

If it is supposed that $g_{11} = 0$ (for example), then we will find:

$$g_{11} = 0 \Rightarrow \begin{cases} g_{22} = 0 \\ g_{12} = 0 \\ m_{12} = m_{22} = 0. \end{cases} \quad (21)$$

The last solution in (21) gives a constraint for the coefficient of mixture matrix: from the model (1), the constraint is equivalent to the case where there is only one source. This special situation can be easily detected from the cumulant mixture signals properties. If it is assumed that the other coefficients of global matrix are equal to zero, then there is a similar result. If it is supposed now that (g_{ij}) and (m_{ij}) are different from zero, then the equation of the global maximum (19) is found.

REFERENCES

- [1] R. L. L. Tong and Y. H. V. C. Soon, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 499–509, May 1991.
- [2] D. T. Pham and P. Garat, "Séparation aveugle de sources temporellement corrélées," in *XIV Colloque GRETSI*, Juan-Les Pins, France, Sept. 1993, pp. 317–320.
- [3] A. Swami, G. Giannakis, and S. Shamsunder, "A unified approach to modeling multichannel ARMA processes using cumulants," *IEEE Trans. Signal Processing*, pp. 898–913, Apr. 1994.
- [4] Y. Inouye and T. Umeda, "Parameter estimation multichannel ARMA processes using fourth-order cumulants," in *Proc. Int. Signal Processing Workshop Higher Order Statistics*, Chamrousse, France, July 1991, pp. 249–252.
- [5] J. Héroult, C. Jutten, and B. Ans, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé," in *Actes du Xème colloque GRETSI*, Nice, France, May 20–24, 1985, pp. 1017–1022.
- [6] P. Comon, C. Jutten, and J. Héroult, "Blind separation of sources, part II: Statement problem," *Signal Processing*, vol. 24, no. 1, pp. 11–20, Nov. 1991.
- [7] J.-C. Fort, "Stabilité de L'algorithme de séparation de sources de Jutten et Héroult," *Traitement du signal*, vol. 8, no. 1, pp. 35–42, 1991.
- [8] J.-L. Lacoume and P. Ruiz, "Sources identification: A solution based on cumulants," in *IEEE ASSP Workshop V*, Minneapolis, MN, Aug. 1988.
- [9] P. Comon, "Separation of sources using higher-order cumulants," in *SPIE Vol. 1152 Advanced Algorithms Architectures Signal Processing IV*, San Diego, CA, Aug. 8–10, 1989.
- [10] J.-F. Cardoso, "Source separation using higher order moments," in *Proc. ICASSP*, Glasgow, Scotland, May 1989, pp. 2109–2212.
- [11] P. Comon, "Independent component analysis, A new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287–314, Apr. 1994.
- [12] E. Moreau and O. Macchi, "New self-adaptive algorithms for source separation based on contrast functions," in *IEEE Signal Processing Workshop Higher-Order Statistics*, South Lake Tahoe, CA, June 1993, pp. 215–219.
- [13] H. L. Nguyen Thi, "Séparation aveugle de sources à large Bande dans un mélange convolutif," Ph.D. thesis, INP, Grenoble, Jan. 1993.
- [14] X. Oliva Galvan, "Blind separation of sources: Some adaptive algorithms," Tech. Rep., Lab. CÉPHAG, Grenoble, Apr. 1993.
- [15] P.-O. Amblard and J.-M. Brossier, "Adaptive estimation of the fourth-order cumulant of an i.i.d. stochastic process," *IEEE Trans. Signal Processing*, vol. 42, no. 1, pp. 37–43, 1995.

Optimal Dimension Reduction for Array Processing—Generalized

Sören Anderson and Arye Nehorai

Abstract—This correspondence extends previously reported work on the problem, or rather possibility, of achieving optimality of beamspace (BS) array processing parameter estimators, where use is made of dimensionally reduced data vectors. The optimality here is with respect to the best possible element space (ESP) parameter estimation accuracy, i.e., the Cramér-Rao bound, and is valid for multiple parameters per emitter.

I. INTRODUCTION

The topic of sensor array signal processing deals with methods for processing measurements of an array of sensors. Some typical applications are radio telescopes, where signals emitted by radio sources are measured by means of antenna groups and passive sonar, where hydrophone arrays are used for collecting data. The problem in these applications is to detect and estimate incoming signals in order to determine, for example, the locations of the emitting sources.

The interest in performance improvement for model parameter estimation algorithms can result in antenna arrays that are composed of a large number of sensor elements. Since the computational requirements are directly affected by the dimension of the collected data, the burden increases rapidly with the number of sensors. This is the case, e.g., for the maximum likelihood estimator as well as for the more recently developed high-resolution methods based on eigendecomposition of the array covariance matrix. These latter methods, which by a "direct" approach (using the nonreduced or element space (ESP) data vectors), require a number of operations proportional to the cube of the dimension of the ESP data vectors [6]. Therefore, in order to secure efficient computation of the algorithms, it is useful to introduce some sort of mapping that reduces the dimension of the data set before applying signal processing algorithms. Another advantage is that if the transformation matrix is implemented using analog technology, there is a need for a smaller number of analog-to-digital converters, as compared with the number needed for ESP computations. The space into which the set of full dimensional data is mapped is referred to as the (reduced dimension) beamspace (BS).

One approach to reduce the amount of computations is to employ a linear (matrix) transformation for mapping the full-dimension ESP data into the lower dimensional BS and then apply a signal processing algorithm to this new set of data. The design of the matrix transformation is guided by such subjective criterion as, for instance, selection of spatial sector [3]–[5], maximization of average signal-to-noise ratio within a specified sector [12], [13], or minimization of output interference power (assuming *a priori* knowledge of the interference scenario) [11]. In [7], the design is made by analytic determination of a transformation matrix through minimization of the

Manuscript received December 31, 1993; revised January 24, 1995. This work was supported by the Air Force Office of Scientific Research under Grant no. F49620-93-1-0096, the Office of Naval Research under Grant no. N00014-91-J-1298, and the National Science Foundation under Grant no. MIP-9122753. The associate editor coordinating the review of this paper and approving it for publication was Prof. Douglas Williams.

S. Anderson is with Ericsson Radio Systems AB, Stockholm, Sweden, on leave from the Department of Electrical Engineering, Linköping University, Linköping, Sweden.

A. Nehorai is with the Department of Electrical Engineering, Yale University, New Haven, CT 06520 USA.

IEEE Log Number 9412608.

BS signal-to-noise ratio (SNR) at which two closely spaced emitters can be resolved. The derivation is carried out using the framework of the MUSIC algorithm [10].

The main purpose of this correspondence is to generalize the previously reported analysis [1], [2] under circumstances by which it is possible to retain the property of asymptotically achieving the full-dimension Cramér-Rao bound (CRB) while reducing the dimension of the data set used for the estimation of the underlying model parameters. As shown in [1] and [2], it is possible to obtain analytical expressions for the requirements to be met by an optimal transformation matrix. These expressions rely on knowledge of the functional form of the array (calibrated array of arbitrary geometry) and the true values on the parameters of interest. This last condition does not, however, limit the possibilities of using the results derived to outline a design method.

The generalization considered herein consists of the employment of a more general data model and the corresponding CRB. This model was used in [9] for the modeling of *electromagnetic* wavefronts by means of vector sensors. Therein, the derivation of the CRB for this more general model can be found. In [8], the same model is used to model *acoustic* vector sensor measurements.

The main difference between the model considered here and the one employed in [1] and [2] is that we now can allow for multiple *parameters* per emitter as well as for vector sensor data.

II. THE DATA MODEL AND ITS CRB

In this section, the employed measurement model is briefly reviewed. Moreover, the corresponding CRB expression, as derived in [9], is also given.

A Generalized Measurement Model

The basic (narrowband) data model we will consider here is of the following form, where we assume that the number of vector sensors is m , and the number of sources is n . The employed *vector* sensor is a device that is able to deliver, at its output, multiple measurements of different wave phenomena (as in [9], for example), the complete electric and magnetic field components at the sensor, or (as in [8]) the acoustic pressure and particle velocity vector. The model is the following:

$$\mathbf{y}(t) = \mathbf{A}(\theta)\mathbf{x}(t) + \mathbf{e}(t) \quad (1)$$

where $\mathbf{y}(t) \in \mathbb{C}^{\bar{m} \times 1}$ is the observed array output vector at time instant t , whereas $\mathbf{e}(t) \in \mathbb{C}^{\bar{m} \times 1}$ is an additive measurement noise vector. Assuming that the n emitters each generate unknown vectors of wave fronts $\mathbf{x}^{(k)}(t) \in \mathbb{C}^{\nu_k \times 1}$, $k = 1, \dots, n$, where $\mathbf{x}^{(k)}(t)$ is the vector signal of the k th source, we have $\mathbf{x}(t) \in \mathbb{C}^{\bar{\nu} \times 1}$, where $\bar{\nu} = \sum_{k=1}^n \nu_k$. We thus obtain the vector $\mathbf{x}(t)$ by stacking the $\mathbf{x}^{(k)}(t)$'s on top of each other; see (5). The transfer matrix $\mathbf{A}(\theta) \in \mathbb{C}^{\bar{m} \times \bar{\nu}}$ and the parameter vector $\theta \in \mathbb{R}^{\bar{q} \times 1}$ are given by

$$\mathbf{A}(\theta) = \left[\mathbf{A}_1(\theta^{(1)}), \dots, \mathbf{A}_n(\theta^{(n)}) \right] \quad (2)$$

$$\theta = \left[\left(\theta^{(1)} \right)^T, \dots, \left(\theta^{(n)} \right)^T \right]^T \quad (3)$$

with $\mathbf{A}_k(\theta^{(k)}) \in \mathbb{C}^{\bar{m} \times \nu_k}$ and the parameter vector for the k th source $\theta^{(k)} \in \mathbb{R}^{q_k \times 1}$ (thus, $\bar{q} = \sum_{k=1}^n q_k$). We also define

$$\mathbf{y}(t) = \left[\left(\mathbf{y}^{(1)}(t) \right)^T, \dots, \left(\mathbf{y}^{(m)}(t) \right)^T \right]^T \quad (4)$$

$$\mathbf{x}(t) = \left[\left(\mathbf{x}^{(1)}(t) \right)^T, \dots, \left(\mathbf{x}^{(n)}(t) \right)^T \right]^T \quad (5)$$

where $\mathbf{y}^{(j)}(t) \in \mathbb{C}^{\mu_j \times 1}$ is the vector measurement of the j th sensor (hence, $\bar{\mu} = \sum_{j=1}^m \mu_j$).

The CRB

Consider the model (1) and the problem of estimating the parameter vector θ of that model, with θ , $\mathbf{P} = \mathbb{E}[\mathbf{x}(t)\mathbf{x}(t)^T]$, and $\sigma^2 \mathbf{I} = \mathbb{E}[\mathbf{e}(t)\mathbf{e}(t)^T]$ unknown. Observe that the noise model actually can be even less restrictive than in the present case without changing the CRB expression; see [9] for those details. From [9], we also have the following theorem (where N is the number of array output measurements or *snapshots*): The noise $\mathbf{e}(t)$ as well as the source signal $\mathbf{x}(t)$ are both assumed to be temporally uncorrelated stationary (complex) Gaussian processes. See [8], [9] for further assumptions.

Theorem 1: The CRB on the estimation error covariance matrix of any (locally) unbiased estimator of the vector θ in the model (1) with θ , \mathbf{P} , σ^2 unknown and $\nu_k = \nu$ for all k is

$$\text{CRB}(\theta) = \frac{\sigma^2}{2N} \left\{ \text{Re} \left[\text{btr} \left((\mathbf{1} \boxtimes \mathbf{U}) \boxtimes (\mathbf{D}^* \mathbf{I}_c \mathbf{D})^{bT} \right) \right] \right\}^{-1} \quad (6)$$

where

$$\mathbf{U} = \mathbf{P}(\mathbf{A}^* \mathbf{A} \mathbf{P} + \sigma^2 \mathbf{I})^{-1} \mathbf{A}^* \mathbf{A} \mathbf{P} \quad (7)$$

$$\mathbf{I}_c = \mathbf{I} - \mathbf{I} \quad (8)$$

$$\mathbf{I} = \mathbf{A}(\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^* \quad (9)$$

$$\mathbf{D} = \left[\mathbf{D}_1^{(1)}, \dots, \mathbf{D}_{q_1}^{(1)}, \dots, \mathbf{D}_1^{(n)}, \dots, \mathbf{D}_{q_n}^{(n)} \right] \quad (10)$$

$$\mathbf{D}_i^{(k)} = \frac{\partial \mathbf{A}_k}{\partial \theta_i^{(k)}} \quad (11)$$

and where $\mathbf{1}$ denotes a $\bar{q} \times \bar{q}$ matrix with all entries equal to one, and the block trace operator $\text{btr}(\cdot)$, the block Kronecker matrix product \boxtimes the block Schur-Hadamard product \boxplus and the block transpose operator bT are as defined in [9] with blocks of dimensions $\nu \times \nu$, except for the matrix $\mathbf{1}$ that has blocks of dimensions $q_i \times q_j$. Furthermore, the CRB remains the same independent of whether σ^2 is known or unknown. \square

If $\nu_i \neq \nu_j$ for at least one pair (i, j) , we shall instead consider the expression

$$[F_0'']_{\langle ij \rangle l p} = \frac{2}{\sigma^2} \text{Re} \left\{ \text{Tr} \left[\mathbf{U}_{\langle ij \rangle} \left(\mathbf{D}_p^{(i)} \right)^* \mathbf{I}_c \left(\mathbf{D}_l^{(j)} \right) \right] \right\} \quad (12)$$

where $[F_0'']_{\langle ij \rangle l p}$ is the (l, p) th entry of the matrix $[F_0'']_{\langle ij \rangle} \in \mathbb{C}^{q_i \times q_j}$, and $[F_0'']_{\langle ij \rangle}$, in turn, denotes the (i, j) th block entry of the Fisher information matrix

$$F_0''(\theta) = \frac{1}{N} \text{CRB}^{-1}(\theta). \quad (13)$$

Similarly, $\mathbf{U}_{\langle ij \rangle}$ is the (i, j) th block entry of the matrix \mathbf{U} defined in (7).

Observe here that the matrix $\mathbf{D}_i^{(i)} \in \mathbb{C}^{\bar{\mu} \times \nu_i}$, that is, it is the $\left(\sum_{k=1}^{i-1} q_k + i \right)$ th block column (composed of ν_i columns) of \mathbf{D} ; see [9] for more details.

III. OPTIMAL TRANSFORMATION MATRIX

In order to obtain the beamspace (BS) observations, we first define a transformation matrix $\mathbf{T} \in \mathbb{C}^{\bar{m} \times r}$. The observed data in BS are then related to the element space (ESP) measurements according to $\mathbf{z}(t) = \mathbf{T}^* \mathbf{y}(t) \in \mathbb{C}^{r \times 1}$. If $r = 1$, this is more commonly referred to as *beamforming* in the array processing literature.

The BS data model can now be written, with θ_0 denoting the *true* value of the parameter vector θ , as

$$\mathbf{z}(t) = \mathbf{T}^* \mathbf{A}(\theta_0) \mathbf{x}(t) + \mathbf{T}^* \mathbf{e}(t) \quad (14)$$

where we also shall require that the transformation matrix satisfies $\mathbf{T}^* \mathbf{T} = \mathbf{I}_r$, that is, it shall have orthonormal columns. This requirement is for guaranteeing that the BS noise $\mathbf{T}^* \mathbf{e}(t)$ is white whenever the ESP noise $\mathbf{e}(t)$ is also white.

In order to assess the optimal performance for this new model, the only thing we have to do is to replace A and D in Theorem 1 (or in (12)) with T^*A and T^*D , respectively. If we denote by $\text{CRB}_{\text{ESP}}(\theta_0)$ the ESP CRB and the corresponding BS bound by $\text{CRB}_{\text{BS}}(\theta_0, T)$, we find the BS CRB expression to be

$$\text{CRB}_{\text{BS}}(\theta_0, T) = \frac{\sigma^2}{2N} \left\{ \text{Re} \left[\text{btr} \left((1 \otimes U_{\text{BS}}) \square (D^* T [H_c]_{\text{BS}} T^* D)^{bT} \right) \right] \right\}^{-1} \quad (15)$$

where we now have

$$U_{\text{BS}} = P(A^* T T^* A P + \sigma^2 I)^{-1} A^* T T^* A P \quad (16)$$

$$[H_c]_{\text{BS}} = I - T^* A (A^* T T^* A)^{-1} A^* T. \quad (17)$$

Provided that there exist *efficient* estimation algorithms, that is algorithms that attain the CRB, the main objective would be to find a transformation matrix T that makes the ESP and BS bounds coincide for the true value on the parameter vector, i.e., $\text{CRB}_{\text{BS}}(\theta_0, T) = \text{CRB}_{\text{ESP}}(\theta_0)$. This would then guarantee the best possible performance of any efficient algorithm applied to BS data.

Now, let $\dim\{T\}$ denote the number of columns, or the dimension, of the matrix T . Recall also that ν_k is the number of components of the vector signal $x^{(k)}(t)$ of the k th source. We have the following theorem.

Theorem 2: Suppose that $\nu_k = \nu$, $k = 1, \dots, n$. Then, the equality

$$\text{CRB}_{\text{BS}}(\theta_0, T) = \text{CRB}_{\text{ESP}}(\theta_0) \quad (18)$$

holds true, provided that

$$T T^* [A(\theta_0), D(\theta_0)] = [A(\theta_0), D(\theta_0)] \quad (19)$$

and $\dim\{T\} \geq \nu(n + \sum_{k=1}^n q_k)$.

Proof: The theorem follows directly by making use of the assumption (19) in (16) and (17) since (19) implies that

$$U_{\text{BS}} = U \quad (20)$$

$$[H_c]_{\text{BS}} = T^* [H_c] T. \quad (21)$$

The equality (21) follows if use is made of the requirement that the columns of T shall be orthonormal $T T^* = I$. A comparison with Theorem 1 now shows that the BS CRB coincides with the ESP one. Note that the number of columns in T is allowed to be larger than $\nu(n + \sum_{k=1}^n q_k)$, as long as (19) is satisfied. This is also the lower bound on the number of sufficient conditions on T that guarantee that (19) can be fulfilled when $[A(\theta_0), D(\theta_0)]$ has full rank. \square

The final notes in the proof above were utilized in [1] and [2] for devising a method for design of transformation matrices, using a less general model than (1). The argument is based on the fact that the transformation matrix may have more columns than the lower bound indicated in the theorem above. This means that instead of knowing *exactly* the true value of the parameter vector, we can design a transformation matrix satisfying the assumption (19) in some parameter intervals, "hopefully" containing also the true parameter vector value.

Corollary 1: Suppose now that $\nu_i \neq \nu_j$ for at least one pair (i, j) . Then, if $\dim\{T\} \geq \sum_{k=1}^n (1 + q_k) \nu_k$ and

$$T T^* [A(\theta_0), D(\theta_0)] = [A(\theta_0), D(\theta_0)] \quad (22)$$

we have

$$\text{CRB}_{\text{BS}}(\theta_0, T)_{\langle ij \rangle lp} = \text{CRB}_{\text{ESP}}(\theta_0)_{\langle ij \rangle lp} \quad (23)$$

for the (l, p) th position of the (i, j) th block entry of the ESP CRB matrix $\text{CRB}_{\text{ESP}}(\theta_0)$.

Proof: The proof follows immediately from (12) and the assumption (19) together with the proof of Theorem 2. Observe that

the number of columns are $\bar{\nu} = \sum_{k=1}^n \nu_k$ for $A(\theta)$ and $\sum_{k=1}^n \nu_k q_k$ for $D(\theta)$. \square

Theorem 2 and its corollary generalize the result in [1], [2], where it was assumed that both $q_k = 1$ and $\nu_k = 1$, $k = 1, \dots, n$ (that is, only one signal and one signal parameter, the azimuth angle, per source). In those references, a method was proposed for design of a transformation matrix satisfying assumptions similar to those stated herein. Some "rough" prior information about the true value of the parameter vector θ is necessary, however. That design method can, fairly easily, also be applied to this generalized model (with some obvious changes). The only thing needed is specified intervals for all different physical parameters (such as azimuth and elevation, etc.), within which those parameters, by the assumed prior knowledge, can be assumed to be located.

Finally, observe that the bounds on $\dim\{T\}$, which is the number of columns in T , stated in Theorem 2 and its corollary are indeed *lower* bounds, provided that the matrix $[A(\theta_0) D(\theta_0)]$ has *full rank*.

IV. SUMMARY

This correspondence extends previously reported work on the possibility of optimality in beamspace array signal processing. The presented extension is in terms of a more general measurement model. The main result is the statement of sufficient conditions for being able to have the *same* model parameter estimation accuracy in beamspace as that possible in element space. These conditions include bounds on the dimension of the beam space, above which it is always possible to attain the element space CRB also in beam space.

The implication of this work is that, provided there exist *efficient* parameter estimation algorithms for a model that fits into the employed model structure, any of these algorithms can be used for element space efficient parameter estimation using the reduced-dimension beam space data. The reduction of the data dimension imply a reduction in computational requirements for many popular estimation methods.

REFERENCES

- [1] S. Anderson, "Optimal dimension reduction for sensor array signal processing," in *Proc. 25th Ann. Asilomar Conf. Signals, Syst. Comput.*, Monterey, CA, Nov. 4-6, 1991, pp. 918-922.
- [2] —, "On optimal dimension reduction for sensor array signal processing," *Signal Processing*, vol. 30, no. 2, pp. 245-256, Jan. 1993.
- [3] K. Buckley and X. L. Xu, "Reduced dimension beam-space broad-band source localization: preprocessor design," in *Proc. SPIE Adv. Algorithms Architectures Signal Processing*, vol. 975, pp. 368-376, 1988.
- [4] —, "Spatial-spectrum estimation in a location sector," *IEEE Trans. Acoust., Speech, Signal Processing*, pp. 1842-1852, Nov. 1988.
- [5] P. Forster and G. Vezzosi, "Application of spheroidal sequences to array processing," in *Proc. ICASSP '87*, Apr. 1987, pp. 2268-2271.
- [6] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd edition. Baltimore, MD: Johns Hopkins Univ. Press, 1989.
- [7] H. B. Lee and M. S. Wengrovitz, "Resolution threshold of beamspace MUSIC for two closely spaced emitters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 9, pp. 1545-1559, Sept. 1990.
- [8] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," Tech. Rep. 9213, Yale Univ., Dept. of Elect. Eng., New Haven, CT, Nov. 1992; a short version appeared in *Proc. 26th Asilomar Conf. Signals, Syst. Comput.*, Oct. 1992.
- [9] —, "Vector-sensor array processing for electromagnetic source localization," *IEEE Trans. Signal Processing*, Feb. 1994.
- [10] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. thesis, Stanford Univ., Nov. 1981.
- [11] B. D. Van Veen and R. A. Roberts, "Partially adaptive beamformer design via output power minimization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 11, pp. 1524-1532, Nov. 1987.
- [12] B. D. Van Veen and B. Williams, "Dimensionality reduction in high resolution direction of arrival estimation," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Monterey, CA, Oct. 1988, pp. 588-591.
- [13] —, "Structured covariance matrices and dimensionality reduction in array processing," in *Proc. 4th Workshop Spectrum Estimation Modeling*, Aug. 1988, pp. 168-171.