Analysis of adaptive multichannel filters trained with covariance mismatched samples

Olivier Besson, ISAE-SUPAERO

submitted to 5^{th} SONDRA workshop

An ubiquitous task in radar systems consists in detecting the presence of a target with known signature $\mathbf{v}$ and/or estimating its amplitude α from a corrupted version $\mathbf{x} = \alpha\mathbf{v} + \mathbf{n}$ where $\mathbf{n}$ stands for the disturbance (clutter, thermal noise and possibly jammers). A most usual approach relies on designing an adaptive filter $\mathbf{w}$ whose output signal to noise ratio (SNR) is given by $\text{SNR}(\mathbf{w}) = P|\mathbf{w}^H\mathbf{v}|^2/(\mathbf{w}^H\boldsymbol{\Sigma}\mathbf{w})$ where $\boldsymbol{\Sigma}$ is the disturbance covariance matrix. $\text{SNR}(\mathbf{w})$ is always inferior to $\text{SNR}_{\text{opt}} = P\mathbf{v}^H\boldsymbol{\Sigma}^{-1}\mathbf{v}$ which is obtained with $\mathbf{w}_{\text{opt}} = (\mathbf{v}^H\boldsymbol{\Sigma}^{-1}\mathbf{v})^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{v}$ and thus $\ell(\mathbf{w}) = \text{SNR}(\mathbf{w})/\text{SNR}_{\text{opt}}$ is a traditional figure of merit for any adaptive filter.

When $\boldsymbol{\Sigma}$ is unknown it is common practice to use $\mathbf{w} = (\mathbf{v}^H\mathbf{S}_t^{-1}\mathbf{v})^{-1}\mathbf{S}_t^{-1}\mathbf{v}$ where $\mathbf{S}_t = \mathbf{X}_t\mathbf{X}_t^H$ and $\mathbf{X}_t$ is the $N \times K$ matrix of training samples. The corresponding *SNR loss* becomes

$$\ell = \frac{(\mathbf{v}^H\mathbf{S}_t^{-1}\mathbf{v})^2}{(\mathbf{v}^H\boldsymbol{\Sigma}^{-1}\mathbf{v})(\mathbf{v}^H\mathbf{S}_t^{-1}\boldsymbol{\Sigma}\mathbf{S}_t^{-1}\mathbf{v})}$$

When $\mathbf{X}_t \stackrel{d}{=} \mathbb{C}\mathcal{N}_{N,K}(\mathbf{0}, \boldsymbol{\Sigma}, \mathbf{I}_K)$ Reed Mallett and Brennan (RMB) showed that ℓ follows a Beta distribution.

The aim of this presentation is to examine the distribution of ℓ when there is a mismatch in the training samples. More precisely we will consider two cases. First we assume that $\mathbf{X}_t \stackrel{d}{=} \mathbb{C}\mathcal{N}_{N,K}(\mathbf{0}, \boldsymbol{\Sigma}_t, \mathbf{I}_K)$ with $\boldsymbol{\Sigma}_t \neq \boldsymbol{\Sigma}$. Next we assume that the training samples have covariance matrix $\boldsymbol{\Sigma}$ but follow a *Student distribution*. Note that it can also be viewed as a covariance mismatch, i.e., $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}^{1/2}\mathbf{W}^{-1}\boldsymbol{\Sigma}^{H/2}$ where $\mathbf{W}$ follows a Wishart distribution $\mathbf{W} \stackrel{d}{=} \mathbb{C}\mathcal{W}_N(\nu, \mu^{-1}\mathbf{I}_N)$.

Let us note $\boldsymbol{\Omega} = \mathbf{Q}^H\boldsymbol{\Sigma}_t^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}_t^{-H/2}\mathbf{Q} = \begin{pmatrix} \Omega_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} \end{pmatrix}$ where $\mathbf{Q}$ is such that $\mathbf{Q}^H\boldsymbol{\Sigma}_t^{-1/2}\mathbf{v} = (\mathbf{v}^H\boldsymbol{\Sigma}_t^{-1}\mathbf{v})^{1/2}\mathbf{e}_1$ where $\mathbf{e}_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T$. Then, *when* $\mathbf{X}_t \stackrel{d}{=} \mathbb{C}\mathcal{N}_{N,K}(\mathbf{0}, \boldsymbol{\Sigma}_t, \mathbf{I}_K)$, *for any* $\boldsymbol{\Sigma}$ *and* $\boldsymbol{\Sigma}_t$ *we have*

$$\boxed{\ell \stackrel{d}{=} \left[1 + \frac{\mathbf{v}^H\boldsymbol{\Sigma}^{-1}\mathbf{v}}{\mathbf{v}^H\boldsymbol{\Sigma}_t^{-1}\mathbf{v}} \frac{\sum_{i=1}^{N-1}\lambda_i\mathbb{C}\chi_1^2(V_{21}\delta_i)}{V_{21}}\right]^{-1}} \tag{1}$$

where $V_{21} \stackrel{d}{=} \mathbb{C}\chi^2_{K-N+2}(0)$, $\delta_i = |\mathbf{u}_i^H\boldsymbol{\Omega}_{22}^{-1}\boldsymbol{\Omega}_{21}|^2$ and $\lambda_i, \mathbf{u}_i$ are the eigenvalues and eigenvectors of $\boldsymbol{\Omega}_{22} = (\mathbf{V}_\perp^H\boldsymbol{\Sigma}_t\mathbf{V}_\perp)^{-1/2}(\mathbf{V}_\perp^H\boldsymbol{\Sigma}\mathbf{V}_\perp)(\mathbf{V}_\perp^H\boldsymbol{\Sigma}_t\mathbf{V}_\perp)^{H/2}$.

The very general expression (1) allows to cover a number of special cases of interest. First, when $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_t$, $\boldsymbol{\Omega} = \mathbf{I}_N$ and (1) reduces to the RMB formula:

$$\ell_{\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}} \stackrel{d}{=} \left[1 + \frac{\mathbb{C}\chi^2_{N-1}(0)}{\mathbb{C}\chi^2_{K-N+2}(0)}\right]^{-1}$$

Next, when training samples contain the signal of interest, i.e., $\boldsymbol{\Sigma}_t = \gamma\boldsymbol{\Sigma} + P\mathbf{v}\mathbf{v}^H$ (1) becomes

$$\ell_{\boldsymbol{\Sigma}_t = \gamma\boldsymbol{\Sigma} + P\mathbf{v}\mathbf{v}^H} \stackrel{d}{=} \left[1 + (1 + \gamma^{-1}\text{SNR}_{\text{opt}})\frac{\mathbb{C}\chi^2_{N-1}(0)}{\mathbb{C}\chi^2_{K-N+2}(0)}\right]^{-1}$$

When $\boldsymbol{\Sigma}_t \neq \boldsymbol{\Sigma}$, if the so-called *generalized eigen-relation* (GER) -which states that $\boldsymbol{\Sigma}_t^{-1}\mathbf{v} = \lambda\boldsymbol{\Sigma}^{-1}\mathbf{v}$- holds then

$$\ell_{\text{GER}} \stackrel{d}{=} \left[1 + \frac{\sum_{i=1}^{N-1}\lambda_i\mathbb{C}\chi_1^2(0)}{\lambda\mathbb{C}\chi^2_{K-N+2}(0)}\right]^{-1}$$

A special case of the GER is when $\mathbf{\Sigma} = \mathbf{\Sigma}_t + \mathbf{q}\mathbf{q}^H$ (surprise interference) and $\mathbf{q}^H \mathbf{\Sigma}_t^{-1}\mathbf{v} = 0$, where ℓ_{GER} can be simplified to

$$\ell_{\mathbf{\Sigma}=\mathbf{\Sigma}_t+\mathbf{q}\mathbf{q}^H,\, \mathbf{q}^H\mathbf{\Sigma}_t^{-1}\mathbf{v}=0} \overset{d}{=} \left[1 + \frac{\mathbb{C}\chi^2_{N-2}(0) + (1+\mathbf{q}^H\mathbf{\Sigma}_t^{-1}\mathbf{q})\mathbb{C}\chi^2_1(0)}{\mathbb{C}\chi^2_{K-N+2}(0)}\right]^{-1}$$

Finally we show that the distribution of ℓ in (1) can be accurately approximated by

$$\ell \overset{d}{\simeq} \left[1 + a\frac{\mathbb{C}\chi^2_p(0)}{\mathbb{C}\chi^2_q(0)}\right]^{-1}$$

In the Student case where $p(\mathbf{X}_t) \propto |\mu\mathbf{\Sigma}|^{-K}|\mathbf{I}_N + (\mu\mathbf{\Sigma})^{-1}\mathbf{X}_t\mathbf{X}_t^H|^{-(\nu+K)}$ we show that

$$\ell_{\text{Student},\mathbf{\Sigma}_t=\mathbf{\Sigma}} \overset{d}{=} \left[1 + \left(1 + \frac{\mathbb{C}\chi^2_{K-N+1}(0)}{\mathbb{C}\chi^2_\nu(0)}\right)\frac{\mathbb{C}\chi^2_{N-1}(0)}{\mathbb{C}\chi^2_{K-N+2}(0)}\right]^{-1}$$

where the term in blue concentrates the difference compared to the Gaussian case. Numerical illustrations of the impact of these mismatches (either covariance or distribution) on the SNR loss will be given.

In a second part we examine the impact of a covariance mismatch $\mathbf{\Sigma}_t \neq \mathbf{\Sigma}$ on classical adaptive detectors such as Kelly's GLRT or the AMF. Both depend on two statistics $\beta = (1 + s_1 - s_2)^{-1}$ and $\tilde{t} = s_2/(1 + s_1 - s_2)$ where $s_1 = \mathbf{x}^H\mathbf{S}_t^{-1}\mathbf{x}$ and $s_2 = |\mathbf{x}^H\mathbf{S}_t^{-1}\mathbf{v}|^2/(\mathbf{v}^H\mathbf{S}_t^{-1}\mathbf{v})$. For instance Kellys's GLRT is $1+\tilde{t}$ and the AMF is $\tilde{t}/\beta$. The distributions of β and $\tilde{t}$ for any $\mathbf{\Sigma}_t$ are given in the following table:

$\mathbf{\Sigma}_t = \mathbf{\Sigma}$	$\mathbf{\Sigma}_t \neq \mathbf{\Sigma}$
$\beta \overset{d}{=} \left[1 + \frac{\mathbb{C}\chi^2_{N-1}(0)}{\mathbb{C}\chi^2_{K-N+2}(0)}\right]^{-1}$	$\beta \overset{d}{=} (1 + \tilde{\mathbf{x}}_2\mathbf{W}_{22}^{-1}\tilde{\mathbf{x}}_2)^{-1} \overset{d}{=} \left[1 + \frac{\sum_{i=1}^{N-1}\lambda_i\mathbb{C}\chi^2_1(0)}{\mathbb{C}\chi^2_{K-N+2}(0)}\right]^{-1}$
$\tilde{t}\vert\beta \overset{d}{=} \mathbb{C}\mathcal{F}_{1,K-N+1}(\beta\vert\alpha\vert^2\mathbf{v}^H\mathbf{\Sigma}^{-1}\mathbf{v},)$	$\tilde{t}\vert\tilde{\mathbf{x}}_2, \mathbf{W}_{22} \overset{d}{=} \left[1 + \beta\left(\frac{\mathbf{v}^H\mathbf{\Sigma}_t^{-1}\mathbf{v}}{\mathbf{v}^H\mathbf{\Sigma}^{-1}\mathbf{v}} - 1\right)\right]$ $\times\mathbb{C}\mathcal{F}_{1,K-N+1}\left(\beta\frac{\Vert\alpha\vert(\mathbf{v}^H\mathbf{\Sigma}_t^{-1}\mathbf{v})^{1/2}+\mathbf{\Omega}_{21}\mathbf{\Omega}_{11}^{-1}\tilde{\mathbf{x}}_2\vert^2}{1+\beta\left(\frac{\mathbf{v}^H\mathbf{\Sigma}_t^{-1}\mathbf{v}}{\mathbf{v}^H\mathbf{\Sigma}^{-1}\mathbf{v}}-1\right)},\right)$

where $\mathbf{W}_{22} \overset{d}{=} \mathbb{C}\mathcal{W}_{N-1}(K, \mathbf{I}_{N-1})$ is independent of $\tilde{\mathbf{x}}_2 \overset{d}{=} \mathbb{C}\mathcal{N}_{N-1}(\mathbf{0}, \mathbf{\Omega}_{22})$. These expressions allow assessment of most adaptive detectors under covariance mismatch. For instance, Figure 1 shows the actual P_{fa} in case of $\mathbf{\Sigma}_t \neq \mathbf{\Sigma}$ when the nominal $\bar{P}_{fa} = 10^{-3}$.

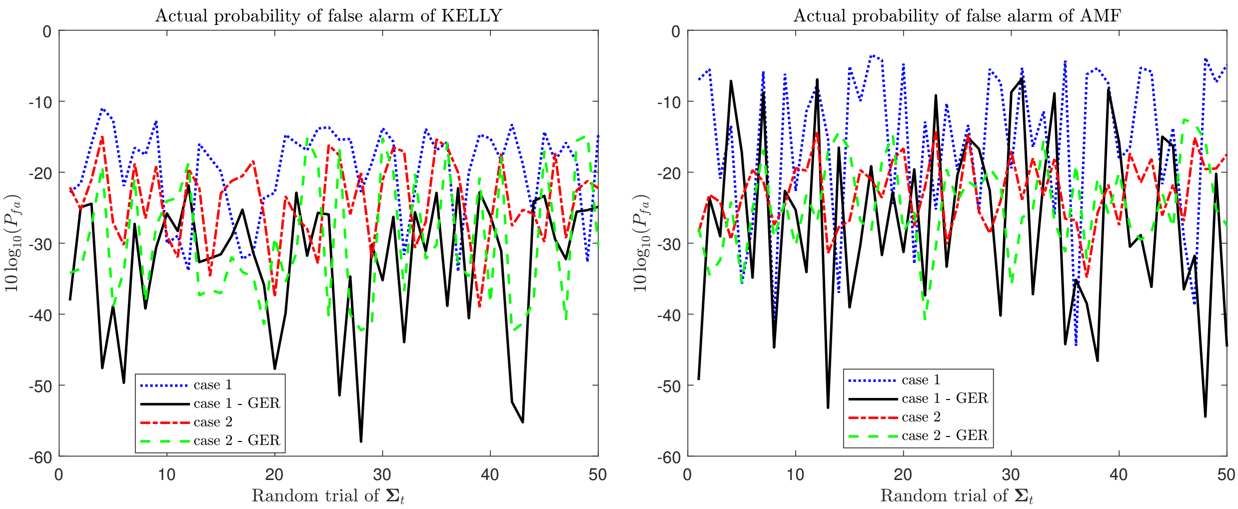

Figure 1: Actual probability of false alarm of Kelly and AMF detectors under covariance mismatch. Case 1: $\mathbf{\Sigma}_t = \mathbf{\Sigma}^{1/2}\mathbf{W}^{-1}\mathbf{\Sigma}^{H/2}$ with $\mathbf{W} \overset{d}{=} \mathbb{C}\mathcal{W}_N(\nu, \mu^{-1}\mathbf{I}_N)$. Case 2: $\mathbf{\Sigma} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, $\mathbf{\Sigma}_t = \mathbf{U}\mathbf{\Lambda}^{1/2}\text{diag}(\gamma_n)\mathbf{\Lambda}^{1/2}\mathbf{U}^H$. $P_{fa}(\mathbf{\Sigma}_t = \mathbf{\Sigma}) = 10^{-3}$.