gamma_v}{1 - P_v}$$

Since $|e^{j\phi}| = 1$, we have

$$\left|\frac{\hat{\gamma}_x - P_v\gamma_v}{1 - P_v}\right| = 1$$

We can solve for P_v from the above equation. There will be two solutions for P_v. But, the one greater than 1 should be discarded as $P_s = 1 - P_v \geq 0$.