ASYMPTOTIC PERFORMANCE ANALYSIS OF DIRECTION FINDING ALGORITHMS BASED ON FOURTH-ORDER CUMULANTS.

Jean-François Cardoso, Éric Moulines
Ecole Nationale Supérieure des Télécommunications / CNRS URA 820 / GDR TdSI, 46 rue Barrault, 75634 PARIS
Email: cardoso.moulines@ens-lyon.fr, Tel: 33-1-47-81-78-59, Fax: 33-1-47-88-79-35

Abstract — In the narrow band array processing context, the use of higher-order statistics has been often advocated because consistent and asymptotically unbiased parameter estimates can be obtained without it being necessary to know, to model or to estimate the spatial covariance of the noise as long as it is normally distributed. However, experimentation shows that this "noise insensitivity" is traded for increased variability of the parameter estimates. The main purpose of this contribution is to derive and work out closed form expressions of the asymptotic covariance of MUSIC-like direction-of-arrival estimates based on two fourth-order cumulant matrices: the diagonal slice and the contracted quadrivariate. This is compared to the standard covariance based MUSIC estimate establishing on a rational basis the domain of applicability of higher-order statistics for DOA estimation. In particular, the actual impact of the noise variance and of the dynamic range of the sources is investigated. This asymptotic performance analysis is achieved within a general framework, which we believe to be of general interest.

Keywords — Direction finding, array processing, performance analysis, higher order statistics, cumulants.

INTRODUCTION

Current narrow-band array processing techniques are based on the second-order statistics of the received signals. In many situations, the received signals are non-Gaussian: they convey valuable statistical information in their moments of order greater than two (this is in particular true when considering communications signals.) In these circumstances, it makes sense to develop array processing techniques using this higher-order information. Of particular interest are the algorithms based on higher-order cumulants of the array output, since these statistics show the distinctive property of being, in a certain sense, insensitive to Gaussian noise, making it possible to devise consistent parameter estimates without it being necessary to know, to model or to estimate the noise covariance.

The signals of interest, in the context of narrow-band array processing, are most of the time complex ‘circular’, which implies that their third-order cumulants are zeros. To cope with these signals, it is thus necessary to use even order cumulants. Computationally considerations as well as robustness to noise dictate the use of 4th-order cumulants. Of course, there are more ‘quadruples’ than ‘pairs’ of indices; this larger number of statistics can be exploited to find the solution of identification problems which could not have been solved using 2nd-order statistics only (such as blind identification of the directional vectors [1]).

In many situations, however, it is sufficient to deal with a reduced set of cumulants. This is in particular the case of direction of arrival (DOA) estimation using a calibrated array, which is the problem addressed in this contribution. In the sequel, we consider two 4th-order cumulant-based matrices, the diagonal slice [2], [3] and the contracted quadrivariate [4], [5]. These two matrices show the desirable properties of retaining sufficient information for identification of the DOA parameters in Gaussian noise with unknown spatial covariance, requiring an overall computational effort similar to the second-order techniques. These cumulant matrices have the same size as the second-order covariance: from their column space, the DOAs can be estimated using the MUSIC algorithm just as with covariance-based MUSIC [6]. This offers the opportunity for a close comparison between 2nd- and 4th-order techniques. Fourth-order methods are generically ‘inensitive’ to additive Gaussian noise in the sense that asymptotically unbiased estimates can be obtained without modelling the noise structure, but this desirable feature is traded for increased variability. This effect has been investigated in several contributions, mainly based on numerical simulations and evaluations (see for example [7], [8], [9]) but closed-form expressions for the covariance of the DOAs have never been worked out. It is the main purpose of this paper to derive such expressions and to discuss the respective merits of 2nd- and 4th-order statistics. In particular, the actual impact on asymptotic performance of the noise variance and of the dynamic range of the sources is evidenced, helping to determine the domain of applicability of 4th-order techniques.

The paper is organized as follows. In section I, the generic problem of DOA estimation in narrow-band array processing context is reviewed. In section II, the two cumulant matrices under consideration are defined; MUSIC-like algorithms for DOA estimation are then derived. In section III, a general functional approach providing a common unifying framework for asymptotic performance analysis is presented. It is used in section IV to derive closed-form expressions of the asymptotic covariance of the DOA estimates. These expressions are further investigated in section V via explicit workouts in some simple scenarios. Some technical details and calculations are given in appendix.

I. Problem formulation.

Let an array of m sensors receive n narrow-band plane waves from n discrete sources. In complex notations, the m-vector of sensor outputs is modelled by the following relation

\[ y(t) = x(t) + b(t) = A(\theta) s(t) + b(t) \]  

(1)

The n × 1 real vector \( \theta = [\theta_1, \ldots, \theta_n]^T \) corresponds to the unknown signal parameters, referred to as the direction-of-arrivals (DOA); the columns of the matrix \( A = A(\theta) \) are the directional vectors \( A = [a(\theta_1), \ldots, a(\theta_n)]^T \). The n × 1 vector \( s(t) = [s_1(t), \ldots, s_n(t)]^T \) contains the complex envelopes of the emitted signals and \( b(t) \) is a complex m × 1 vector of additive noise. The m × 1 vector \( a(\theta_i) \) models the array response to a unit amplitude wavefront having signal parameter \( \theta_i \). The array manifold is defined to be the set \( A = \{a(\theta) | \theta \in \Theta \} \) for some
region $\Theta$ in DOA space. The set $\mathcal{A}$ is assumed to be known, either analytically or via some calibration procedure. It is also assumed to be unambiguous, i.e., for any collection of $n$, $n \leq m$, distinct DOAs $\theta_i$, the corresponding vectors $s(\theta_i)$ are linearly independent. Equivalently, the matrix $A(\theta)$ is assumed to be full rank, for all values of $\theta$.

Each emitter waveform $s_p(t)$ is modelled as a sequence of independent and identically distributed (i.i.d.) complex random variables. The distribution of each random variable $s'_p(t)$ is assumed to be non-Gaussian circular $^1$ with finite moments up to the 4th-order and non-vanishing 4th-order cumulant. We also assume that the signal emitted by the sources are not fully correlated, i.e., the matrix $\mathbf{R}_s \overset{\text{def}}{=} E(s(t)s^*(t)^*)$ is full-rank (where $s(t)^*$ denotes the conjugate-transpose of $s(t)$).

The additive noise, $b(t)$, is modelled as an i.i.d. sequence of complex zero-mean Gaussian variables. The noise is assumed to be statistically independent from the signal waveforms; in contrast with 2nd-order methods, its spatial covariance, $\mathbf{R}_b \overset{\text{def}}{=} E(b(t)b(t)^*)$, is not assumed to be known. Finally, the array output is sampled at $N$ discrete time instants $\{y(1), \ldots, y(N)\}$. Based on this snapshot, the number of signals and their directions of arrival must be estimated. We assume in the following that the number of signals is known and concern ourselves only with the problem of estimating the DOAs.

The second-order moments of the random vector $s(t)$ are defined as

$$
\mu_i^j = E(s_i(t)s^*_j(t)) \quad 1 \leq i, j \leq n
$$

where $s_i(t)$ denotes the $i$-th component of the vector $s(t)$ and $s^*_j(t)$ is the $j$-th component of the vector $s(t)^*$ (this notational convention applies throughout). Similarly, the 4th-order moments are

$$
\mu_{ik}^{jl} = E(s_i(t)s^*_k(t)s_j(t)s^*_l(t)) \quad 1 \leq i, j, k, l \leq n
$$

Note that the 4th-order moments are defined so that two out of the four factors are account for, to be otherwise equal to zero, because of the circularity condition. The 2nd- and 4th-order cumulants of the random vector $s(t)$ are defined as [10], [9]

$$
\gamma_i^j = \text{Cum}(s_i(t), s^*_j(t)) \quad 1 \leq i, j \leq n
$$

$$
\kappa_{ik}^{jl} = \text{Cum}(s_i(t), s^*_k(t), s_j(t), s^*_l(t)) \quad 1 \leq i, j, k, l \leq n
$$

Extension of these notations to moments of higher order is obvious and omitted for brevity. Under the circularity assumptions, the definition of 2nd- and 4th-order cumulants reduces to

$$
\gamma_i^j = \mu_i^j
$$

$$
\kappa_{ik}^{jl} = \mu_{ik}^{jl} - \mu_i^k \mu_j^l - \mu_i^l \mu_j^k
$$

II. Cumulant statistics for DOA estimation.

A. Cumulant matrices.

Under the above assumptions, the array output $y(t)$, $1 \leq t \leq N$, is an i.i.d sequence of non-Gaussian $m$-dimensional complex random vectors with covariance given by:

$$
\mathbf{R} = \mathbf{R}(\theta) = E(y(t)y(t)^*) = \mathbf{AAR}^{H} + \mathbf{R}_b
$$

Denoting $a_i^j$ and $a_i^\gamma$ respectively the $(i, j)$-th entry of matrices $\mathbf{A}$ and $\mathbf{A}^H$ (the transpose-conjugate of $\mathbf{A}$), the array output covariance $\mathbf{R} = (r_{ij})_{1 \leq i, j \leq m}$ is equivalently expressed as

$$
r_{ij} = \sum_{\alpha, \beta = 1}^{n} a_i^\alpha a_j^\beta \gamma_{\alpha \beta} + \xi_{ij}^\gamma, \quad 1 \leq i, j \leq m
$$

where $\xi_{ij}^\gamma$ denotes the $(i, j)$-th entry of the noise covariance matrix $\mathbf{R}_b$.

The array output covariance may be thought of as the set of all 2nd-order cumulants of the random vector $y(t)$. As a natural extension, the quadracovariance is defined as the following set of 4th-order cumulants of the array output:

$$
q_{ik}^{jl} = \text{Cum}(y_i(t), y_j^*(t), y_k(t), y_l^*(t)) \quad 1 \leq i, j, k, l \leq m
$$

The quadracovariance structure is readily obtained by rewriting the model (1) for the array output $y(t)$ as

$$
y_i(t) = \sum_{\alpha = 1}^{n} a_i^\alpha s_\alpha(t) + b_i(t) \quad \text{and} \quad y_j(t) = \sum_{\beta = 1}^{n} a_j^\beta s_\beta(t) + b_j^*(t)
$$

where $b_i(t)$ is the $i$-th component of the noise vector $b(t)$. Additivity of the cumulants in the addition of independent variables [10] splits the array output quadracovariance into a signal term and a noise term according to:

$$
q_{ik}^{jl} = \sum_{\alpha = 1}^{n} a_i^\alpha a_j^\beta a_k^\gamma a_l^\delta \gamma_{\alpha \beta \gamma \delta}
$$

$$
\quad + \text{Cum}(b_i(t), b_j^*(t), b_k(t), b_l^*(t))
$$

The last term of this expression vanishes, since the noise has been assumed to be normally distributed. Using then the multilinearity of the cumulants [10] and the definition (5), it comes

$$
q_{ik}^{jl} = \sum_{\alpha, \beta, \gamma, \delta = 1}^{n} a_i^\alpha a_j^\beta a_k^\gamma a_l^\delta \gamma_{\alpha \beta \gamma \delta} + \sum_{\alpha = 1}^{n} a_i^\alpha a_j^\beta a_k^\gamma a_l^\delta \gamma_{\alpha \beta \gamma \delta}
$$

Expressions (7) and (12) are the counterpart of [9], Eqs.(3.6) - (3.7) in indexed notations. As demonstrated in [9], [11], [12], the whole set of 4th-order cumulants of the array output can be exploited to estimate the DOAs. However, these techniques involve the estimation and processing of a set of $m^4$ statistics which may be impractical even for arrays of moderate size. In the following, we focus on techniques based on $m \times m$ matrices formed from the 4th-order cumulants of the array output: the diagonal slice and the contracted quadracovariance.

The diagonal slice, $\mathbf{D} = (d_{ij})_{1 \leq i, j \leq m}$, originally proposed by Pan & Nkias [2] for DOA estimation, is obtained by selecting the following subset of the 4th-order cumulants of the array output:

$$
d_{ij} = \text{Cum}(y_i(t), y_j^*(t), y_k(t), y_l^*(t)) \quad 1 \leq i, j \leq m
$$

Plugging expression (12) in the above definition yields the following identities

$$
d_{ij} = \sum_{\alpha, \beta, \gamma, \delta = 1}^{n} a_i^\alpha a_j^\beta a_k^\gamma a_l^\delta \gamma_{\alpha \beta \gamma \delta}
$$

$$
\quad + \sum_{\alpha = 1}^{n} a_i^\alpha a_j^\beta a_k^\gamma a_l^\delta \gamma_{\alpha \beta \gamma \delta}
$$

This may be more conveniently rewritten as a product of two matrices $\mathbf{D} = \mathbf{L} \mathbf{L}^H$ where $\mathbf{L}$ is the $m \times m$ matrix defined as

$$
l_{ij}^\alpha \overset{\text{def}}{=} \sum_{\alpha, \beta, \gamma, \delta = 1}^{n} a_i^\alpha a_j^\beta a_k^\gamma a_l^\delta \gamma_{\alpha \beta \gamma \delta}
$$

$$
1 \leq i \leq m \quad 1 \leq j \leq n
$$

$^1$A complex random variable is said to be circular if its distribution is invariant under the multiplication by an arbitrary unit-modulus complex number; the same definition applies for complex random vectors.
Since $D = AL^\dagger$, the range of the diagonal slice $D$ is seen to be included in the signal subspace: this key property, obtained in [2] can be exploited to estimate the DOAs, using subspace-based methods.

The **contraction quadricovariance**, $C = (c^i_j)_{i,j \leq m}$, originally considered in [4], is obtained by summing the quadricovariance of the array output on one lower index and one upper index, according to

$$c^i_j = \sum_{k=1}^m \text{Cum}(y_i(t), y_j^k(t)) = \sum_{k=1}^m q^i_k h_j \quad 1 \leq i, j \leq m$$ \hspace{1cm} (16)

This particular summation on indices is called **contraction** in tensor terminology, hence the name of this statistic. Substituting the array output quadricovariance Eq.(12) in the above expression yields

$$c^i_j = \sum_{k=1}^m \sum_{a,\beta,\gamma=1}^n a^i a^j a^k a_\beta a_\gamma K^{a_\beta a_\gamma}$$ \hspace{1cm} (17)

$$= \sum_{a,\beta,\gamma=1}^n a^i a^j \sum_{k=1}^m a^k a_\beta a_\gamma K^{a_\beta a_\gamma}$$ \hspace{1cm} (18)

Similar to the diagonal slice, the latter expression may be expressed as a product of three matrices: $Z = AZ^\dagger$ where $Z = (z^i_j)_{i,j \leq m}$ is an $n \times n$ hermitian matrix defined as

$$z^i_j = \sum_{k=1}^m \sum_{a,\beta,\gamma=1}^n K^{a_\beta a_\gamma} a^i a^j a^k$$

The contracted quadricovariance has a structure similar to a ‘noise-free’ 2nd-order covariance, the matrix $Z$ playing the role of the covariance of the source signals (note, however, that this matrix depends on the DOAs, in contrast to the second-order case). This structure may be exploited for DOA estimation using standard subspace techniques.

### B. Sample cumulant matrices.

Fourth-order cumulant matrices are estimated from the snapshot $[y_1(1), \cdots, y_N(N)]$ in the following way. For the diagonal slice, the estimate $C_N$ is the direct sample counterpart of definition (13). Since 4th-order circular signals are considered, the 4th-order cumulants of the array output may be expressed in terms of 4th-order and 2nd-order moments as:

$$q^i_k = E(y_i(t)y_j(t)y_k(t))$$

$$E(y_i(t)y_j(t)t)E(y_k(t)y_j(t)t) - E(y_i(t)y_j(t)y_k(t))E(y_i(t)y_j(t))$$ \hspace{1cm} (19)

Using definition (13) and substituting the moments by their sample estimates, the following sequence of estimator $D_N$ for the diagonal slice is obtained

$$d^i_k = \frac{1}{N} \sum_{t=1}^N |y_i(t)|^2 y_i(t)y_j(t)^k(t) - \frac{2}{N} \sum_{t=1}^N y_i(t)y_j(t)^k(t) \sum_{t=1}^N |y_i(t)|^2$$ \hspace{1cm} (20)

where $|y_i(t)|$ denotes the modulus of $y_i(t)$. An estimate for the contracted quadricovariance can be derived along the same lines. Pinning the decomposition Eq (19) into definition (16) yields

$$c^i_j = E(y_i(t)y_j(t) \sum_{k=1}^m y_i(t)^k(t)) - \sum_{k=1}^m \sum_{k=1}^m y_i(t)^k(t)$$

$$C = E(||y(t)||^2 y(t)y(t)^*) - R \text{Tr}(R) - R^2$$ \hspace{1cm} (21)

where $||y(t)||$ stands for the norm of the vector $y(t)$, and $\text{Tr}(R)$ for the trace of the covariance matrix $R$. The contracted quadricovariance can thus be estimated as

$$c_N = \frac{1}{N} \sum_{t=1}^N ||y(t)||^2 y(t)y(t)^* - R_N \text{Tr}(R_N) - R_N^2$$ \hspace{1cm} (22)

where $R_N$ denotes the standard estimate of the 2nd-order covariance, i.e. $R_N \equiv N^{-1} \sum_{t=1}^N y(t)y(t)^*$. The diagonal slice and the contracted quadricovariance are estimated at a cost similar to the 2nd-order covariance.

### C. Parameter estimation with cumulant matrices.

If no noise is present, the observation $y(t)$ is entirely confined to the $n$-dimensional subspace spanned by the columns of the matrix $A$. Determining the DOAs from noise free observations is simply matter of finding the $n$ unique elements of $A$ that intersect this subspace. A different approach is necessary in presence of noise, since the observations are full-rank. The approach of 2nd-order as well as higher-order MUSIC is to first estimate the dominant subspace of the observations and then to find the elements of $A$ which are in some sense closest to this subspace.

As outlined above, both the diagonal slice and the contracted quadricovariance are rank-defective, and the dominant subspace corresponds to the range of these matrices. In contrast to second-order methods, this property holds true regardless of the noise spatial covariance $R_N$: consistent estimates of the DOAs can be obtained without any a priori information about the noise spatial structure, as long as the noise is Gaussian and is independent from the signals. This property appears to be the strongest motivation for using higher-order statistics in the DOA estimation problem [2, 7, 4, 9, 12].

For simplicity, we assume in the sequel that the matrices $Z$ and $LL^\dagger$, defined in Eq.(18) and Eq.(15) respectively, are full rank for all values of the parameter $\theta \in \Theta$. (this is true in particular when the signals emitted by the sources are statistically independent; see Appendix A, for a more general discussion, see [11]). In this case, the **signal subspace** corresponds to the span of the $n$ left singular vector of $D$ (respectively the $n$ eigenvectors of $C$) associated to the $n$ singular values of $D$ (respectively the $n$ eigenvalues of $C$) of largest magnitude. The orthogonal complement of the signal subspace is referred to as the **noise subspace**.

Whenever an estimate of the noise subspace has been computed from some cumulant matrix, the DOAs may be estimated by finding the vectors of $A$ that have the smallest projection onto it. In the MUSIC technique, DOAs estimates $\theta = [\theta_1, \cdots, \theta_N]^T$ are obtained as those values that minimize the so-called **null spectrum**:

$$l(\alpha) \overset{\text{def}}{=} ||\Pi\alpha||^2 = \text{Tr}(\Pi M(\alpha))$$ \hspace{1cm} (23)

where $\Pi$ is the orthogonal projector on the estimated noise subspace and where $M(\alpha)$ is the $m \times m$ **steering matrix** defined as:

$$M(\alpha) = a(\alpha)a^*(\alpha)$$ \hspace{1cm} (24)

Practical implementations of these algorithms involve the following steps:

1. Forming estimates of the cumulant matrices using either Eq.(20) or Eq.(22).
2. a. **For the diagonal slice:** Computing the singular value decomposition of $D_N$; Selecting the $m-n$ left singular vectors associated to the singular values of lowest magnitude.
2.b. For the contracted quadracovariance: Computing the eigendecomposition of $C_N$; Selecting the $m-n$ eigenvectors associated to the eigenvalues of lowest magnitude.

3. Forming the sample noise projector $\Pi_N$ from the previously selected vectors and then searching for local minima.

The procedure for estimating the sample noise subspace for the diagonal slice is equivalent to (i) computing the eigendecomposition of the hermitian matrix (referred to in the sequel as the symmetrized diagonal slice)

$$P_N = D_N D_N^H$$

and (ii) extracting the $(m-n)$ eigenvectors associated with the $(m-n)$ eigenvalues of smallest magnitudes. In the following, we consider this equivalent procedure rather than the original one, since it allows to derive the performance of the DOA estimators in an unified framework.

III. Asymptotic performance analysis

In this section, we establish the asymptotic normality and derive closed form expressions for the MUSIC DOA estimates based on the 2nd-order covariance, the symmetrized diagonal slice and the contracted quadracovariance. To this purpose, we adopt the functional approach which consists in recognizing that the whole process of constructing DOA estimates is equivalent to defining a functional relation between the DOA estimates $\theta_N$ and the sample statistics $S_N$ they are inferred from. Herein, $S_N$ refers generically to the statistics under consideration: $R_N$, $P_N$ or $C_N$. This functional dependence is denoted $\theta_N = \theta(S_N)$. It is shown below that the sequence of the sample statistics $S_N$ is asymptotically normal and that it transmits this property to the sequence of DOA estimates because the mapping $\theta$ is sufficiently regular in a neighborhood of the true value of the statistic $S = S(\theta)$. More specifically, it is established that $\theta_N$ is differentiable at $S$ so that the asymptotic covariance of $\theta_N$ is linearly related to the asymptotic covariance of $S_N$ (see below theorem 1). Obtaining the asymptotic covariance of the DOA estimates is thus just a matter of (i) defining the function $\theta$, (ii) computing its differential at point $S$ (iii) computing the asymptotic covariance of the sample statistics (iv) combining the results.

This section is devoted to introducing the definitions and theorems related to this general framework. In the next section, it is instantiated to the case of cumulant based MUSIC estimates.

Definition 1 (Covariance of a random matrix.)

Let $S = \{S_{ij}\}_{1 \leq i,j \leq m}$ be a random $m \times m$ hermitian matrix such that $E(S_{ij}^2) < +\infty$ for $1 \leq i, j \leq m$. Its covariance $\Delta$ is defined as $\Delta = \{\Delta_{ij}\}_{1 \leq i,j \leq m}$ with

$$\Delta_{ij}^{kl} = E(S_{ij}^l S_{jk}^k) - E(S_{ij}^l)E(S_{jk}^k) = \text{Cov}(S_{ij}^l, S_{jk}^k) \quad 1 \leq i, j, k, l \leq m$$

The space of all $m \times m$ hermitian matrices is denoted $\mathbb{C}^{m \times m}$. It is a normed space when equipped with the norm defined by $\|A\| = \text{Tr}(A^H A)^{1/2} = \text{Tr}(A^2)^{1/2}$. For the sake of completeness, we proceed by introducing the notion of differential of a function $f : \mathbb{C}^{m \times m} \to \mathbb{C}$, where $\mathbb{C}$ is a normed space (see for example [13]).

Definition 2 (Differential of a function of an hermitian matrix.)

Let $f$ be a function defined on an open set $U \subset \mathbb{C}^{m \times m}$ with values in a normed space $\mathbb{F}$, $f : U \to \mathbb{F}$. The function $f$ is said to be differentiable at point $S \in U$ if it exists a linear application $Df(S) : \mathbb{C}^{m \times m} \to \mathbb{F}$ such that

$$f(S + \delta S) = f(S) + Df(S) \cdot \delta S + o(\delta S)$$

Here $\bullet$ denotes the composition of the linear application $Df(S)$ with the matrix $\delta S$ and $o(\delta S)$ is a function verifying

$$\lim_{\|\delta S\| \to 0} \frac{\|o(\delta S)\|}{\|\delta S\|} = 0$$

$Df(S)$ is referred to as the differential of the function $f$ at point $S$.

In the case of real valued functions, $\mathbb{F} = \mathbb{R}$, the differential $Df(S)$ is known to be canonically associated with a $m \times m$ hermitian matrix, also denoted for simplicity by $Df(S) = (Df(S))_{ij} \in \mathbb{C}$, so that

$$Df(S) \cdot \delta S = \text{Tr}(Df(S)\delta S)$$

In practice, the matrix $Df(S)$ is often conveniently computed from a first order expansion of $f$ at point $S$ (see Eq.(32) for a particular instance).

Our derivations ultimately rely on the following theorem (directly adapted from theorem A of [14], pp. 122; see also [15], pp. 541-543 for related results).

Theorem 1: Let $S_N$ be an asymptotically normal sequence of $m \times m$ random hermitian matrices, with mean $S$ and covariance $\Delta, \text{ i.e.}$

$$\sqrt{N}(S_N - S) \rightarrow N(0, \Delta) \quad \Delta = \{\Delta_{ij}^{kl}\}_{1 \leq i,j,k,l \leq m}$$

Let $f = (f_1, \cdots, f_n)$ be a vector-valued function defined on a neighborhood $U$ of $S$, such that each component function $f_i$ has a non-zero differential at point $S$, i.e. $Df_i(S) \neq 0, 1 \leq i \leq n$. Then, $f(S_N)$ is an asymptotically normal sequence of $n$-dimensional random vectors with mean $f(S)$ and covariance matrix

$$\Sigma = \{\Sigma_{ij}\}_{1 \leq i,j \leq n}$$

given by

$$\Sigma_{ij} = \sum_{a,b,m=1}^{m} Df_a(S_N)\Delta_{ab}^{kl}Df_b(S_N)^k_l$$

IV. Asymptotic performance of MUSIC estimates

The asymptotic performance of DOA estimators based on the covariance, the contracted quadracovariance and the diagonal slice can be established along a common line because in the three cases, the DOA estimates are obtained from the sample estimates $R_N$, $C_N$ and $P_N$ according to the same procedure, namely the MUSIC algorithm.

A. The derivative of the MUSIC estimator.

The construction of MUSIC estimates is a two step procedure. First, an estimate of the noise projector is obtained from the eigendecomposition of the sample statistic. The DOA estimates are then found as the minima of the null spectrum. The definition of the function $\theta$ arises naturally from this construction. We first define, in a neighborhood $\mathcal{V}$ of the true value of the statistic $S = S(\theta)$ a matrix-valued function $\Pi$, which associates to each $S \in \mathcal{V}$ the orthogonal projector on the subspace spanned by the eigenvectors corresponding to its eigenvalues of smallest magnitudes. The function $\theta$ is then defined, in a neighborhood $\mathcal{U}$ of $S$, by a direct application of the implicit function theorem [13].

Definition of $\Pi$. The noise projector $\Pi = \Pi(\theta)$ is defined as the orthogonal projector on the eigensubspace of $S$ associated with

- $\lambda_m = \sigma$ when $S$ is chosen to be the second-order covariance (in this case only, the noise is assumed to be spatially white, i.e. $R_0 = \sigma I$).
\* \*$\lambda_m = 0$ when $S$ is either the symmetrized diagonal slice or the contracted quadricovariance.

Let $\gamma$ be a positive number such that no other eigenvalue $\lambda$ of $S$ verifies $|\lambda - \lambda_m| < \gamma$. For any hermitian $m \times m$ matrix $S \in \mathcal{E}^{m,m}$, define $P(S)$ as the subset of its eigenvalues $\lambda$ verifying $|\lambda - \lambda_m| < \gamma$. Define finally $\Pi : \mathcal{E}^{m,m} \to \mathcal{E}^{m,m}$ to be the matrix-valued function which associates to $S \in \mathcal{E}^{m,m}$ the orthogonal projector $\Pi(S)$ on the space spanned by the subset of its eigenvectors corresponding to the eigenvalues in $P(S)$.

**Theorem 2:** It exists an open neighborhood $\mathcal{V} \subset \mathcal{E}^{m,m}$ of $S$, such that the function $\Pi$ is (infinitely) differentiable on $\mathcal{V}$. The first-order differential of $\Pi$ at $S$ is given by

$$D\Pi(S) : \delta S = -\Pi \delta S \Gamma - \Gamma \delta S \Pi$$

where $\Gamma = (S - \lambda_m I)^{-1}$ is the pseudo-inverse of the matrix $(S - \lambda_m I)$, i.e., is the unique matrix verifying

$$\Gamma(S - \lambda_m I) = (S - \lambda_m I) \Gamma = I - \Pi \equiv \Pi^- \Gamma \Pi = 0$$

The values of $\Gamma$ for the three statistics under consideration are summarized in the table below.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$S$</th>
<th>$\lambda_m$</th>
<th>$\Gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd-order covariance</td>
<td>$R$</td>
<td>$\sigma$</td>
<td>$(R - \sigma I)^{-1}$</td>
</tr>
<tr>
<td>Contracted quadricovariance</td>
<td>$C$</td>
<td>$C$</td>
<td>$C^{-1}$</td>
</tr>
<tr>
<td>Symmetrized diagonal slice</td>
<td>$DD^\dagger$</td>
<td>$0$</td>
<td>$(DD^\dagger)^{-1}$</td>
</tr>
</tbody>
</table>

**Definition of $\hat{\theta}$.** The construction of the function $\hat{\theta}$ goes as follows. Since the steering matrix $M(\alpha)$ depends smoothly on the DOA $\alpha$, we assume that the mapping $\alpha \to M(\alpha) = s(\alpha)s(\alpha)^*$ is twice differentiable w.r.t. $\alpha$. We denote $\hat{M}(\alpha)$ and $\tilde{M}(\alpha)$ its first and second-order derivative. Define the function $f : \mathcal{V} \times \Theta \to \mathbb{R}$ as

$$f : (S, \alpha) \in \mathcal{V} \times \Theta \to f(S, \alpha) \equiv \text{Tr}(\Pi(S)\tilde{M}(\alpha))$$

The differentiability of both $S \to \tilde{\Pi}(S)$ (Theorem 2) and $\alpha \to M(\alpha)$ implies that the function $f(S, \alpha)$ is differentiable w.r.t. $S$ and $\alpha$ on $\mathcal{V} \times \Theta$. At $S$, the true value of the statistic, the function $\alpha \to f(S, \alpha)$ is zero whenever $\alpha$ is equal to one of the true DOAs:

$$f(S, \theta_i) = \text{Tr}(\tilde{\Pi}(S)\hat{M}(\theta_i)) = \text{Tr}(\tilde{M}(\theta_i)) = 0$$

where $M_i$ denotes the first order derivative of the steering matrix at the true DOA $\theta_i$, i.e., $M_i \equiv \hat{M}(\theta_i)$. Assume that the derivative of the function $\alpha \to f(S, \alpha)$ w.r.t. $\alpha$ does not vanish at the true value of the parameter:

$$h_i \equiv h_i(\theta) = \frac{\partial f}{\partial \alpha}(S, \theta_i) = \text{Tr}(\Pi(S)\hat{M}(\theta_i)) \neq 0, \quad 1 \leq i \leq n$$

The implicit function theorem [13] ensures that

**Theorem 3:** It exists $(i)$ an open set $\{(V_1, \ldots, V_n) : \theta_i \in V_i \subset \Theta, \quad 1 \leq i \leq n, (ii)$ an open neighborhood $\mathcal{W}$ of $S$, $S \in \mathcal{W} \subset \mathcal{V}$, (iii) a vector-valued function $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_n) : \mathcal{W} \to \Theta$, such that the relations

$$(S, \alpha) \in \mathcal{W} \times \Theta \quad \text{and} \quad f(S, \alpha) = 0, \quad 1 \leq i \leq n$$

are equivalent to

$$S \in \mathcal{W} \quad \text{and} \quad \alpha = \tilde{\theta}_i(S), \quad 1 \leq i \leq n.$$

In addition, the function $\tilde{\theta}$ is differentiable at $S$.

The differential of $\tilde{\theta}$ at point $S$, which is required in the sequel, is obtained in the following lemma

**Lemma 1:** The differential of the function $\tilde{\theta}$ at $S$, $D\tilde{\theta}(S) = (D\tilde{\theta}_1(S), \ldots, D\tilde{\theta}_n(S))$ is a collection of $n$ hermitian matrices $D\tilde{\theta}_i(S)$ given by

$$D\tilde{\theta}_i(S) = \frac{1}{n} \sum_{k=1}^{n} (\Pi M, \Gamma + \Gamma M, \Gamma) \quad 1 \leq i \leq n$$

**Proof:** For all $\delta S$ such that $S = S + \delta S \in \mathcal{W}$, we have, by definition of the functions $\tilde{\theta}_i$,

$$\text{Tr}(\Pi(S + \delta S)\hat{M}(\tilde{\theta}_i(S + \delta S))) = 0$$

The functions $S \to \tilde{\Pi}(S)$ and $S \to \tilde{\theta}_i(S)$ being differentiable at the true value of the statistic $S$, Eq (32) yields

$$h_i \text{Tr}(D\tilde{\theta}_i(S)\delta S) = \text{Tr}(\Pi M, \Gamma + \Gamma M, \Gamma)\delta S + o(\delta S) = 0$$

which establishes the result.

**Definition of the MUSIC functional.** The MUSIC estimator $\hat{\theta}_N, \quad 1 \leq i \leq n$ for the DOA $\theta_i, \quad 1 \leq i \leq n$ is defined as the minimum of the null spectrum Eq (23) belonging to the neighborhood $\mathcal{V}_i$ of $\theta_i$.

$$\hat{\theta}_N = \arg \min_{\alpha \in \mathcal{V}_i} \text{Tr}(\Pi N M(\alpha)) \quad 1 \leq i \leq n$$

By differentiating the above, this minimum also verifies

$$\text{Tr}(\Pi N \hat{M}(\hat{\theta}_N)) = 0 \quad 1 \leq i \leq n$$

Now, assume that for the particular realization under consideration, the sample statistics $S_N$ belongs to the neighborhood $\mathcal{W}$ of $S$. Since $(S_N, \hat{\theta}_N) \in \mathcal{W} \times \mathcal{V}_i$, it comes by application of theorem 3, that $\hat{\theta}_N = \hat{\theta}(S_N)$. This property holds for $1 \leq i \leq n$, showing that the MUSIC estimator $\hat{\theta}_N$ is related to the sample statistics $S_N$ by a differentiable function $\theta$ defined on a neighborhood $\mathcal{W}$ of $S$:

$$\hat{\theta}_N = \hat{\theta}(S_N) \quad \mathcal{V}_N \in \mathcal{W}$$

**B. Asymptotic covariance of sample cumulant matrices.**

The second step in our procedure consists in evaluating the covariance of the sample statistics of interest, namely the second-order covariance $R_N$, the symmetrized diagonal slice $P_N$ and the contracted quadricovariance $C_N$.

**Second-order covariance.** Since the samples $y(t), 1 \leq t \leq N$ have been assumed to be independent and identically distributed random vectors with finite moments up to order 8, the multivariate version of the central limit theorem (CLT) [14] may be applied to establish that $R_N$ is an asymptotically normal sequence of random hermitian matrices, with mean $R = R(\theta)$ and covariance $\Delta = (\Delta_{\theta N})_{1 \leq i,j,k,l \leq m}^N$.

$$\sqrt{N}(R_N - R) \to N(0, \Delta)$$

The $(i, j, k, l)$-th entry of $\Delta$ may be expressed as (see [16] for instance)

$$\Delta_{\theta N} = \lim_{N \to +\infty} N E((r_i^T - \bar{r}_i^T)(r_j^T - \bar{r}_j^T) \bar{r}_k^T \bar{r}_l^T) \quad 1 \leq i, j, k, l \leq m$$

where $R = (r_i^T)_{1 \leq i,j \leq m}$ and $Q = (q_{\theta N}^T)_{1 \leq i,j,k,l \leq m}$ denote the entries of the array output covariance and quadricovariance at the true value of the parameter $\theta$.

**Fourth-order covariance.** Computations are more intricate for the fourth-order case but the same arguments apply. Since $y(t), 1 \leq t \leq N$ is an i.i.d sequence of random vectors with
finite sample estimates up to order 8, the CLT theorem states that the sample estimate of the quadracovariance of the array output 
\[ Q_N = \sum_{i,j=1}^{N} y_i(t) y_j(t) \] 
where 
\[ y_i(t) = \frac{1}{N} \sum_{n=1}^{N} y_i(n) \] 
is an asymptotically normal sequence of random elements with mean 
\[ Q = \sum_{i,j=1}^{N} y_i(t) y_j(t) \] 
and covariance 
\[ H \] 
Here, the covariance \( H \) is a quantity indexed by 8 indices, i.e., 
\[ H = (H_{ijkl})_{i,j,k,l} \] 
whose entries are defined as

\[ H_{ijkl} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} y_i(n) y_j(n) y_k(n) y_l(n) \] 

The expansion of \( H_{ijkl} \) in terms of the cumulants of the array output (up to order 8) is given in appendix. Now, since the two 4th-order cumulant matrices under consideration (the symmetrized diagonal slice and the contracted quadracovariance) depend simply on the quadracovariance of the array output, the asymptotic covariance of their sample estimates also depend simply on \( H_{ijkl} \).

**Contracted quadracovariance.** Consider first the case where the sample statistics are the contracted quadracovariance \( S_N = C_N \). From the definition (22), the \((1, j)\)th entry of \( C_N \) is \( c_{1j} = \frac{1}{N} \sum_{n=1}^{N} y_1(n) y_j(n) \); it thus depends linearly from the sample estimate of the 4th-order cumulants of the array output. Since \( Q_N \) is an asymptotically normal sequence of estimates of the array output quadracovariance, the sequence of hermitian matrices \( C_N \) also, by linearity, asymptotically normal with mean \( C = C(\theta) \) and covariance \( \Delta = (\Delta_{ijkl})_{i,j,k,l} \) given by

\[ \Delta_{ijkl} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} y_i(n) y_j(n) y_k(n) y_l(n) \]

**Symmetrized Diagonal slice.** Consider now the case where the sample statistics are \( S_N = D_N = D_N(D)^H \). From the definition \( (13) \), the \((1, j)\)th entry of the matrix \( D_N \) is \( d_{1j} = \frac{1}{N} \sum_{n=1}^{N} y_1(n) y_j(n) \). Since the matrix \( D_N \) is a particular subset of the array output samples cumulants, it is also an asymptotically normal sequence of complex random matrices, with mean \( D = D(\theta) \); more generally, by Eq. (47), we obtain

\[ \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} y_i(n) y_j(n) y_k(n) y_l(n) = \Delta_{ijkl} \]

To extend these results to the symmetrized diagonal slice, we proceed as follows. First note that

\[ \sqrt{N} (P_N - P) = D \sqrt{N} (D_N - D)(D)_H + \sqrt{N}(D_N - D)(D)_H \]

where \( P \) denotes the true value of the symmetrized diagonal slice, i.e., \( P = D(D)_H \). Using Slutsky’s theorem in the last term in this expression converges in probability to zero and that the two sequences of hermitian random matrices \( \sqrt{N}(P_N - P) \) and \( \sqrt{N}(D_N - D)(D)_H + \sqrt{N}(D_N - D)(D)_H \) have the same limiting distribution. Since \( D_N \) is an asymptotically normal sequence of matrices with mean \( D \), it comes from this decomposition that \( P_N \) is an asymptotically normal sequence of hermitian matrices, with mean \( P \) and covariance \( \Delta = (\Delta_{ijkl})_{i,j,k,l} \), given by

\[ \Delta_{ijkl} = \sum_{n=1}^{m} F_{ijkl} + \sum_{n=1}^{m} F_{ijkl} \] 

C. The asymptotic covariance of DOA estimates.

At this point, the following properties have been established:
- The estimator \( \hat{\theta}_N \) depends on the sample statistics \( S_N \) by a differentiable function mapping \( \hat{\theta} \), defined on a neighborhood \( W \) of \( \theta \), as \( \hat{\theta}_N = \hat{\theta}(S_N) \) for \( S_N \in W \).
- At point \( S \), the function \( \theta(S) \) is equal to \( \theta \), the true value of the parameters: \( \theta = \theta(S) \).
- The sequence of sample statistics \( S_N \) is asymptotically normal with mean \( S \) and covariance \( \Delta \).

We may thus conclude by applying Theorem 1 that \( \hat{\theta}_N \) is an asymptotically normal sequence of estimators of \( \theta = \theta(S) \).

**Second-order method:** provided that the noise is spatially white: \( R_o = \sigma I \).

**Fourth-order method:** regardless of the noise spatial covariance.

The expression for the covariance of the DOA estimates is obtained by plugging in (31) the expression (38) of the differential of the mapping \( \hat{\theta} \) at \( S \) and the expression of the covariance of the sample statistics, i.e., (44) for the second-order covariance; (48) for the contracted quadracovariance; (49) for the diagonal slice. In summary, the covariance of the DOA estimates takes the following form:

\[ \Sigma_{ij} = \lim_{N \to \infty} \sum_{n=1}^{m} \sum_{\alpha, \beta=1}^{m} h_{\alpha \beta} \langle \Pi_{\alpha} \Gamma + \Gamma \Pi_{\beta} \rangle_{\alpha} \Delta_{\beta} \]

where \( \Gamma \) and \( \Delta \) given by the following table.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>( R )</th>
<th>( \Gamma )</th>
<th>( \Delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R )</td>
<td>( \sigma I )</td>
<td>( (R - \sigma I) # )</td>
<td>Eq. (44)</td>
</tr>
<tr>
<td>( C )</td>
<td>arbitrary</td>
<td>( C^{#} )</td>
<td>Eq. (48)</td>
</tr>
<tr>
<td>( DD^{H} )</td>
<td>arbitrary</td>
<td>( (DD^{H})^{#} )</td>
<td>Eq. (49)</td>
</tr>
</tbody>
</table>

V. Discussion

To get some insights into the actual performance of the estimators, we now present a work-out of these expressions under the simplifying assumption that the noise is spatially white: \( R_o = \sigma I \).

A. Second-order performance.

Regarding the second-order method, several simplifications occur due to the special structure (12) of the quadracovariance of the array output. In particular, the contribution of 4th-order terms vanishes in the composition (50) with the noise projector thanks to the following cancellations

\[ \sum_{a=1}^{m} \Pi_{a} \Pi_{a}^{\star \gamma} = 0 \] 

\[ \sum_{a=1}^{m} \Pi_{a} \Pi_{a}^{\star \gamma} = 0 \]
In addition, the orthogonality between the noise and the signal subspace implies that
\[ \Pi R = \sigma \Pi = R \Pi \] and \( \Pi \Gamma = \Gamma \Pi \), \( \Pi \Gamma = \Gamma \Pi = 0 \) (53)

Due to these multiple cancellations, the covariance of the DOA estimates takes the simpler form
\[ \Sigma_{ij} = \frac{\sigma}{h_i h_j} \left[ \text{Tr}(M \Pi M \Gamma R \Gamma') + \text{Tr}(M \Pi M \Gamma R \Gamma') \right] \] (54)

Finally, this expression can be recast into a more familiar form (as, for instance, in [17]), by splitting the derivative of the steering matrix \( \dot{M}(a) \) into \( \dot{M}(a \dot{a}(a)) + \ddot{a}(a) \cdot \dot{a}(a) \):
\[ \Sigma_{ij} = \frac{2}{h_i h_j} R(\dot{a}(\theta_j))^T \Pi \dot{a}(\theta_j) a(\theta_j)^T \dot{a}(\theta_j) \] (55)

where \( R(.) \) denotes the real part of a complex number and \( U = \sigma \Gamma R \). Since \( \Gamma = (R - \sigma \Gamma)^\# \), matrix \( U \) is more interestingly seen to be equal to:
\[ U = \sigma \Gamma R \sigma = \sigma \Gamma + \sigma^2 \Gamma^2 . \] (56)

Remark 1. Since the 4th-order cumulant term vanishes by Eqs. (51, 52), the asymptotic covariance of DOA estimates is the same whether or not the source signals are Gaussian. This property, sometimes referred to as 'asymptotic robustness' [18], [19], holds for a wide-class of DOA estimation methods [20].

Remark 2. In the limit of large SNR, the first term of the r.h.s. of (56) is dominant, so that \( a(\theta_j)^T U_a(\theta_j) \simeq \sigma a(\theta_j)^T \Gamma a(\theta_j) \). Recalling that \( \Gamma = (R - \sigma \Gamma)^\# = (AR\Gamma^H)^\# \), we find that \( a(\theta_j)^T U_a(\theta_j) \) is approximately equal to the \( (i, j) \)-th entry of matrix \( \sigma R \Gamma^2 \). In particular, for uncorrelated source signals, matrix \( R \), is diagonal with \( \sigma \), as its \( i \)-th diagonal entry so that in the large SNR limit:
\[ a(\theta_j)^T U_a(\theta_j) \simeq \sigma a(\theta_j)^T \Gamma a(\theta_j) = \frac{\sigma}{\sigma_i} k_i^2 \] (57)

We conclude that, for high enough SNR and uncorrelated sources, the MUSIC DOA estimates are uncorrelated (see [15], [17] for similar results). Moreover, the asymptotic variance of the DOA estimate of each source depends only on the signal-to-noise ratio for that particular source (the presence of other sources manifests itself only via the geometrical factor \( h_i \)); as shown below, this property does not extend to the higher-order methods herein investigated.

B. Fourth-order performance.

For the contracted quadracovariance, the derivation can be done along the same lines, let alone the larger number of terms involved in the covariance of the sample statistic. Again, many terms cancel due to the orthogonality between signal and noise subspaces. This happens whenever an index belonging to the noise projector is composed with a quantity belonging to the true signal-subspace, such as the cumulants of the array output of order greater than 2. A detailed investigation of the many terms in the expression (74) of the covariance of the array output of the contracted quadracovariance shows that only 8 terms do not vanish. After simplification, the covariance of the DOA estimator takes the form (55), the matrix \( U \) being now defined as \( U = \sigma C^F F C^\# \) with the form \( F = (f^i_j)_{1 \leq i, j \leq 4} \) given by
\[ f^i_j = \sum_{a, b = 1}^{m} h_{i\alpha} h_{j\beta} + r_{i\alpha} h_{j\beta}^* + r_{j\beta} h_{i\alpha}^* + r_{i\alpha} h_{j\beta}^* + r_{j\beta} h_{i\alpha}^* + r_{i\alpha} h_{j\beta} + r_{j\beta} h_{i\alpha} + \sigma_\alpha^2 r_{i\alpha} r_{j\beta} \] (58)

Here, \( h_{i\alpha} \) denotes the \( (i, j, k, l, m, n) \)-th entry of the array output quadracovariance:
\[ h_{i\alpha}^{im} = \text{Cum}(y_i, y_j, y_k, y_l, y_m, y_n). \] (59)

For the diagonal slice, the workaround does not lead to a better understanding. The analytical expression is not included here for brevity.

C. Single source case.

To get a better insight, we now specialize to the very simple case of a single source impinging on linear array with \( m \) sensors spaced at regular intervals of half a wavelength
\[ a(\alpha) = [1, \exp(j \pi \sin(\alpha)), \cdots, \exp(j \pi (m-1) \sin(\alpha))]^T \] (60)

We denote respectively the 3-t, the power of the source, \( k_3 \neq 0 \) its 4th-order cumulant (kurtosis) and \( k_1 \) its 6th-order cumulant.

The noise is assumed to be spatially white, i.e. \( R_0 = \sigma^2 \). The signal-to-noise ratio (SNR) is \( p_1 = k_3 / \sigma^2 \). The asymptotic variance of the DOA estimate \( \hat{\theta}_N \) for 2nd-order and 4th-order MUSIC is given by
\[ \Sigma(\theta) = \frac{1}{h_1 p_1} \left[ a_0 + \frac{a_1}{m p_1} + \frac{a_2}{(m p_1)^2} + \frac{a_3}{(m p_1)^3} \right] \] (61)

where \( h_1 = \text{Tr}((U(\hat{\theta}_3), M(\hat{\theta}_3)) \) and where the coefficients \( a_0, a_1, a_2, a_3 \) are given below for the three statistics under consideration.

<table>
<thead>
<tr>
<th>Coeff.</th>
<th>( R )</th>
<th>( C )</th>
<th>( DD^H )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_0 )</td>
<td>1 ((2 + 5 k_3 + h_1)/k_3^2 )</td>
<td>((2 + 5 k_3 + h_1)/k_3^2 )</td>
<td></td>
</tr>
<tr>
<td>( a_1 )</td>
<td>1 ((6 + 5 k_3)/k_3^2 )</td>
<td>((6 + 5 k_3)/k_3^2 )</td>
<td></td>
</tr>
<tr>
<td>( a_2 )</td>
<td>0 ((m + 6)/k_3^2 )</td>
<td>((3 m - 2)/k_3^2 )</td>
<td></td>
</tr>
<tr>
<td>( a_3 )</td>
<td>0 ((m + 2)/k_3^2 )</td>
<td>(2m(m - 1)/k_3^2 )</td>
<td></td>
</tr>
</tbody>
</table>

Here, \( k_3 \) and \( h_1 \) are the standardized 4th-order and 6th-order cumulants of the source, defined as \( k_3 = k_3 / \sigma^4 \) and \( h_1 = h_1 / \sigma^2 \).

The geometrical factor \( h_1 \) is intrinsic to MUSIC-C-type estimators; it appears to be the same for 2nd-order and 4th-order methods. The only difference lies in the coefficients of the expansion of the variance of the DOA estimator in powers of the signal-to-noise ratio.

When the SNR is 'high enough', the performance of the contracted quadracovariance and of the diagonal slice are dominated by the \( a_0 \) factor, hence appearing to be equivalent. As for 2nd-order methods, the performance in this limit is inversely proportional to the SNR. Further insights are gained by rewriting \( a_0 \) in terms of the moments of the source signal \( s_1(t) \). Since \( s_1(t) \) has been assumed to be complex circular, it comes
\[ a_0 = 1 + \frac{E[s_1(t)^4]E[s_1(t)^4] - E^2[s_1(t)^2]^2}{E[s_1(t)^4] - 2E^2[s_1(t)^2]^2} \] (62)

Applied to the complex circular random variables \( s_1(t) \) and \( s_1(t)s_1(t)^* \), the Cauchy-Schwarz inequality yields
\[ E[s_1(t)^4]E[s_1(t)^4] \geq E^2[s_1(t)^2] \] with equality when these two variables are proportional. In the circular case, this latter condition implies that \( s_1(t) \) is distributed according to
\[ s_1(t) = \begin{cases} 0 & \text{with probability } p, 0 \leq p \leq 1 \\ \lambda e^{j\theta} & \text{with probability } 1 - p \end{cases} \] (63)
where $\lambda$ is a deterministic constant and $\phi$ is random variable uniformly distributed over $[-\pi, \pi]$. As a consequence, $a_0$ is always greater than 1: at the limit of high SNR, performance of the 4th-order based MUSIC techniques under consideration are bounded from below by performance of 2nd-order MUSIC. But, for the particular class of source signals distributed according to (63), the 4th-order techniques under consideration achieve the same performance as 2nd-order MUSIC in the high SNR limit.

We now turn to the case where noise is dominant. For 2nd-order MUSIC, expression (61) shows a "threshold effect". At the limit of high SNR, the performance varies essentially as $1/\ln(\rho_1)$; at the limit of low SNR, the performance is dominated by the second term $1/m\ln(\rho_1)$. The transition between these two distinct regions occurs at the SNR value where the two contributions are equal, i.e., $1/\ln(\rho_1) = 1/m\ln(\rho_1)$. For the ULA linear array under consideration, the threshold is thus given by

$$\rho_{th} = \frac{1}{m} \tag{64}$$

The threshold effect is stronger for 4th-order MUSIC, since the variance of the DOA estimate is dominated, in the low SNR limit, by the fourth power of the noise-to-signal ratio. Defining now the threshold $\rho_{th}$ as the value of the SNR for which the contribution of the first term (high-SNR limit) becomes equivalent to the contribution of the last term (low-SNR limit), we get

Cont. quadricov.: $\rho_{th} = \frac{1}{m} \left( \frac{m + 2}{2 + 5k + k_1} \right)^{\frac{1}{3}}$

Sym. diagonal slice: $\rho_{th} = \frac{1}{m} \left( \frac{2m(m - 1)}{2 + 5k + k_1} \right)^{\frac{1}{3}}$

It follows that the threshold for the contracted quadrivariate is significantly lower than the threshold for the diagonal slice, especially for large array size.

In contrast to the high-SNR limit, where large source kurtosis does not improve the performance, large values of $k_1$ affects favorably the variance of the DOA estimator at the limit of low SNR; it is indeed proportional to the inverse of the squared standardized source 4th-order kurtosis. Since the value of the standardized 4th-order kurtosis are not bounded, the DOA covariance may be arbitrarily low [21]. This would be for example the case for an impulsive source; nonetheless, one should be aware that in such MUSIC technique is likely to be very sub-optimal.

D. Multiple sources, high SNR case.

We now address the case of $n$ statistically independent sources impinging on the ULA array defined in Eq.(60). We more specifically discuss the influence of the other sources on the variance of a particular DOA estimate. For the sake of simplicity, we only deal with the high SNR case: only the term of the first order term in the noise power $\sigma$ is evaluated. As previously, $\sigma$, $k$, and $h_i$ respectively denote the variance, the 4th-order and the 6th-order cumulant of the $i$-th source ($1 \leq i \leq n$). $\rho_i$ denote the signal to noise ratio for the $i$-th source, $\rho_i = \sigma_i/\sigma$ and $k_i$ and $h_i$ refer to the standardized 4th-order and 6th-order cumulants.

For 2nd-order MUSIC, the first order term in the asymptotic variance of a particular DOA estimate is (see Eq.(57)):

$$\Sigma_{ii} = \lim_{N \rightarrow +\infty} NE(\hat{\theta}_{i,N} - \theta_i)^2 \approx \frac{1}{\ln(\rho_i)} \quad 1 \leq i \leq n \tag{65}$$

This is just the expression obtained for a single source: influence of other sources manifests itself only in the geometrical factor $\ln(\rho_i)$. Thanks to Eq.(57), one may also assert that the DOA estimates for two distinct DOAs $\theta_i,N$ and $\theta_j,N$ are asymptotically uncorrelated.

The behavior of 4th-order methods is significantly different. Denote $\gamma_{ij}$ the square cosine of the directional vectors $a(\theta_i)$ and $a(\theta_j)$: $\gamma_{ij} = |a(\theta_i) a(\theta_j)^T| / \sigma_i$. For the contracted quadrivariate, the asymptotic variance of the DOA estimator $\theta_{i,N}$, $1 \leq i \leq n$ may be expressed as:

$$\Sigma_{ii} = \lim_{N \rightarrow +\infty} NE(\hat{\theta}_{i,N} - \theta_i)^2 = \frac{1}{\ln(\rho_i)} (a_0 + b_0 + c_0 + d_0) \tag{66}$$

where the coefficients $a_0$, $b_0$, $c_0$ and $d_0$ are respectively given by

$$a_0 = \frac{k_i^2}{2} (2 + 5k_1 + h_1) \tag{67}$$

$$b_0 = \frac{k_i^2}{2} (3 + 2k_1) \sum \rho_j \gamma_{ij} \tag{68}$$

$$c_0 = \frac{k_i^2}{2} \sum \rho_j \gamma_{ij} \tag{69}$$

$$d_0 = \frac{m^2}{k_i^2} \sum \rho_j \rho_q \gamma_{pq} \tag{70}$$

Note that the first term $a_0$ corresponds to the variance of the DOA estimator when only one source emitting at DOA $\theta_i$ is present. A simple inspection of the other terms shows that they all are positive. Some of these are likely to be predominate, especially if the variance of a source at $\theta_p$ ($p \neq i$) is significantly larger than $\sigma_i$. Hence, the variance of the DOA estimate of a weak source is severely affected by the presence of strong sources; as stressed above, this is not the case for 2nd-order MUSIC. Note finally that the terms $b_0$, $c_0$, $d_0$ go to zero when the standardised kurtosis $k_i$ of the i-th source goes to infinity, which means that the effect of strong interferers may be, to some extent, alleviated if the kurtosis of the source of interest is large enough.

The asymptotic covariance of estimates $\hat{\theta}_i,N$ and $\hat{\theta}_j,N$ associated with two distinct DOAs $\theta_i \neq \theta_j$ is given by

$$\Sigma_{ij} = \frac{2}{\ln(\rho_j)} \Re \{a(\theta_i)^T \Pi_a(\theta_j) \times \left[ \frac{\sigma_j}{\sigma_i^2 m k_i} a(\theta_i) a(\theta_j)^T + \frac{1}{m^2 \sigma_i \sigma_j k_i k_j} a(\theta_i)^T R_s a(\theta_j) \right] \} \tag{71}$$

In contrast to 2nd-order MUSIC, the estimates for two distinct DOAs are not asymptotically uncorrelated: this may appear as a severe drawback of these methods, especially when (i) the sources are close in DOA space and (ii) the dynamic range of the sources is large. Here, again, the correlation of DOA estimates is inversely proportional to the standardised kurtosis of the sources: the DOA estimates appear uncorrelated for sources with very large kurtosis.

We have only discussed here the case of the contracted quadrivariate, because it is more simply analytically handled. Numerical evaluations of the covariance of the DOA for the diagonal slice show that the performance of the DOA estimator of a weak source is even more severely affected by the presence of strong sources (numerical examples are presented in [5]).

VI. Conclusion

This contribution addressed the problem of DOA estimation in the narrow-band array processing context using the 4th-order
cumulants of the array output. We concerned ourselves mainly with two matrix-valued statistics: the diagonal slice and the contracted quadricovariance, which can be estimated at a cost similar to the 2nd-order covariance. Fourth-order techniques yield consistent estimation of the direction-of-arrivals without any a priori information about the noise spatial covariance, as long as the noise is Gaussian and independent from the source signals.

The functional approach was used to provide a common unifying framework in deriving closed form expressions for asymptotic performance of DOA estimates. Based on these expressions, a detailed comparison of 2nd-order and 4th-order techniques was undertaken in some simple scenarios: a single source/multiple independent sources impinging on a uniform linear array in spatially white noise.

The influence of the SNR on the variance of the estimator was precisely quantified in the single source case. It appeared that 2nd-order and 4th-order techniques show similar performance at the limit of high SNRs. Beyond a certain threshold (typically related to the number of sensors in the ULA case), the variance of the estimates grows steeply with the noise level for 4th-order methods: the expansion of the variance in the inverse of the SNR shows terms up to the power of four for 4th-order methods while only a power of two entails the performance of 2nd-order method.

The study of the multiple independent sources case (under the assumption of large SNRs for all the sources) has evidenced, for 4th-order methods, a significant increase in the variance of the DOA estimate of a weak source in presence of stronger sources, a phenomenon which does not occur for 2nd-order MUSIC. This increase seems to prohibit the use of 4th-order method, even for large SNR, when the dynamic range of the sources is important.

Finally, the comparison of the asymptotic performance of DOA for 4th-order cumulant based statistics shows that the contracted quadricovariance outperforms the diagonal slice in many respects: it shows lower estimation variance and significantly so for low SNR or for non equipowered sources.

Appendix

I. Independent Sources

Since both $\mathbf{Z}$ and $\mathbf{LL}^H$ depend on $\mathbf{A}$ and on the joint 4th-order cumulants of the signal emitted by the sources, it is difficult to state general conditions for these matrices to be full rank. This property can however be asserted for certain simple scenarios, for instance when the signals emitted by the sources are statistically independent.

In these circumstances, the cross-cumulants of the sources vanish, so that $k_{ij}^{l}$ boils down to

$$k_{ik}^{lj} = k_{kl}^{ij} \quad 1 \leq i, j, k, l \leq n$$

where $\delta_{ik}$ is the 4th-order Kronecker $\delta$ symbol (i.e., $\delta_{ik}$ is equal to 1 if all the indices are equal $i = j = k = l$ and is zero otherwise), and $k_{ij}$ is the 4th-order cumulant (the kurtosis) of the $i$-th source, $k_{0} = \text{Cum}(s, s, s, s)$. Pinnin this expression in the definition Eq. (18), one may show that the matrix $\mathbf{Z} = (z_{ij})_{1 \leq i, j \leq n}$ becomes diagonal, and that the $i$-th diagonal elements is given by

$$z_{ii} = \sum_{k=1}^{m} \sum_{a=1}^{m} k_{ik}^{a} a_{a} \sigma_{a}^2 = k_{i} \sum_{a=1}^{m} a_{i}^2 \sigma_{a}^2$$

Provided that the 4th-order cumulants of the sources do not vanish, the matrix $\mathbf{Z}$ is thus full rank.

For the diagonal slice, a similar property can also be derived under the additional condition that all the sensors have identical directivity pattern: $a_i(a) = a(a)$, $1 \leq i \leq m$. In these circumstances, it is easily found that the diagonal slice $\mathbf{D}$ and the contracted quadricovariance $C$ are proportional. For an array of omnidirectional sensors with unit gain for each sensor $(|g(a)| = 1)$, the proportionality factor is equal to $m$: $\mathbf{D} = m \mathbf{C}$.

II. Covariance of the sample cumulants

The covariance of the sample quadricovariance of the array output may be expressed as a sum of homogeneous products of cumulants up to order $8$. The list of all possible such terms can easily be adapted from [16], pp 259-260; even for zero-mean variables, it shows an impressive 569 terms (or partitions). Fortunately, since we deal only with circularly distributed random variables, many of these terms vanish. The remaining 50 terms sum up as:

$$\lim_{N \to \infty} \frac{N}{E}((q_{ik}^{l})_t - (q_{ik}^{l})_0)\sum_{j=1}^{n} (g_{ij} - g_{ij}) = o_{ik}^{l} +$$

$$h_{ik}^{l} r_{i}^l + h_{ik}^{l} r_{k}^l + h_{ik}^{l} r_{l}^l + k_{ik}^{l} r_{i}^l + h_{ik}^{l} r_{k}^l + h_{ik}^{l} r_{l}^l + i_k^{l} r_{i}^l + i_k^{l} r_{k}^l + i_k^{l} r_{l}^l + q_{ik}^{l} r_{i}^l + q_{ik}^{l} r_{k}^l + q_{ik}^{l} r_{l}^l + \alpha_{ik}^{l} r_{i}^l + \alpha_{ik}^{l} r_{k}^l + \alpha_{ik}^{l} r_{l}^l$$

where $h_{ik}^{l}$ and $o_{ik}^{l}$ denote, generalizing definition (10), the 6th-order and 8th-order cumulants of the array output:

$$h_{ik}^{l} = \text{Cum}(y, y', y, y', y', y')$$

$$o_{ik}^{l} = \text{Cum}(y, y', y, y', y', y', y')$$

References

Jean-François Cardoso was born in Tunis on March 1, 1958. He received the Agrégation de Physique degree from the École Normale Supérieure de Saint-Cloud in 1981 and the Doctorat de Physique degree from the University of Paris in 1984. He is currently with the CNRS and works at the "Signal" department of Télécom Paris (former ENST). He is a member of the board of the GDR TOSI (the CNRS research group on signal processing) where he is in charge of the working group on higher order statistics. His research interests are in statistical signal processing, with emphasis on (blind) array processing and performance analysis.

Éric Moulines was born in Bordeaux, France, in 1963. He received his M.S. degree from École Polytechnique in 1984 and his PhD degree from Télécom-Paris in 1990 in Signal Processing. From 1984 until 1990, he was a member of the technical staff at CNRS, working on signal processing applied to low-bit rate speech coding and text-to-speech synthesis. Since 1990, he has been an assistant professor at Télécom-Paris, in the Signal Processing Department. His research interests are in statistical signal processing and include parametric time-series analysis, narrow-band array processing, system identification and speech processing. He is on the editorial board of Speech Communication.