

HIGH RESOLUTION DIRECTION FINDING : FROM PERFORMANCE TOWARD ANTENNA ARRAY OPTIMIZATION - THE MONO-SOURCE CASE

Anne Ferréol¹ and Pascal Chevalier²

THALES Communications

160 boulevard de Valmy, 92700, Colombes, France

phone¹: + (33)1 46 13 23 29 , fax^{1,2}: + (33)1 46 13 25 55, email¹: anne.ferreol@fr.thalesgroup.com

phone²: + (33)1 46 13 26 98 , email²: pascal.chevalier@fr.thalesgroup.com

ABSTRACT

Many direction finding methods have been developed these last decades and the performance of some of the latter have been derived with or without modelling errors, jointly with some Cramer Rao bounds (CRB) on the estimated Direction of Arrival (DOAs). However, despite of these works, the link between the array geometry and these performances, or CRB, has been scarcely analyzed, which prevents from optimizing the array design to obtain a specified level of performance. We consider in this paper the general 2D DOA estimation problem from 3D, 2D or 1D arrays and we limit the analysis to the single source case. In this context, the first purpose of this paper is to show that both the deterministic and the stochastic CRB, jointly with the variance of DOAs obtained from the MUSIC algorithm, with and without modelling errors, are proportional to the same term which depends on the sensors location. This term analysis allows to develop the second purpose of the paper, i.e. the first tools for an array design methodology for the performance optimization of DOA estimation methods.

1. INTRODUCTION

Direction finding or DOA estimation algorithms finds applications in many fields such as radar, sonar or spectrum monitoring. These last decades, many High Resolution (HR) direction finding methods, such as MUSIC [6], have been developed to mitigate the limitations of conventional methods in multiple sources contexts. Performance of some of these methods have been computed analytically with or without modelling errors, showing off the weak robustness of these methods to both finite sample effects [5], [7] and modelling errors [3], [8], [2]. These performance analyses have been completed by the computation of both deterministic [1] and stochastic [7], CRB on the DOAs estimates. However, despite of these works, the link between the array geometry and these performance criteria, or CRB, has been scarcely analyzed, which prevents from optimizing the array design to obtain a specified level of performance. To our knowledge, one of the very first studies about this link for the general 2D DOA estimation problem, has been presented very recently in [4] for planar arrays (2D arrays) and from the stochastic CRB computation in the single source case. However one may wonder if the same link holds true for other performance criteria or in the presence of modelling errors. The first purpose of this paper is to get more insights into this link by considering, on one hand, both 3D, 2D and 1D arrays and, on the other hand, both CRB and MUSIC performance criteria. More precisely, we show in this paper that both the deterministic and the stochastic CRB, jointly with the variance of DOAs estimates obtained from the MUSIC algorithm, with and without modelling errors, are proportional to the same term which depends on the sensors location. This result is completely new for 2D DOA estimation problems. The analysis of this term allows to develop the second

purpose of the paper, i.e. the first tools for an array design methodology for the performance optimization of direction finding methods from 3D, 2D or 1D arrays, which will be presented elsewhere.

2. OBSERVATION MODEL AND PROBLEM FORMULATION

We consider an array of N narrow-band sensors and we denote by $\mathbf{x}(t)$ the vector of the complex envelopes of the signals at the output of sensors. The array is assumed to receive a single source with additive noise. Under these assumptions, the vector $\mathbf{x}(t)$ can be written as

$$\mathbf{x}(t) = \tilde{\mathbf{a}}(\Theta_0) s(t) + \mathbf{n}(t) \quad (1)$$

where $\mathbf{n}(t)$ is the additive noise vector, which is assumed to be spatially white, $\Theta_0 = [\theta_0 \Delta_0]^T$, $(\cdot)^T$ denotes transpose, θ_0 , Δ_0 and $s(t)$ are the azimuth, the elevation and the complex envelope of the source respectively and $\tilde{\mathbf{a}}(\Theta_0)$ is the observed steering vector of a signal source in the direction Θ_0 . In the presence of modelling errors, $\tilde{\mathbf{a}}(\Theta_0)$ is linked to the normalized theoretical steering vector $\mathbf{a}(\Theta_0)$ by the relation

$$\tilde{\mathbf{a}}(\Theta_0) = \mathbf{a}(\Theta_0) + \mathbf{e}. \quad (2)$$

where \mathbf{e} is the modelling error vector, $\mathbf{a}(\Theta_0)^H \mathbf{a}(\Theta_0) = N$ where $(\cdot)^H$ denotes transpose and conjugate. This model is valid for both homogeneous and heterogeneous arrays. According to Figure-1 and for arrays with identical sensors, the n^{th} component, $a_n(\Theta)$, of the normalized theoretical steering vector $\mathbf{a}(\Theta)$ is

$$a_n(\Theta) = \exp\left(j \frac{2\pi}{\lambda} \mathbf{k}(\Theta)^T \mathbf{p}_n\right) \quad (3)$$

where λ is the wavelength, $\mathbf{p}_n = [x_n \ y_n \ z_n]^T$ is the sensor coordinates vector and

$$\mathbf{k}(\Theta) = [u \ v \ w]^T \quad (4)$$

is the normalized wave vector of DOA Θ such that $u = \cos(\theta) \cos(\Delta)$, $v = \sin(\theta) \cos(\Delta)$ and $w = \sin(\Delta) = \pm\sqrt{1 - u^2 - v^2}$. The array is said to be qD , $1 \leq q \leq 3$, if $\mathbf{p}_n \in \mathbb{R}^q$. In particular, for a 2D array, $z_n = 0$, whereas for a 1D array, $y_n = z_n = 0$. Moreover, whatever the kind of array, the DOA of the source may be characterized, either by the (2×1) DOA vector Θ or the (3×1) normalized wave vector $\mathbf{k}(\Theta)$ or the (2×1) vector $\mathbf{k}_{2D}(\Theta) = [u \ v]^T$. To unify the following developments, whatever the choice of the DOA representation, we denote by Ψ_0 the $(l \times 1)$ vector which characterizes the DOA of the source, where $l = 2$ or 3.

For a given DOA estimation algorithm, we denote by $\hat{\Psi}_0$ the estimate of Ψ_0 , obtained from the K observation

vectors $[\mathbf{x}(t_1), \dots, \mathbf{x}(t_K)]$ and we denote by $\Delta\boldsymbol{\Psi}_0 = \hat{\boldsymbol{\Psi}}_0 - \boldsymbol{\Psi}_0$ the estimation error. The random vector $\Delta\boldsymbol{\Psi}_0$ is non zero due to both the finite sample effect ($K < +\infty$) and the presence of modelling error ($\mathbf{e} \neq \mathbf{0}$), which are both considered in this paper. The accuracy of the DOA estimate is characterized by the Mean Square Error (MSE) matrix $\mathbf{MS}_{\boldsymbol{\Psi}_0} = \mathbb{E}[\Delta\boldsymbol{\Psi}_0(\Delta\boldsymbol{\Psi}_0)^T]$ where $\mathbb{E}[\cdot]$ denotes the expectation operation. This matrix is derived in section 4 for the MUSIC algorithm, with and without modelling errors and is lower bounded, for an unbiased estimate $\hat{\boldsymbol{\Psi}}_0$ of $\boldsymbol{\Psi}_0$ in absence of modelling error ($\mathbf{e} = 0$), by the CRB derived in section 3.

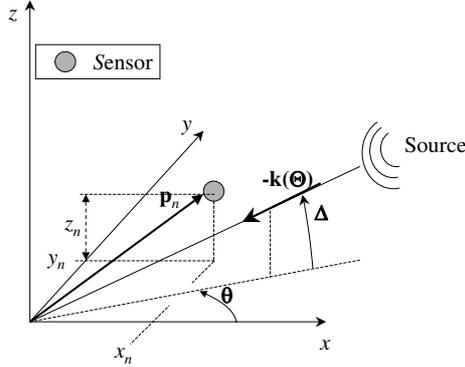


Figure 1: Source wave-vector $\mathbf{k}(\Theta)$ of direction Θ impinging the n -th sensor located at $\mathbf{p}_n = [x_n \ y_n \ z_n]^T$.

3. STOCHASTIC AND DETERMINISTIC CRAMER RAO BOUNDS

We denote by $\boldsymbol{\Omega}_0 = [\boldsymbol{\Psi}_0^T \ \boldsymbol{\beta}_0^T]^T$ the vector of the unknown parameters in the observations, where the vector $\boldsymbol{\beta}_0$ depends on the assumptions which are made on both the source and the noise (see the following sub-sections for the stochastic and deterministic models). Then, if $\hat{\boldsymbol{\Omega}}_0$ is an unbiased estimate of $\boldsymbol{\Omega}_0$ obtained from the observation vectors $[\mathbf{x}(t_1) \ \dots \ \mathbf{x}(t_K)]$, the covariance matrix $\mathbf{C}_{\boldsymbol{\Omega}_0} = \mathbb{E}[(\mathbb{E}[\hat{\boldsymbol{\Omega}}_0] - \boldsymbol{\Omega}_0)(\mathbb{E}[\hat{\boldsymbol{\Omega}}_0] - \boldsymbol{\Omega}_0)^T]$ also corresponds to the MSE matrix $\mathbf{MS}_{\boldsymbol{\Omega}_0}$ of the vector $\boldsymbol{\Omega}_0$. Under these assumptions, in the absence of modelling errors ($\mathbf{e} = 0$), $\mathbf{C}_{\boldsymbol{\Omega}_0} = \mathbf{MS}_{\boldsymbol{\Omega}_0}$ is lower-bounded by the CRB of the vector $\boldsymbol{\Omega}_0$, denoted by $\mathbf{CRB}(\boldsymbol{\Omega}_0)$ as

$$\mathbf{C}_{\boldsymbol{\Omega}_0} \geq \mathbf{CRB}(\boldsymbol{\Omega}_0) = \mathbf{FIM}(\boldsymbol{\Omega}_0)^{-1} \quad (5)$$

where $\mathbf{FIM}(\boldsymbol{\Omega}_0)$ is the so-called Fisher Information Matrix of the vector $\boldsymbol{\Omega}_0$, whose ij^{th} element is

$$\mathbf{FIM}(\boldsymbol{\Omega}_0)(i, j) = -\mathbb{E} \left[\frac{\partial^2 J_{ML}(\boldsymbol{\Omega}_0)}{\partial \boldsymbol{\Omega}_0(i) \partial \boldsymbol{\Omega}_0(j)} \right] \quad (6)$$

where ∂^i denotes the partial i^{th} derivative, $\boldsymbol{\Omega}_0(i)$ is the i^{th} component of $\boldsymbol{\Omega}_0$ and $J_{ML}(\boldsymbol{\Omega}_0)$ is the following Maximum-Likelihood (ML) criterion

$$J_{ML}(\boldsymbol{\Omega}) = \log(p(\mathbf{x}(t_1), \dots, \mathbf{x}(t_K) | \boldsymbol{\Omega})) \quad (7)$$

where $p(\mathbf{x}(t_1), \dots, \mathbf{x}(t_K) | \boldsymbol{\Omega})$ is the joint probability density function of $[\mathbf{x}(t_1), \dots, \mathbf{x}(t_K)]$ conditionally to the vector $\boldsymbol{\Omega}_0$. We then deduce from the previous results that

$\mathbf{C}_{\boldsymbol{\Psi}_0} = \mathbf{MS}_{\boldsymbol{\Psi}_0}$ is lower-bounded by the CRB of the vector $\boldsymbol{\Psi}_0$, denoted by $\mathbf{CRB}(\boldsymbol{\Psi}_0)$ and obtain

$$\mathbf{C}_{\boldsymbol{\Psi}_0} \geq \mathbf{CRB}(\boldsymbol{\Psi}_0)$$

where $\mathbf{CRB}(\boldsymbol{\Psi}_0)$ is the $(l \times l)$ submatrix of $\mathbf{CRB}(\boldsymbol{\Omega}_0)$ associated with the sub-vector $\boldsymbol{\Psi}_0$ of $\boldsymbol{\Omega}_0$. In particular, denoting by $\hat{\boldsymbol{\Psi}}_0(k)$, an unbiased estimate of the k^{th} component $\boldsymbol{\Psi}_0(k)$ of $\boldsymbol{\Psi}_0$, the Root Mean Square (RMS) error of $\hat{\boldsymbol{\Psi}}_0(k)$ is

$$RMS(\hat{\boldsymbol{\Psi}}_0(k)) = \sqrt{\mathbb{E}[\Delta\boldsymbol{\Psi}_0(k)^2]} \geq \mathbf{CRB}(\boldsymbol{\Psi}_0)(k, k) \quad (8)$$

where $\Delta\boldsymbol{\Psi}_0(k) = \hat{\boldsymbol{\Psi}}_0(k) - \boldsymbol{\Psi}_0(k)$ is the k^{th} component of the error vector $\Delta\boldsymbol{\Psi}_0$ and where $\mathbf{CRB}(\boldsymbol{\Psi}_0)(k, k)$ is the k^{th} diagonal element of $\mathbf{CRB}(\boldsymbol{\Psi}_0)$.

3.1 Stochastic Cramer Rao Bound

We talk about Stochastic CRB when the signal $s(t)$ and the noise vector $\mathbf{n}(t)$ are assumed to be Gaussian and zero mean with second order statistics defined by $\mathbb{E}[|s(t)|^2] = \gamma$ and $\mathbb{E}[\mathbf{n}(t)\mathbf{n}(t)^H] = \mathbf{R}_{nn}$. Under these assumptions, the unknown vector $\boldsymbol{\Omega}_0$ which has to be estimated is $\boldsymbol{\Omega}_0 = [\boldsymbol{\Psi}_0^T \ \gamma \ \text{vec}(\mathbf{R}_{nn})]^T$ with $\text{vec}([\mathbf{r}_1 \ \dots \ \mathbf{r}_N]) = [\mathbf{r}_1^T \ \dots \ \mathbf{r}_N^T]^T$. In [7][9] the Fisher Information matrix for the parameter vector $\boldsymbol{\Omega}_0$ is derived under these assumptions.

Assuming a spatially white noise, i.e $\mathbf{R}_{nn} = \sigma^2 \mathbf{I}_N$ where \mathbf{I}_N is the $N \times N$ identity matrix and σ^2 a scalar parameter corresponding of the mean power of the noise per sensor, the parameter $\boldsymbol{\Omega}_0$ is reduced to $\boldsymbol{\Omega}_0 = [\boldsymbol{\Psi}_0^T \ \gamma \ \sigma^2]^T$. In this context, it can be shown, after tedious derivations, that the stochastic CRB on $\boldsymbol{\Psi}_0$ denoted by $\mathbf{CRB}_{stoc}(\boldsymbol{\Psi}_0)$ is

$$\mathbf{CRB}_{stoc}(\boldsymbol{\Psi}_0) = \frac{1 + \frac{\sigma^2}{\gamma N}}{K \left(\frac{\gamma}{\sigma^2} \right)} \mathbf{H}(\boldsymbol{\Psi}_0)^{-1} \quad (9)$$

$$\mathbf{H}(\boldsymbol{\Psi}_0) = 2\dot{\mathbf{A}}^H \boldsymbol{\Pi}(\boldsymbol{\Psi}_0) \dot{\mathbf{A}} \quad (10)$$

$$\dot{\mathbf{A}} = \begin{bmatrix} \frac{\partial \mathbf{a}(\boldsymbol{\Psi}_0)}{\partial \boldsymbol{\Psi}_0(1)} & \dots & \frac{\partial \mathbf{a}(\boldsymbol{\Psi}_0)}{\partial \boldsymbol{\Psi}_0(L)} \end{bmatrix} \quad (11)$$

$$\boldsymbol{\Pi}(\boldsymbol{\Psi}_0) = \mathbf{I}_N - \frac{\mathbf{a}(\boldsymbol{\Psi}_0) \mathbf{a}(\boldsymbol{\Psi}_0)^H}{\mathbf{a}(\boldsymbol{\Psi}_0)^H \mathbf{a}(\boldsymbol{\Psi}_0)} \quad (12)$$

Note that for the 2D DOA estimation problem and for $\boldsymbol{\Psi}_0 = \boldsymbol{\Theta}_0$, expression (9) has also been obtained in [4] for a 2D array with identical sensors. Expression (9) is now generalizing these results for a general array with potentially different sensors and also for $\boldsymbol{\Psi}_0 = \mathbf{k}(\boldsymbol{\Theta}_0)$ or $\mathbf{k}_{2D}(\boldsymbol{\Theta}_0)$.

3.2 Deterministic Cramer Rao Bound

We talk about Deterministic CRB when the signal $s(t)$ is assumed to be deterministic and the noise vector $\mathbf{n}(t)$ is assumed to be zero mean, Gaussian with a covariance matrix defined by $\mathbb{E}[\mathbf{n}(t)\mathbf{n}(t)^H] = \sigma^2 \mathbf{I}_N$. Under these assumptions, the unknown parameter vector which has to be estimated is $\boldsymbol{\Omega}_0 = [\boldsymbol{\Psi}_0^T \ \sigma^2 s(t_k)]^T$ for $(1 \leq k \leq K)$. In this context, it can be shown, after tedious computations, that the deterministic CRB on $\boldsymbol{\Psi}_0$, denoted by $\mathbf{CRB}_{det}(\boldsymbol{\Psi}_0)$, is

$$\mathbf{CRB}_{det}(\boldsymbol{\Psi}_0) = \frac{1}{K \left(\frac{\hat{r}_{ss}}{\sigma^2} \right)} \mathbf{H}(\boldsymbol{\Psi}_0)^{-1} \quad (13)$$

where $\hat{r}_{ss} = (1/K) \sum_{k=1}^K |s(t_k)|^2$ and where $\mathbf{H}(\boldsymbol{\Psi}_0)$ is defined by (10). Note that, to our knowledge, for the 2D DOA estimation problem, expression (13) is new. Note that for

the 1D DOA estimation problem the deterministic CRB is given in [1]. Comparing (9) and (13) we obtain

$$\mathbf{CRB}_{\text{det}}(\Psi_0) = \frac{\left(\frac{\gamma}{\bar{r}_{ss}}\right)}{1 + \frac{\sigma^2}{\gamma N}} \mathbf{CRB}_{\text{stoc}}(\Psi_0)$$

which shows that both $\mathbf{CRB}_{\text{det}}(\Psi_0)$ and $\mathbf{CRB}_{\text{stoc}}(\Psi_0)$ are proportional to $\mathbf{H}(\Psi_0)^{-1}$ and that they become approximately equal if $\gamma N/\sigma^2 \gg 1$ provided that K is not too small.

4. MUSIC PERFORMANCE DERIVATION

We assume in this section that the noise is spatially white ($\mathbf{R}_{nn} = \sigma^2 \mathbf{I}_N$) and we consider the presence of modelling errors ($\mathbf{e} \neq 0$). Under these assumptions, in the presence of a single source of direction Ψ_0 , the estimated DOA, $\hat{\Psi}_0$, obtained by the MUSIC [6] algorithm from the K observation vectors $[\mathbf{x}(t_1), \dots, \mathbf{x}(t_K)]$ corresponds to the vector Ψ which minimizes the criterion

$$J(\Psi) = \mathbf{a}(\Psi)^H \hat{\Pi}(K, \mathbf{e}) \mathbf{a}(\Psi) \quad (14)$$

where $\hat{\Pi}(K, \mathbf{e})$ is the orthogonal projection matrix onto the noise subspace of the estimated correlation matrix of the observations, $\hat{\mathbf{R}}_{xx} = (1/K) \sum_{k=1}^K \mathbf{x}(t_k) \mathbf{x}(t_k)^H$. The estimated direction is then

$$\hat{\Psi}_0 = \min_{\Psi} J(\Psi)$$

In the absence of modelling error ($\mathbf{e} = 0$) and for an infinite value of K , $\hat{\Pi}(K, \mathbf{e})$ tends toward $\Pi(\Psi_0)$ defined by (12) and $\Delta\Psi_0 = \hat{\Psi}_0 - \Psi_0 = 0$. However, for a finite number of samples K or in the presence of modelling errors, the projector error $\Delta\Pi(K, \mathbf{e}) = \Pi(\Psi_0) - \hat{\Pi}(K, \mathbf{e})$ is not zero.

To our knowledge, the performance analyses of MUSIC available in the literature assume that Ψ_0 is a scalar. For the 2D DOA estimation problem, no result seem to be published. Assuming that the error vector $\Delta\Psi_0$ has a small norm, which requires that \mathbf{e} has a small norm and K is not too small, and using a first order series expansion of the Gradient of $J(\hat{\Psi}_0)$ around $\hat{\Psi}_0 = \Psi_0$, the DOA error is

$$\Delta\Psi_0 = \hat{\Psi}_0 - \Psi_0 = -\mathbf{H}_0 \left(\hat{\Pi}(K, \mathbf{e}) \right)^{-1} \nabla_0 \left(\hat{\Pi}(K, \mathbf{e}) \right) \quad (15)$$

where $\mathbf{H}_0 \left(\hat{\Pi}(K, \mathbf{e}) \right)$ and $\nabla_0 \left(\hat{\Pi}(K, \mathbf{e}) \right)$ are the Hessian and Gradient respectively of the MUSIC criterion (14) such that

$$\begin{aligned} \mathbf{H}_0 \left(\hat{\Pi}(K, \mathbf{e}) \right) (i, j) &= 2 \left[\frac{\partial \mathbf{a}(\Psi_0)^H}{\partial \Psi_0(i)} \hat{\Pi}(K, \mathbf{e}) \frac{\partial \mathbf{a}(\Psi_0)}{\partial \Psi_0(j)} \right] \quad (16) \\ \nabla_0 \left(\hat{\Pi}(K, \mathbf{e}) \right) (i) &= 2\Re \left(\frac{\partial \mathbf{a}(\Psi_0)^H}{\partial \Psi_0(i)} \hat{\Pi}(K, \mathbf{e}) \mathbf{a}(\Psi_0) \right) \end{aligned}$$

To be able to derive the MSE matrix $\mathbf{MS}_{\Psi_0} = \mathbb{E}[\Delta\Psi_0(\Delta\Psi_0)^T]$, it is necessary to use a first order series expansion of $\hat{\Pi}(K, \mathbf{e})$ around $\Pi(\Psi_0)$ in (15) and (16), as small errors are assumed. Using such an expansion and using the fact that $\mathbf{H}(\Psi_0) = \mathbf{H}_0(\Pi(\Psi_0))$, it is possible to show, after tedious computations, that $\Delta\Psi_0$ is approximated by

$$\Delta\Psi_0 \approx -\mathbf{H}(\Psi_0)^{-1} \nabla_0(\Delta\Pi(K, \mathbf{e})) \quad (17)$$

and

$$\begin{aligned} \mathbf{MS}_{\Psi_0} &= \mathbb{E}[\Delta\Psi_0(\Delta\Psi_0)^T] \approx \quad (18) \\ \mathbf{H}(\Psi_0)^{-1} \mathbb{E}[\nabla_0 \left(\hat{\Pi}(K, \mathbf{e}) \right) (\nabla_0 \left(\hat{\Pi}(K, \mathbf{e}) \right))^T] \mathbf{H}(\Psi_0)^{-1} \end{aligned}$$

4.0.1 Performance with finite number of samples

In this section, the MSE matrix \mathbf{MS}_{Ψ_0} (18) is derived for a finite number of samples K and without modelling errors. Under these assumptions, it can be shown [5] that $\Delta\Pi(K, \mathbf{e}=0) = \Delta\Pi(K) \approx \Pi(\Psi_0) \Delta\mathbf{R}_{xx} \mathbf{R}_{yy}^\# + (\mathbf{R}_{yy}^\#)^H \Delta\mathbf{R}_{xx} \Pi(\Psi_0)$ where $\mathbf{R}_{yy} = \mathbf{R}_{xx} - \sigma^2 \mathbf{I}_N$, $\Delta\mathbf{R}_{xx} = \mathbf{R}_{xx} - \hat{\mathbf{R}}_{xx}$ and where $(\cdot)^\#$ is the Moore-Penrose pseudo-inverse. The i^{th} component of $\nabla_0(\Delta\Pi(K, \mathbf{e}=0))$ is

$$\nabla_0(\Delta\Pi(K))(i) = 2\Re \left(\frac{\partial \mathbf{a}(\Psi_0)^H}{\partial \Psi_0(i)} \Pi(\Psi_0) \Delta\mathbf{R}_{xx} \mathbf{R}_{yy}^\# \mathbf{a}(\Psi_0) \right) \quad (19)$$

Assuming Gaussian observations, the random variable $\Delta\mathbf{R}_{xx}$ has a Wishart statistic and we obtain $\mathbb{E}[\Delta\mathbf{R}_{xx}] = 0$ and

$$\mathbb{E}[\Delta\mathbf{R}_{xx}(i, j) \Delta\mathbf{R}_{xx}(m, n)] = \frac{\mathbf{R}_{xx}(m, j) \mathbf{R}_{xx}(i, n)}{K} \quad (20)$$

From (18)(19)(20), we deduce that

$$\mathbf{MS}_{\Psi_0} = \mathbf{MS}(K) = \frac{1}{K \left(\frac{\gamma}{\sigma^2}\right)} \mathbf{H}(\Psi_0)^{-1} \quad (21)$$

where $\gamma = \mathbb{E}[|s(t)|^2]$, which is a new result and which is proportional to $\mathbf{H}(\Psi_0)^{-1}$.

4.0.2 Performance with modelling errors

In this section, the MSE matrix \mathbf{MS}_{Ψ_0} (18) is derived for an infinite number of samples K but in the presence of modelling errors. Under these assumptions, it can be shown [2] that $\Delta\Pi(K \rightarrow \infty, \mathbf{e}) = \Delta\Pi(\mathbf{e}) \approx \Pi(\Psi_0) \mathbf{e} \mathbf{a}(\Psi_0)^\# + \mathbf{a}(\Psi_0)^\# \mathbf{e}^H \Pi(\Psi_0)$ where $\mathbf{a}(\Psi_0)^\# = \mathbf{a}(\Psi_0)^H / (\mathbf{a}(\Psi_0)^H \mathbf{a}(\Psi_0))$. The i^{th} component of $\nabla_0(\Delta\Pi(\mathbf{e}))$ is

$$\nabla_0(\Delta\Pi(\mathbf{e}))(i) = 2\Re \left(\frac{\partial \mathbf{a}(\Psi_0)^H}{\partial \Psi_0(i)} \Pi(\Psi_0) \mathbf{e} \right) \quad (22)$$

Assuming that \mathbf{e} is a zero-mean, circular and Gaussian random vector such that $\mathbb{E}[\mathbf{e}\mathbf{e}^H] = \sigma_e^2 \mathbf{I}_N$, expression (18) becomes

$$\mathbf{MS}_{\Psi_0} = \mathbf{MS}(\sigma_e) = \sigma_e^2 \mathbf{H}(\Psi_0)^{-1} \quad (23)$$

Comparing (21) and (23), we find that

$$\mathbf{MS}(K) = \frac{1}{\sigma_e^2 K \left(\frac{\gamma}{\sigma^2}\right)} \mathbf{MS}(\sigma_e)$$

which is a new result and which shows that $\mathbf{MS}(K)$ and $\mathbf{MS}(\sigma_e)$ are proportional to each other and to $\mathbf{H}(\Psi_0)^{-1}$.

5. FIRST TOOLS FOR AN ANTENNA ARRAY OPTIMIZATION

In the previous sections, the accuracy of the vector $\hat{\Psi}_0$ has been characterized by the MSE matrix $\mathbf{MS}_{\Psi_0} = \mathbb{E}[\Delta\Psi_0(\Delta\Psi_0)^T]$. Considering the stochastic or the deterministic CRB or the MUSIC algorithm, with or without modelling errors, it has been shown that \mathbf{MS}_{Ψ_0} is, in all cases, proportional to the matrix $\mathbf{H}(\Psi_0)^{-1}$, which is the only parameter depending on the sensors location. In this section, we investigate more precisely the links between $\mathbf{H}(\Psi_0)^{-1}$ and the sensors location in order to generate tools for an array design optimization.

In practical situations, we may prefer to characterize DOA estimation accuracy by a scalar quantity, easier to handle and interpret. In practice, the only important quality criterion is the angle between the wave-vector $\mathbf{k}(\Theta_0)$ and its estimate $\hat{\mathbf{k}}(\Theta_0)$. For this reason, for given values of $\mathbf{k} = \mathbf{k}(\Theta_0)$ and $\hat{\mathbf{k}} = \hat{\mathbf{k}}(\Theta_0)$, we propose the following scalar criterion

$$C(\mathbf{k}) = \mathbb{E} \left[d(\mathbf{k}, \hat{\mathbf{k}}) \right] \text{ where}$$

$$d(\mathbf{k}, \hat{\mathbf{k}}) = \frac{\mathbf{k}^H \mathbf{\Pi}(\hat{\mathbf{k}}) \mathbf{k}}{\mathbf{k}^H \mathbf{k}} = 1 - \frac{|\mathbf{k}^H \hat{\mathbf{k}}|^2}{(\mathbf{k}^H \mathbf{k})(\hat{\mathbf{k}}^H \hat{\mathbf{k}})}$$

where $\mathbf{\Pi}(\hat{\mathbf{k}}) = \mathbf{I}_d - \hat{\mathbf{k}}\hat{\mathbf{k}}^H$ and $0 \leq C(\mathbf{k}) \leq 1$. Note that \mathbf{k} is a normalized vector but $\hat{\mathbf{k}}$ may not be a unit norm one. It is in particular the case when $\mathbf{k} = \mathbf{k}(\Theta_0)$, since it is implicitly assumed in this case that the three components of $\mathbf{k}(\Theta_0)$ are independent in the search procedure. Assuming small errors, $\Delta \mathbf{k} = \hat{\mathbf{k}} - \mathbf{k}$ and using a second order serial expansion of $\mathbf{\Pi}(\hat{\mathbf{k}})$ around $\hat{\mathbf{k}} = \mathbf{k}$ according to [2], it is possible to show that $D(\mathbf{k})$ is

$$C(\mathbf{k}) \approx \mathbf{k}^H (\text{trace}(\mathbf{MS}_{\mathbf{k}}) \mathbf{I}_d - \mathbf{MS}_{\mathbf{k}}) \mathbf{k}$$

where $\text{trace}(\cdot)$ is the matrix trace. As it has been shown in the previous sections that $\mathbf{MS}_{\mathbf{k}} = \alpha \mathbf{H}(\mathbf{k})^{-1}$ where α is a scalar quantity which is independent of the sensors location and where $\mathbf{H}(\mathbf{k})$ depends on the sensors location, we obtain

$$C(\mathbf{k}) \approx \alpha \mathbf{k}^H (\text{trace}(\mathbf{H}(\mathbf{k})^{-1}) \mathbf{I}_d - \mathbf{H}(\mathbf{k})^{-1}) \mathbf{k} \quad (24)$$

For a given number of sensors N , the typical problem of antenna array design optimization consists to find the array geometry, or the sensors' locations vectors $\mathbf{p}_n (1 \leq n \leq N)$, which minimizes $C(\mathbf{k})$, over a given space area, under some possible constraints. One possible constraint may be to limit to a given level, the rank-1 ambiguity of the array. An other possible constraint may be to impose an omnidirectional behaviour of the array, i.e a constant value of $C(\mathbf{k})$, inside the selected space area. In the following section, we analyze the link between $\mathbf{H}(\mathbf{k})^{-1}$ and the vectors $\mathbf{p}_n (1 \leq n \leq N)$ for 3D, 2D and 1D arrays respectively. This links allow to generate tools for an array design optimization, not presented in this paper.

5.1 3D array case

For a general 3D array, according to (11), the derivative matrix $\dot{\mathbf{A}}$ with $\Psi_0 = \mathbf{k}(\Theta_0) = [u \ v \ w]^T$ is

$$\dot{\mathbf{A}}(\mathbf{k}(\Theta_0)) = j \frac{2\pi}{\lambda} \begin{bmatrix} a_1(\Theta_0) \mathbf{p}_1^H \\ \vdots \\ a_N(\Theta_0) \mathbf{p}_N^H \end{bmatrix} \quad (25)$$

and according to (10), it is possible to verify that the inverse of matrix $\mathbf{H}(\mathbf{k}(\Theta_0))$, which is full rank for a general 3D array, is

$$\mathbf{H}(\mathbf{k}(\Theta_0))^{-1} = \frac{1}{2N} \left(\frac{2\pi}{\lambda} \right)^{-2} (\mathbf{D}_{pp})^{-1} \quad (26)$$

$$\mathbf{D}_{pp} = \frac{1}{N} \sum_{n=1}^N (\mathbf{p}_n - \bar{\mathbf{p}})(\mathbf{p}_n - \bar{\mathbf{p}})^T \text{ with } \bar{\mathbf{p}} = \frac{1}{N} \sum_{n=1}^N \mathbf{p}_n \quad (27)$$

which only depends on matrix $\mathbf{P} = [\mathbf{p}_1 \ \cdots \ \mathbf{p}_N]$ of the sensors positions vectors. Expression (26) shows that

$\mathbf{H}(\mathbf{k}(\Theta_0))^{-1}$ and then $\mathbf{MS}_{\mathbf{k}(\Theta_0)}$ is not dependent of $\mathbf{k}(\Theta_0)$ and only depends, to within a scalar term, to the sensors locations.

Let us now get some insights into the conditions that have to be satisfied by the 3D array to become omnidirectional. To this aim, we denote by $\mathbf{M}(\mathbf{P})$ the hermitian matrix $\mathbf{M}(\mathbf{P}) = \text{trace}(\mathbf{H}(\mathbf{k}(\Theta_0))^{-1}) \mathbf{I}_d - \mathbf{H}(\mathbf{k}(\Theta_0))^{-1}$ which only depends on \mathbf{P} according to (26). Introducing the minimal, $\lambda_{\min}(\mathbf{M}(\mathbf{P}))$, and maximal, $\lambda_{\max}(\mathbf{M}(\mathbf{P}))$, eigenvalue of $\mathbf{M}(\mathbf{P})$, it is well known, from (24)(26), that

$$\alpha \lambda_{\min}(\mathbf{M}(\mathbf{P})) = C(\mathbf{k}_{\min}) \leq C(\mathbf{k}) \leq \alpha \lambda_{\max}(\mathbf{M}(\mathbf{P})) = C(\mathbf{k}_{\max}) \quad (28)$$

where \mathbf{k}_{\min} , \mathbf{k}_{\max} are the eigenvectors of $\mathbf{M}(\mathbf{P})$ associated with the eigenvalue $\lambda_{\min}(\mathbf{M}(\mathbf{P}))$, $\lambda_{\max}(\mathbf{M}(\mathbf{P}))$ respectively. This expression shows that, for a given array geometry \mathbf{P} , the directions associated with the lowest and the highest estimation quality are orthogonal to each other, result already found in [4] for a 2D array. We deduce from (28) that the array is omnidirectional if and only if $\lambda_{\min}(\mathbf{M}(\mathbf{P})) = \lambda_{\max}(\mathbf{M}(\mathbf{P}))$, which means that $\mathbf{M}(\mathbf{P})$ is proportional to the identity matrix. In such a case, $\mathbf{D}_{pp} = D_{xx} \mathbf{I}_3$ and

$$\mathbb{E}[(u - \hat{u})^2] = \mathbb{E}[(v - \hat{v})^2]$$

$$= \mathbb{E}[(w - \hat{w})^2] = \left(\frac{\sqrt{\alpha}}{\pi \left(\frac{D_{3,e}}{\lambda} \right) \sqrt{N/2}} \right)^2$$

where $D_{3,e}$ is the equivalent aperture of the 3D array. This quantity is defined from the Uniform Spherical Array (USA) of radius R , which is omnidirectional since $\mathbf{D}_{pp} = ((2R)^2/16) \mathbf{I}_3$. Then, $D_{3,e}$ is such that $(D_{3,e})^2 = \nu D_{xx}$ in order to have $D_{3,e} = 2R$ for a USA of radius R , and is given by

$$D_{3,e} = \sqrt{16 \frac{\text{trace}(\mathbf{D}_{pp})}{3}} \quad (29)$$

From these results and using the fact that, according to (4)(11), the link between $\dot{\mathbf{A}}(\Theta_0)$ and $\dot{\mathbf{A}}(\mathbf{k}(\Theta_0))$ is for a 3D array

$$\dot{\mathbf{A}}(\Theta_0) = \dot{\mathbf{A}}(\mathbf{k}(\Theta_0)) \mathbf{J}_3(\Theta_0) \quad (30)$$

$$\mathbf{J}_3(\Theta_0) = \begin{bmatrix} -\sin(\theta_0) \cos(\Delta_0) & -\cos(\theta_0) \sin(\Delta_0) \\ \cos(\theta_0) \cos(\Delta_0) & -\sin(\theta_0) \sin(\Delta_0) \\ 0 & \cos(\Delta_0) \end{bmatrix}$$

the Mean Square error matrix of Θ_0 is

$$\mathbb{E}[\Delta \Theta_0 (\Delta \Theta_0)^T] = \frac{\alpha}{2N \left(\frac{2\pi}{\lambda} \right)^2} \left(\mathbf{J}_3(\Theta_0)^T \mathbf{D}_{pp} \mathbf{J}_3(\Theta_0) \right)^{-1} \quad (31)$$

We then deduce from (29)(31) that, for a 3D omnidirectional array, the Θ_0 components mean square error are

$$\mathbb{E}[(\theta_0 - \hat{\theta}_0)^2] = \left(\frac{\sqrt{\alpha}}{\pi \left(\frac{D_{3,e}}{\lambda} \right) \cos(\Delta_0) \sqrt{N/2}} \right)^2 \quad (32)$$

$$\mathbb{E}[(\Delta_0 - \hat{\Delta}_0)^2] = \left(\frac{\sqrt{\alpha}}{\pi \left(\frac{D_{3,e}}{\lambda} \right) \sqrt{N/2}} \right)^2$$

We note that the mean square error of $\hat{\theta}_0$ and $\hat{\Delta}_0$ are independent of θ_0 , hence the so-called azimuthal omnidirectionality of these estimates. However the mean square error of $\hat{\theta}_0$ depends on Δ_0 and increases to infinite for $\Delta_0 = \pi/2$, direction for which the concept of azimuth is no longer defined. To our knowledge, the results of section 5.1 have never been explicitly published for 3D arrays.

5.2 2D array case

For a 2D array, the vector \mathbf{p}_n is reduced to $\mathbf{p}_n = [x_n \ y_n]^T$, the steering vector of a source does no longer depend on w and the matrix $\mathbf{H}(\mathbf{k}(\Theta_0))$ defined previously is not full rank. This means that it is no longer possible to estimate uniquely and independently the three components of vector $\mathbf{k}(\Theta_0)$ from a 2D array. However, limiting the estimation to the two first independent components (u, v) of $\mathbf{k}(\Theta_0)$, the results developed in sections 3, 4 and 5 can still be applied. Under these assumptions, it is possible to show that $\mathbf{H}(\mathbf{k}(\Theta_0))^{-1}$, which is now a square matrix, is still given by (26) where \mathbf{D}_{pp} , defined by (27), is

$$\mathbf{D}_{pp} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (33)$$

The omnidirectionality condition is obtained in a similar way as for the 3D array. Introducing from (33) the coefficients ρ and η defined by

$$\rho^2 = \frac{(D_{xy})^2}{D_{xx}D_{yy}} \quad \text{and} \quad \eta^2 = \frac{D_{yy}}{D_{xx}} \quad (34)$$

A 2D array is then *omni-directional* when $\rho = 0$ and $\eta = 1$. According to (33)(34) and the definition of $\mathbf{H}(\mathbf{k}(\Theta_0))$, assuming \mathbf{D}_{pp} is full rank, the mean square error of the wave-vector components are, for a 2D omnidirectional array

$$\begin{aligned} \mathbb{E}[(u - \hat{u})^2] &= \frac{\alpha}{N(\frac{\pi}{\lambda})^2} \frac{1+\eta^2}{2(D_{2,e})^2(1-\rho^2)} \\ \mathbb{E}[(v - \hat{v})^2] &= \frac{\mathbb{E}[(u - \hat{u})^2]}{\eta^2} \\ \mathbb{E}[(u - \hat{u})(v - \hat{v})] &= -\mathbb{E}[(u - \hat{u})^2] \frac{\rho}{\eta} \end{aligned} \quad (35)$$

where $D_{2,e}$ is the equivalent aperture of the 2D array. This value is defined from the Uniform Circular Array (UCA) of radius R , which is omnidirectional since $\mathbf{D}_{pp} = ((D_{2,e})^2/8)\mathbf{I}_2$. Then, $D_{2,e}$ is such that $(D_{2,e})^2 = \nu D_{xx}$ in order to have $D_{2,e} = 2R$ for a UCA of radius R , and is given by

$$D_e^2 = \sqrt{8 \frac{\text{trace}(\mathbf{D}_{pp})}{2}} \quad (36)$$

Expression (35) shows that the mean square errors on the estimate of (u, v) do not depend on the angle of arrival. From these results and using the fact that, according to (4)(11), the link between $\dot{\mathbf{A}}(\Theta_0)$ and $\dot{\mathbf{A}}(\mathbf{k}(\Theta_0))$ is, for a 2D array

$$\dot{\mathbf{A}}(\Theta_0) = \dot{\mathbf{A}}(\mathbf{k}(\Theta_0)) \mathbf{J}_2(\Theta_0) \quad (37)$$

$$\mathbf{J}_2(\Theta_0) = \begin{bmatrix} -\sin(\theta_0) \cos(\Delta_0) & -\cos(\theta_0) \sin(\Delta_0) \\ \cos(\theta_0) \cos(\Delta_0) & -\sin(\theta_0) \sin(\Delta_0) \end{bmatrix}$$

the mean square error of Θ_0 is given, for a 2D omnidirectional array, by

$$\mathbb{E}[\Delta\Theta_0(\Delta\Theta_0)^T] = \frac{\alpha}{2N(\frac{2\pi}{\lambda})^2} \left(\mathbf{J}_2(\Theta_0)^T \mathbf{D}_{pp} \mathbf{J}_2(\Theta_0) \right)^{-1} \quad (38)$$

For a 2D array, the mean square error of the direction Θ_0 components are

$$\begin{aligned} \mathbb{E}[(\theta_0 - \hat{\theta}_0)^2] &= \left(\frac{RMS}{\cos(\Delta_0)} \right)^2 f(\theta_0) \\ \mathbb{E}[(\Delta_0 - \hat{\Delta}_0)^2] &= \left(\frac{RMS}{\sin(\Delta_0)} \right)^2 f\left(\theta_0 + \frac{\pi}{2}\right) \\ RMS &= \frac{\sqrt{\alpha}}{\pi \left(\frac{D_{2,e}}{\lambda}\right) \sqrt{N}} \sqrt{\frac{1+\eta^2}{2(1-\rho^2)}} \\ f(\theta) &= \sin(\theta)^2 + \frac{\cos(\theta)^2}{\eta^2} + \sin(2\theta) \frac{\rho}{\eta} \end{aligned} \quad (39)$$

which shows that the DOA accuracy depends on Θ_0 and the array parameters such as $D_{2,e}$, ρ and η .

5.3 1D array case

For a 1D array, the vector \mathbf{p}_n is reduced to $\mathbf{p}_n = x_n$, the steering vector of a source does no longer depend on v and w and we have to limit the estimation to the first components u of $\mathbf{k}(\Theta_0)$. In this case, \mathbf{D}_{pp} is a scalar equal to $(D_{1,e})^2/2$ where $D_{1,e}$ is the equivalent aperture of the array and according to (26), the mean square error of the wave-vector component is

$$\mathbb{E}[(u - \hat{u})^2] = \left(\frac{\sqrt{\alpha}}{2\pi \left(\frac{D_{1,e}}{\lambda}\right) \sqrt{N}} \right)^2 \quad (40)$$

5.4 Conclusion

In this paper and for the single source case, it has been shown that both stochastic and deterministic CRB, jointly with the MUSIC performance with and without modelling errors, are proportional to the matrix $\mathbf{H}(\Psi_0)^{-1}$, which is the only parameter depending on the sensors location. The precise links between $\mathbf{H}(\Psi_0)^{-1}$, and thus the DOA estimation precision, and the sensors location has been analyzed in a second time. This gives rise to analytical conditions on the sensors location to generate for example omnidirectional arrays or arrays with a maximal equivalent aperture. These conditions are important tools for an array design methodology which will be presented elsewhere.

REFERENCES

- [1] H. Clergeot, S. Tressens, and A. Ouamri. Performance of high resolution frequencies estimation methods compared to the Cramer-Rao bounds. *IEEE Trans. Acoust., Speech, Signal Processing*, 37(11):1703–1720, Nov. 1989.
- [2] A. Ferréol, P. Larzabal, and M. Viberg. On the asymptotic performance analysis of subspace DOA estimation in the presence of modeling errors: Case of MUSIC. *IEEE Trans. Signal Processing*, 54(3):907–920, Mar. 2006.
- [3] B. Friedlander. Sensitivity analysis of the MUSIC algorithm. *IEEE Trans. Acoust., Speech, Signal Processing*, 38(10):1740–1751, July 1990.
- [4] H. Gazzah and S. Marcos. Cramer-Rao Bounds for Antenna Array Design. *IEEE Trans. Signal Processing*, 54(1):336–345, Jan. 2006.
- [5] H. Krim, P. Forster, and J.G. Proakis. Operator Approach to Performance Analysis of Root-MUSIC and Root-Min-Norm. *IEEE Trans. Acoust., Speech, Signal Processing*, 40(7):1687–1696, July 1992.
- [6] R. O. Schmidt. *A signal subspace approach to multiple emitter location and spectral estimation*. PhD thesis, Stanford University, Stanford, CA, Nov. 1981.
- [7] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and the Cramer-Rao bound: further results and comparisons. *IEEE Trans. Acoust., Speech, Signal Processing*, 38:2140–2150, 1990.
- [8] A. Swindlehurst and T. Kailath. A performance Analysis of Subspace-Based method in the presence of model errors, Part I: The MUSIC algorithm. *IEEE Trans. Signal Processing*, 40(3):1758–1774, July 1992.
- [9] A. Weiss and B. Friedlander. ON the Cramer Rao Bound for direction finding of correlated signals. *Proc. IEEE Int. Conf. Asilomar, Signals, Systems and Computers*, 2, Nov. 1990.