1 Preliminaries

We use the model

$$\boldsymbol{y}(t) = \boldsymbol{A}(\boldsymbol{\omega})\boldsymbol{x}(t) + \boldsymbol{e}(t), \quad t = 1, 2, \ldots, N,$$

where $\boldsymbol{A}(\boldsymbol{\omega}) = [\boldsymbol{a}(\omega_1)\ \boldsymbol{a}(\omega_2)\ \ldots\ \boldsymbol{a}(\omega_n)]$ has full column-rank, $\omega_i \in \mathbb{R}$, $\boldsymbol{y}(t) \in \mathbb{C}^m$, $\boldsymbol{x}(t) \in \mathbb{C}^n$ is a zero-mean random vector with non-singular $\boldsymbol{P} \triangleq \mathbb{E}\left\{\boldsymbol{x}(t)\boldsymbol{x}^H(t)\right\}$, and $\boldsymbol{e}(t)$ is zero-mean additive noise with $\mathbb{E}\left\{\boldsymbol{e}(t)\boldsymbol{e}^H(t)\right\} = \sigma\boldsymbol{I}$, $\mathbb{E}\left\{\boldsymbol{e}(t)\boldsymbol{e}^T(t)\right\} = \boldsymbol{0}$. We also assume that the number of emitters n is known, even though it could be estimated using the MUSIC technique that follows.

2 MUSIC (MUltiple SIgnal Characterization)

The MUSIC method uses the 2:nd order statistics of the measurements $\boldsymbol{y}(t)$. In the ideal case, it decomposes the Hermitian matrix

$$\boldsymbol{R} \triangleq \mathbb{E}\left\{\boldsymbol{y}(t)\boldsymbol{y}^H(t)\right\} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma\boldsymbol{I} \tag{1}$$

using the eig. val. decomposition as follows,

$$\boldsymbol{R} = \underbrace{\begin{bmatrix}\boldsymbol{S} & | & \boldsymbol{G}\end{bmatrix}}_{\text{unitary}} \underbrace{\left[\begin{array}{ccc|ccc} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_n & & & \\ \hline & & & \sigma & & \\ & & & & \ddots & \\ & & & & & \sigma \end{array}\right]}_{\boldsymbol{\Lambda}} \begin{bmatrix}\boldsymbol{S}^H \\ -- \\ \boldsymbol{G}^H\end{bmatrix}. \tag{2}$$

Note that the dependency in $\boldsymbol{A}$ of $\boldsymbol{\omega}$ is assumed to be obvious and therefore is not stated explicitly. From eq. (1) and (2), we can conclude that

$$\begin{aligned} &\boldsymbol{R}\boldsymbol{G} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H\boldsymbol{G} + \sigma\boldsymbol{G} = \begin{bmatrix}\boldsymbol{S} & \boldsymbol{G}\end{bmatrix}\boldsymbol{\Lambda}\begin{bmatrix}\boldsymbol{S}^H \\ \boldsymbol{G}^H\end{bmatrix}\boldsymbol{G} = \sigma\boldsymbol{G} \\ &\Longrightarrow \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H\boldsymbol{G} = \boldsymbol{0} \Longrightarrow \underbrace{\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{P}}_{\text{invertible}}\boldsymbol{A}^H\boldsymbol{G} = \boldsymbol{0} \Longrightarrow \boldsymbol{A}^H\boldsymbol{G} = \boldsymbol{0}. \end{aligned} \tag{3}$$

This property is utilized by the MUSIC method to find the parameters $\boldsymbol{\omega}$ of $\boldsymbol{A}(\boldsymbol{\omega})$. More formally, MUSIC picks n values $\hat{\omega}_1, \hat{\omega}_2, \ldots, \hat{\omega}_n$ that minimize the squared 2-norm of $\boldsymbol{G}^H\boldsymbol{a}(\omega)$. Hence,

$$\hat{\omega}_i = \underset{\omega}{\operatorname{argmin}}\, f(\omega), \quad f \triangleq \boldsymbol{a}(\omega)^H\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{a}(\omega). \tag{4}$$

So far, we have used the true $\boldsymbol{R}$ matrix, which is not available in practice. Therefore, in practice, the MUSIC method performs the procedure above using the estimate $\hat{\boldsymbol{R}} \triangleq \frac{1}{N}\sum_{t=1}^{N} \boldsymbol{y}(t)\boldsymbol{y}^H(t)$ of $\boldsymbol{R}$ instead. The hat-notation $\hat{(\cdot)}$ is introduced to represent approximations of the respective quantities above. Note that the quantities with the hat-notation are stochastic since they all are functions of $\boldsymbol{y}(t)$, which follows from the computation of the estimate $\hat{\boldsymbol{R}}$ of $\boldsymbol{R}$.

3 Asymptotic Distribution of MUSIC

The results below give a detailed step-by-step derivation of the proof in ref. [1] App. B where the asymptotic distribution of the error $\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}$ is derived. First, let us derive some useful relations for large N

$$\hat{\boldsymbol{S}} \overset{\text{Lem.3.1a}}{=} \boldsymbol{S} + \mathcal{O}(1/\sqrt{N}), \quad \hat{\boldsymbol{G}} \neq \boldsymbol{G} + \mathcal{O}(1/\sqrt{N}), \tag{5}$$

$$\hat{\boldsymbol{w}} = \boldsymbol{w} + \mathcal{O}(1/\sqrt{N}) \tag{6}$$

$$\boldsymbol{I} = \underbrace{\begin{bmatrix}\boldsymbol{S} & \boldsymbol{G}\end{bmatrix}}_{\text{unitary}} \begin{bmatrix}\boldsymbol{S}^H \\ \boldsymbol{G}^H\end{bmatrix} = \boldsymbol{S}\boldsymbol{S}^H + \boldsymbol{G}\boldsymbol{G}^H, \quad \boldsymbol{I} = \hat{\boldsymbol{S}}\hat{\boldsymbol{S}}^H + \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H, \tag{7}$$

$$\hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \overset{(7)}{=} \boldsymbol{I} - \hat{\boldsymbol{S}}\hat{\boldsymbol{S}}^H \overset{(5)}{=} \boldsymbol{I} - \boldsymbol{S}\boldsymbol{S}^H + \mathcal{O}(1/\sqrt{N}) \overset{(7)}{=} \boldsymbol{G}\boldsymbol{G}^H + \mathcal{O}(1/\sqrt{N}), \tag{8}$$

$$\boldsymbol{A}^H = \boldsymbol{A}^H\boldsymbol{I} \overset{(7)}{=} \boldsymbol{A}^H(\boldsymbol{S}\boldsymbol{S}^H + \boldsymbol{G}\boldsymbol{G}^H) \overset{(3)}{=} \boldsymbol{A}^H\boldsymbol{S}\boldsymbol{S}^H, \tag{9}$$

$$\boldsymbol{G}^H\hat{\boldsymbol{S}} \overset{(5)}{=} \boldsymbol{G}^H(\boldsymbol{S} + \mathcal{O}(1/\sqrt{N})) = \mathcal{O}(1/\sqrt{N}), \tag{10}$$

$$\begin{aligned}\boldsymbol{S}^H\hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H\boldsymbol{G} &\overset{(7)}{=} \boldsymbol{S}^H(\boldsymbol{I} - \hat{\boldsymbol{S}}\hat{\boldsymbol{S}}^H)\boldsymbol{G} = -\boldsymbol{S}^H\hat{\boldsymbol{S}}\hat{\boldsymbol{S}}^H\boldsymbol{G} \\ &\overset{(5),(10)}{=} -\hat{\boldsymbol{S}}^H\boldsymbol{G} + \mathcal{O}(1/N) \overset{(10)}{=} \mathcal{O}(1/\sqrt{N}),\end{aligned} \tag{11}$$

$$\begin{aligned}\boldsymbol{G}^H\hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H\boldsymbol{G} &\overset{(7)}{=} \boldsymbol{G}^H(\boldsymbol{I} - \hat{\boldsymbol{S}}\hat{\boldsymbol{S}}^H)\boldsymbol{G} = \boldsymbol{I} - \boldsymbol{G}^H\hat{\boldsymbol{S}}\hat{\boldsymbol{S}}^H\boldsymbol{G} \\ &\overset{(10)}{=} \boldsymbol{I} + \mathcal{O}(1/N) \Longrightarrow \text{asympt. unitary } \hat{\boldsymbol{G}}^H\boldsymbol{G},\end{aligned} \tag{12}$$

$$\boldsymbol{S}^H\hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H\boldsymbol{G} \overset{(11)}{=} \mathcal{O}(1/\sqrt{N}) \overset{(12)}{\Longrightarrow} \boldsymbol{S}^H\hat{\boldsymbol{G}} = \mathcal{O}(1/\sqrt{N}), \tag{13}$$

$$\hat{\boldsymbol{G}}^H = \hat{\boldsymbol{G}}^H\boldsymbol{I} \overset{(7)}{=} \hat{\boldsymbol{G}}^H(\boldsymbol{S}\boldsymbol{S}^H + \boldsymbol{G}\boldsymbol{G}^H) \overset{(13)}{=} \hat{\boldsymbol{G}}^H\boldsymbol{G}\boldsymbol{G}^H + \mathcal{O}(1/\sqrt{N}). \tag{14}$$

Now, let us continue with the proof in App. B. Since $\hat{\omega}_i = \operatorname{argmin}_\omega \hat{f}(\omega)$ where $\hat{f}(\omega) \triangleq \boldsymbol{a}(\omega)^H\hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H\boldsymbol{a}(\omega)$, we can conclude that $\hat{f}'(\hat{\omega}_i) = 0$. From Taylor's theorem and the relation in (6), we can write the following

$$\begin{aligned}
0 &= \hat{f}'(\hat{\omega}_i) = \hat{f}'(\omega_i) + \hat{f}''(\omega_i)(\hat{\omega}_i - \omega_i) + \mathcal{O}(1/N) \\
&= 2\mathrm{Re}\left\{\boldsymbol{a}(\omega)^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \boldsymbol{d}(\omega)\right\} \\
&\quad + 2\mathrm{Re}\left\{\boldsymbol{d}(\omega)^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \boldsymbol{d}(\omega) + \boldsymbol{a}(\omega)^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \boldsymbol{d}'(\omega)\right\}(\hat{\omega}_i - \omega_i) + \mathcal{O}(1/N) \\
&\overset{(8),(6)}{=} 2\mathrm{Re}\left\{\boldsymbol{a}(\omega)^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \boldsymbol{d}(\omega)\right\} \\
&\quad + 2\mathrm{Re}\left\{\boldsymbol{d}(\omega)^H \boldsymbol{G}\boldsymbol{G}^H \boldsymbol{d}(\omega) + \boldsymbol{a}(\omega)^H \boldsymbol{G}\boldsymbol{G}^H \boldsymbol{d}'(\omega)\right\}(\hat{\omega}_i - \omega_i) + \mathcal{O}(1/N) \\
&= 2\mathrm{Re}\left\{\boldsymbol{a}(\omega)^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \boldsymbol{d}(\omega)\right\} + 2\mathrm{Re}\left\{\boldsymbol{d}(\omega)^H \boldsymbol{G}\boldsymbol{G}^H \boldsymbol{d}(\omega)\right\}(\hat{\omega}_i - \omega_i) + \mathcal{O}(1/N) \\
&\Longleftrightarrow \hat{\omega}_i - \omega_i = \frac{\mathrm{Re}\left\{\boldsymbol{a}(\omega)^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \boldsymbol{d}(\omega)\right\}}{\mathrm{Re}\left\{\boldsymbol{d}(\omega)^H \boldsymbol{G}\boldsymbol{G}^H \boldsymbol{d}(\omega)\right\}} + \mathcal{O}(1/N),
\end{aligned} \tag{15}$$

where $\boldsymbol{d}(\omega) \triangleq \boldsymbol{a}'(\omega)$. The only stochastic part in the expression above is $\boldsymbol{a}^H(\omega)\hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H$. Hence,

$$\begin{aligned}
\boldsymbol{A}^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H &\overset{(9)}{=} \boldsymbol{A}^H \boldsymbol{S}\boldsymbol{S}^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \overset{(14)}{=} \boldsymbol{A}^H \boldsymbol{S}\boldsymbol{S}^H \hat{\boldsymbol{G}}(\hat{\boldsymbol{G}}^H \boldsymbol{G}\boldsymbol{G}^H + \mathcal{O}(1/\sqrt{N})) \\
&\overset{(13)}{=} \boldsymbol{A}^H \boldsymbol{S}\boldsymbol{S}^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \boldsymbol{G}\boldsymbol{G}^H + \mathcal{O}(1/N) \overset{(11)}{=} -\boldsymbol{A}^H \boldsymbol{S}\hat{\boldsymbol{S}}^H \boldsymbol{G}\boldsymbol{G}^H + \mathcal{O}(1/N)
\end{aligned}$$

where according to App. A in [1], $\hat{\boldsymbol{S}}^H \boldsymbol{G}$ has the same asymptotic distribution as $-\boldsymbol{S}^H \hat{\boldsymbol{G}}$. Therefore, $\boldsymbol{A}^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H$ has the same asymptotic distribution as $\boldsymbol{A}^H \boldsymbol{S}\boldsymbol{S}^H \hat{\boldsymbol{G}}\boldsymbol{G}^H$, where the only stochastic part $\boldsymbol{S}\boldsymbol{S}^H \hat{\boldsymbol{G}}$ has a Gaussian distribution according to Lemma 3.1b) in [1].

References

[1] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood, and Cramer-Rao Bound," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720-741, May 1989.