algorithm," *IEEE Trans. Signal Processing*, vol. SP-40, pp. 1758–1774, July 1992.
[6] J. L. Cadre, "Parametric methods for spatial signal processing in the presence of unknown colored noise fields," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 965–983, July 1989.
[7] M. Viberg and A. Swindlehurst, "Analysis of the combined effects of finite samples and model errors on array processing performance," *IEEE Trans. Signal Processing*, vol. 42, pp. 3073–3083, Nov. 1994.
[8] M. Viberg, "Sensitivity of parametric direction finding to colored noise fields and undermodeling," *Signal Processing*, vol. 34, pp. 207–222, Nov. 1993.
[9] H. Krim and J. G. Proakis, "Smoothed eigenspace-based paramerter estimation," *Automatica*, Special Issue on Statistical Signal Procssing and Control, Jan. 1994.

Second-Order Complex Random Vectors and Normal Distributions

Bernard Picinbono

Abstract— **Complex random vectors are usually described by their covariance matrix. This is insufficient for a complete description of second-order statistics, and another matrix called relation matrix is necessary. Some of its properties are analyzed and used to express the probability density function of normal complex vectors. Various consequences are presented.**

I. Introduction

Complex random vectors (RV's) are widely used in many areas of signal processing such as spectral analysis [1] and array processing [2]. However, the statistical properties of RV's effectively used are essentially limited to those of the covariance matrix. Linear prediction procedures and autoregressive modeling also use only properties of the correlation function of complex signals [1] and [3]. Many questions concerning statistical properties of RV's remain open, however, and some of them will be analyzed in this correspondence. In the first part, we show that the covariance matrix is insufficient to completely describe the statistics of complex RV's, and for this purpose, another matrix is necessary. Its definition and the conditions of its existence are analyzed. By using this matrix, we present the structure of the probability density function (PDF) of normal complex RV's. From this PDF, we deduce the characteristic function and various properties of complex normal random variables. For example, it is shown that contrary to the real case, noncorrelated normal random variables are not generally independent. Conditional PDF's are also analyzed, and the consequences in mean square estimation are presented.

Let us first remind that a complex RV $\mathbf{Z}$ of $\mathbb{C}^n$ is simply a pair of real RV's of $\mathbb{R}^n$ such that $\mathbf{Z} = \mathbf{X} + j\mathbf{Y}$. It is therefore always possible to treat all the problems concerning complex RV's by using a real RV of $\mathbb{R}^{2n}$. However, this procedure is often much more tedious than using directly the RV $\mathbf{Z}$ of $\mathbb{C}^n$.

Manuscript received October 19, 1995; revised March 27, 1996. The associate editor coordinating the review of this paper and approving it for publication was Dr. Monique Fargues.
The author is with the Laboratoire des Signaux et Systèmes, Supélec, Plateau de Moulon, 91190 Gif-sur-Yvette, France.
Publisher Item Identifier S 1053-587X(96)07131-0.

II. Second-Order Properties

Even if the most interesting second-order properties are related to the covariance matrix $\mathbf{\Gamma}$, it does not completely describe the second-order statistical properties of $\mathbf{Z}$. For this, another matrix $\mathbf{C}$, which we refer to as the relation matrix, is necessary. For zero-mean RV's, these matrices are defined by

$$\mathbf{\Gamma} \triangleq E(\mathbf{Z}\mathbf{Z}^H); \quad \mathbf{C} \triangleq E(\mathbf{Z}\mathbf{Z}^T). \tag{1}$$

In these equations, T means transposition, and H means transposition and complex conjugation. The matrix $\mathbf{\Gamma}$ is complex, Hermitian, and nonnegative definite (NND). We assume in the following that there is no zero eigenvalue. The matrix $\mathbf{C}$ is complex and symmetric and therefore satisfies $\mathbf{C}^* = \mathbf{C}^H$, where the star means the complex conjugate. This matrix $\mathbf{C}$ is very rarely introduced in signal processing literature, and the main reason for this fact is that it is explicitly or implicitly assumed to be zero. This characterizes second-order *circularity,* which means that second-order statistics of $\mathbf{Z}$ and $\exp(j\alpha)\,\mathbf{Z}$ are the same for any α. This assumption of circularity [4] is sometimes even introduced in the definition, as, for example, in the normal case (see [1, p. 43] and [5]). In [6], the term "proper" is used instead of "circular." However, circularity is only a particular assumption that is not always valid.

The question that immediately appears is to know whether the matrices $\mathbf{\Gamma}$ and $\mathbf{C}$ must only satisfy the conditions indicated above and deduced from their definition. The answer is no, and we shall establish a necessary and sufficient condition on the pair $(\mathbf{\Gamma}, \mathbf{C})$.

Proposition: Assuming that $\mathbf{\Gamma}$ is complex and positive definite and that $\mathbf{C}$ is complex and symmetric, this matrix $\mathbf{C}$ is a relation matrix of a random vector $\mathbf{Z}$ if and only if the matrix $\mathbf{\Gamma}^* - \mathbf{C}^H\mathbf{\Gamma}^{-1}\mathbf{C}$ is NND.

Proof: Suppose first that $\mathbf{C}$ is the relation matrix of a RV $\mathbf{Z}$. Consider now the RV $\mathbf{W}$ of $\mathbb{C}^{2n}$ defined by $[\mathbf{Z}^T, \mathbf{Z}^H]^T$. Its covariance matrix is a $2n \times 2n$ complex matrix, and a simple calculation yields

$$\mathbf{\Gamma}_2 = \begin{pmatrix} \mathbf{\Gamma} & \mathbf{C} \\ \mathbf{C}^* & \mathbf{\Gamma}^* \end{pmatrix} = \begin{pmatrix} \mathbf{\Gamma} & \mathbf{C} \\ \mathbf{C}^H & \mathbf{\Gamma}^* \end{pmatrix}. \tag{2}$$

As any covariance matrix, it is NND. Its Cholesky block factorization can be written as

$$\mathbf{\Gamma}_2 = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{R} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{R}^H \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \tag{3}$$

where

$$\mathbf{R} = \mathbf{C}^H\mathbf{\Gamma}^{-1}; \quad \mathbf{P} = \mathbf{\Gamma}^* - \mathbf{C}^H\mathbf{\Gamma}^{-1}\mathbf{C}. \tag{4}$$

As $\mathbf{\Gamma}_2$ is NND, the diagonal-block matrix appearing in (3) is also NND. The fact that $\mathbf{\Gamma}$ is PD implies that $\mathbf{P}$ defined by (4) is NND, which gives the only if part.

Suppose now that $\mathbf{C}$ is such that $\mathbf{P}$ is NND. We have to show that there exists a complex RV $\mathbf{Z}$ satisfying (1). It results from (3) that if $\mathbf{\Gamma}$ is positive definite and $\mathbf{P}$ NND, then $\mathbf{\Gamma}_2$, which is defined by (2), is NND. This implies that there exists at least one RV of $\mathbb{C}^{2n}$ such that its covariance matrix is $\mathbf{\Gamma}_2$ (see [3, p. 65]). However, this does not mean that this RV can be partitioned as $[\mathbf{Z}^T, \mathbf{Z}^H]^T$. To arrive at this result, we must introduce the real and imaginary parts $\mathbf{X}$ and $\mathbf{Y}$. For this purpose, let $\mathbf{\Gamma}_{2r}$ be the $2n \times 2n$ matrix defined by

$$\mathbf{\Gamma}_{2r} = \mathbf{M}\mathbf{\Gamma}_2\mathbf{M}^H \tag{5}$$

where $\mathbf{M}$ is defined by

$$\mathbf{M} = \frac{1}{2}\begin{pmatrix} \mathbf{I} & \mathbf{I} \\ -j\mathbf{I} & j\mathbf{I} \end{pmatrix}, \quad \mathbf{M}^{-1} = \begin{pmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I} \end{pmatrix}. \tag{6}$$

It is clear that as $\mathbf{\Gamma}_2$ is NND, $\mathbf{\Gamma}_{2r}$ is also NND. Furthermore, a simple calculation shows that $\mathbf{\Gamma}_{2r}$ is a real symmetric matrix, or

$$\mathbf{\Gamma}_{2r} = \begin{pmatrix} \mathbf{\Gamma}_x & \mathbf{\Gamma}_{xy} \\ \mathbf{\Gamma}_{yx} & \mathbf{\Gamma}_y \end{pmatrix} \tag{7}$$

with

$$\mathbf{\Gamma}_x = \tfrac{1}{2}\mathrm{Re}\,(\mathbf{\Gamma} + \mathbf{C}); \quad \mathbf{\Gamma}_{xy} = \tfrac{1}{2}\mathrm{Im}\,(-\mathbf{\Gamma} + \mathbf{C}) \tag{8}$$

$$\mathbf{\Gamma}_{yx} = \tfrac{1}{2}\mathrm{Im}\,(\mathbf{\Gamma} + \mathbf{C}); \quad \mathbf{\Gamma}_y = \tfrac{1}{2}\mathrm{Re}\,(\mathbf{\Gamma} - \mathbf{C}) \tag{9}$$

where Re and Im stand for real and imaginary parts, respectively.

As $\mathbf{\Gamma}_{2r}$ is symmetric and NND, it is possible to construct at least one vector of $\mathbb{R}^{2n}$ written as $[\mathbf{X}^T, \mathbf{Y}^T]^T$ such that its covariance matrix is $\mathbf{\Gamma}_{2r}$. Taking $\mathbf{Z} = \mathbf{X} + j\mathbf{Y}$, we easily obtain that the covariance and relation matrices of $\mathbf{Z}$ are $\mathbf{\Gamma}$ and $\mathbf{C}$, respectively. This completes the proof.

It is worth pointing out that the matrix $\mathbf{R}$ appears in the linear mean square estimation of $\mathbf{Z}^*$ in terms of $\mathbf{Z}$ by the relation $\widehat{\mathbf{Z}^*} = \mathbf{RZ}$. Furthermore, the corresponding matrix error is $\mathbf{P}$. The innovation $\widetilde{\mathbf{Z}^*} = \mathbf{Z}^* - \mathbf{RZ}$ is uncorrelated with $\mathbf{Z}$. However, it is easy to verify that $E(\widetilde{\mathbf{Z}^*}\mathbf{Z}^T) = \mathbf{P}$. Finally, it is clear that applying the matrix $\mathbf{M}$ defined by (6) to the vector $[\mathbf{Z}^T, \mathbf{Z}^H]^T$ simply gives the vector $[\mathbf{X}^T, \mathbf{Y}^T]^T$.

III. Normal Distributions

Normal RV's arise in many areas of signal processing for well-known reasons. In the complex case, it is almost always assumed that the RV's are also circular, which considerably simplifies the calculation (see [3, p. 118], [5], and [6]). We will present the situation appearing when circularity is not introduced.

A complex RV is said to be normal if its real and imaginary parts $\mathbf{X}$ and $\mathbf{Y}$ are jointly normal. As a consequence, the PDF of such a vector with zero mean value is

$$p(\mathbf{x}, \mathbf{y}) = (2\pi)^{-n}[\det(\mathbf{\Gamma}_{2r})]^{-1/2} \exp[-\tfrac{1}{2}q(\mathbf{x}, \mathbf{y})] \tag{10}$$

with

$$q(\mathbf{x}, \mathbf{y}) = \mathbf{v}^T\mathbf{\Gamma}_{2r}^{-1}\mathbf{v} = \mathbf{v}^H\mathbf{\Gamma}_{2r}^{-1}\mathbf{v}. \tag{11}$$

In this quadratic form, $\mathbf{v}$ is the vector of $\mathbb{R}^{2n}$ defined by $\mathbf{v}^T = [\mathbf{x}^T, \mathbf{y}^T]$, and $\mathbf{\Gamma}_{2r}$ is the covariance matrix defined by (7). It is a $2n \times 2n$ matrix, and the $n \times n$ matrices appearing in its block decomposition are

$$\mathbf{\Gamma}_x = E(\mathbf{XX}^T);\ \mathbf{\Gamma}_y = E(\mathbf{YY}^T);\ \mathbf{\Gamma}_{xy} = E(\mathbf{XY}^T). \tag{12}$$

Note that the last equality of (11) comes from the fact that all the previous elements are real.

As noted, for example, in [7] or [8], it is clear that $\mathbf{x}$ and $\mathbf{y}$ in (10) can be expressed in terms of $\mathbf{z} = \mathbf{x} + j\mathbf{y}$ and of $\mathbf{z}^*$, which introduces another form of the PDF. The calculation of this PDF is given in [7], and we present here a derivation giving the same result but expressed in terms of the matrices $\mathbf{\Gamma}$ and $\mathbf{C}$ previously introduced and not explicitly used in [7].

Let $\mathbf{w}$ be the vector defined by $[\mathbf{z}^T, \mathbf{z}^H]^T$, analog to the RV $\mathbf{W}$ introduced above. It results from this definition that

$$\mathbf{w} = \mathbf{M}^{-1}\mathbf{v}; \quad \mathbf{v} = \mathbf{Mw} \tag{13}$$

where $\mathbf{M}$ is the matrix given by (6). As a consequence, the quadratic form (11) can be expressed as

$$q(\mathbf{x}, \mathbf{y}) = q'(\mathbf{z}, \mathbf{z}^*) = \mathbf{w}^H\mathbf{\Gamma}_w^{-1}\mathbf{w} \tag{14}$$

with $\mathbf{\Gamma}_w^{-1} = \mathbf{M}^H\mathbf{\Gamma}_{2r}^{-1}\mathbf{M}$. Furthermore, the classical rule for the product of determinants yields $\det(\mathbf{\Gamma}_{2r}) = \det(\mathbf{\Gamma}_w)|\det(\mathbf{M})|^2$. By using the fact that a determinant is unchanged by adding rows or columns, it results from (6) that $\det(\mathbf{M}) = j^n 2^{-n}$, and therefore, $[\det(\mathbf{\Gamma}_{2r})]^{-1/2} = 2^n[\det(\mathbf{\Gamma}_w)]^{-1/2}$. By combining all these results, we can express the PDF (10) as

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= p'(\mathbf{z}, \mathbf{z}^*) \\ &= (\pi)^{-n}[\det(\mathbf{\Gamma}_w)]^{-1/2} \exp(-\tfrac{1}{2}\mathbf{w}^H\mathbf{\Gamma}_w^{-1}\mathbf{w}). \end{aligned} \tag{15}$$

This is the result given in [7]. Let us now calculate $\mathbf{\Gamma}_w$ appearing in (14) in terms of the elements of $\mathbf{\Gamma}_{2r}$ defined by (12). By using (6), we obtain that $\mathbf{\Gamma}_w$ is the matrix $\mathbf{\Gamma}_2$ defined by (2), where

$$\begin{aligned} \mathbf{\Gamma} &= \mathbf{\Gamma}_x + \mathbf{\Gamma}_y + j(\mathbf{\Gamma}_{yx} - \mathbf{\Gamma}_{xy}); \\ \mathbf{C} &= \mathbf{\Gamma}_x - \mathbf{\Gamma}_y + j(\mathbf{\Gamma}_{yx} + \mathbf{\Gamma}_{xy}). \end{aligned} \tag{16}$$

These equations are, of course, equivalent to (8) and (9). By using (3) for the determinant and by combining all these results, (15) takes the form

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= p'(\mathbf{z}, \mathbf{z}^*) \\ &= (\pi)^{-n}[\det(\mathbf{\Gamma})\ \det(\mathbf{P})]^{-1/2} \exp[-\tfrac{1}{2}q'(\mathbf{z}, \mathbf{z}^*)] \end{aligned} \tag{17}$$

with

$$q'(\mathbf{z}, \mathbf{z}^*) = \left[\mathbf{z}^H, \mathbf{z}^T\right] \begin{pmatrix} \mathbf{\Gamma} & \mathbf{C} \\ \mathbf{C}^H & \mathbf{\Gamma}^* \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{z} \\ \mathbf{z}^* \end{pmatrix}. \tag{18}$$

The principal interest of this expression is the fact that it uses only the two matrices $\mathbf{\Gamma}$ and $\mathbf{C}$ defined by (1) and having a simple meaning in terms of the complex random vector $\mathbf{Z}$.

There is a case that is especially important. It appears when $\mathbf{C} = \mathbf{0}$, which means that the random vector $\mathbf{Z}$ is *circular*. With this property, the previous equations become

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= p_C(\mathbf{z}) \\ &= (\pi)^{-n}[\det(\mathbf{\Gamma})]^{-1} \exp\left[-\mathbf{z}^H\mathbf{\Gamma}^{-1}\mathbf{z}\right] \end{aligned} \tag{19}$$

which is the classical expression of the PDF of a circular normal vector.

It is now interesting to explicitly express the matrix appearing in the quadratic form (18). By using a simple inverse calculation, we obtain

$$\begin{pmatrix} \mathbf{\Gamma} & \mathbf{C} \\ \mathbf{C}^H & \mathbf{\Gamma}^* \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{P}^{-*} & -\mathbf{R}^H\mathbf{P}^{-1} \\ -\mathbf{R}^T\mathbf{P}^{-*} & \mathbf{P}^{-1} \end{pmatrix} \tag{20}$$

where the matrices $\mathbf{P}$ and $\mathbf{R}$ are defined by (4), and $\mathbf{P}^{-*}$ means $(\mathbf{P}^{-1})^*$. With this matrix, the quadratic form (18) becomes

$$q'(\mathbf{z}, \mathbf{z}^*) = 2\left[\mathbf{z}^H\mathbf{P}^{-*}\mathbf{z} - \mathrm{Re}\,(\mathbf{z}^T\mathbf{R}^T\mathbf{P}^{-*}\mathbf{z})\right]. \tag{21}$$

Finally, the PDF (17) can be written as

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= p(\mathbf{z}, \mathbf{z}^*) \\ &= (\pi)^{-n}[\det(\mathbf{\Gamma})\ \det(\mathbf{P})]^{-1/2} \\ &\quad \cdot \exp\left[-\mathbf{z}^H\mathbf{P}^{-*}\mathbf{z} + \mathrm{Re}\,(\mathbf{z}^T\mathbf{R}^T\mathbf{P}^{-*}\mathbf{z})\right]. \end{aligned} \tag{22}$$

It is possible to put this expression in another form. Applying the matrix inversion lemma (see [1, p. 24]) to the matrix $\mathbf{P}^{-*}$ given by (4) yields

$$\mathbf{P}^{-*} = \mathbf{\Gamma}^{-1} + \mathbf{\Gamma}^{-1}\mathbf{C}\mathbf{P}^{-1}\mathbf{C}^H\mathbf{\Gamma}^{-1}. \tag{23}$$

By inserting this expression in (22), the circular PDF $p_C(\mathbf{z})$ defined by (19) appears, and this gives

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= p_C(\mathbf{z})\left[\det(\mathbf{\Gamma}^{-1}\mathbf{P})\right]^{-1/2} \\ &\quad \cdot \exp\left[-\mathbf{z}^H\mathbf{R}^H\mathbf{P}^{-1}\mathbf{R}\mathbf{z} + \mathrm{Re}\,(\mathbf{z}^T\mathbf{R}^T\mathbf{P}^{-*}\mathbf{z})\right]. \end{aligned} \tag{24}$$

Finally, it can be noted that as $\mathbf{P}$ is Hermitian and positive definite, $\det(\mathbf{P}) = \det(\mathbf{P}^*)$, and by using (4), we obtain $\det(\mathbf{\Gamma}^{-1}\mathbf{P}) = \det(\mathbf{I} - \mathbf{\Gamma}^{-1}\mathbf{C}\mathbf{\Gamma}^{-*}\mathbf{C}^H) = \det(\mathbf{I} - \mathbf{R}^H\mathbf{R}^T)$. As a result, the most general PDF of a complex normal RV can be factorized in a product of the PDF corresponding to the circular case by a function depending only on the matrices $\mathbf{R}$ and $\mathbf{P}$ defined by (4). A similar result is used in [8].

The same procedure can be applied for the characteristic function $\Phi(\mathbf{u}, \mathbf{v})$ defined by

$$\Phi(\mathbf{u}, \mathbf{v}) = E\Big[\exp j(\mathbf{u}^T\mathbf{X} + \mathbf{v}^T\mathbf{Y})\Big]. \tag{25}$$

Because of the normal assumption, this function can be written as

$$\Phi(\mathbf{u}, \mathbf{v}) = \exp\left(-\tfrac{1}{2}\mathbf{a}^H\mathbf{\Gamma}_{2r}\mathbf{a}\right) \tag{26}$$

where $\mathbf{a}$ is the vector of $\mathbb{R}^{2n}$ equal to $[\mathbf{u}^T, \mathbf{v}^T]^T$. By introducing the vector $\mathbf{w} = \mathbf{x} + j\mathbf{y}$, and the vector $\mathbf{b}$ of $\mathbb{C}^{2n}$ equal to $[\mathbf{w}^T, \mathbf{w}^H]^T$, we can write $\mathbf{a} = \mathbf{M}\mathbf{b}$, where $\mathbf{M}$ is defined by (6). The quadratic form appearing in (26) can be expressed in terms of $\mathbf{b}$, which yields

$$\begin{aligned}\Phi(\mathbf{u}, \mathbf{v}) &= \Phi'(\mathbf{w}, \mathbf{w}^*) \\ &= \exp\left\{-\tfrac{1}{4}[\mathbf{w}^H\mathbf{\Gamma}\mathbf{w} + \mathrm{Re}\,(\mathbf{w}^H\mathbf{C}\mathbf{w}^*)]\right\}.\end{aligned} \tag{27}$$

This expression must be compared with (22) giving the PDF.

As a conclusion, we can say that the statistics of a normal complex vector are defined either by the three real $n \times n$ matrices appearing in (12) or by the two complex $n \times n$ matrices $\mathbf{\Gamma}$ and $\mathbf{C}$ defined by (1). If we use these two complex matrices, the PDF and the characteristic function are given by (22) and (27). These expressions are strongly simplified if the circular assumption defined by $\mathbf{C} = \mathbf{0}$ holds, which justifies that in many cases, the circularity is introduced, explicitly or not, in the definition of the complex normality (see [5]). This can be summarized by using the notation $N(\mathbf{m}; \mathbf{\Gamma}, \mathbf{C})$, which means the distribution of a complex normal RV $\mathbf{Z}$ characterized by the mean value $\mathbf{m}$ and the covariance and relation matrices $\mathbf{\Gamma}$ and $\mathbf{C}$. A circular distribution is obviously of the type $N(\mathbf{m}; \mathbf{\Gamma}, \mathbf{0})$ and the real case noted $N(\mathbf{m}; \mathbf{\Gamma})$ appears when $\mathbf{m}$ and $\mathbf{\Gamma}$ are real and $\mathbf{C} = \mathbf{\Gamma}$. At the end of this section, it is worth noting some analogies and differences between the real and complex cases. It is clear that normality is preserved in any linear transformation. The same property is valid for circularity because if $\mathbf{C} = \mathbf{0}$ for $\mathbf{Z}$, this remains valid for $\mathbf{AZ}$ for any matrix $\mathbf{A}$. On the other hand, contrary to the real case, noncorrelation and normality do not imply independence. In fact, the components Z_i of the RV $\mathbf{Z}$ are uncorrelated if and only if the matrix $\mathbf{\Gamma}$ is diagonal. This does not imply that the characteristic function (27) can be factorized as a product of PDF's corresponding to each components, which defines the independence. This factorization, of course, appears when $\mathbf{C} = \mathbf{0}$, i.e., in the circular case. This again shows the analogy between the real and complex circular normal cases. However, this is not the only situation where noncorrelation implies independence. This also appears when $\mathbf{C}$ is diagonal because the last term of (27) can be decomposed in a sum of terms corresponding to each component. Note that we have seen above a more general situation: If $\mathbf{Z}$ is normal, the innovation $\widetilde{\mathbf{Z}^*}$ is normal, as deduced from $\mathbf{Z}$ by a linear transformation, as well as by construction uncorrelated with $\mathbf{Z}$. However, $\widetilde{\mathbf{Z}^*}$ and $\mathbf{Z}$ are not independent because their interrelation matrix is not zero but $\mathbf{P}$.

IV. Conditional Distributions

It is well known that if $\mathbf{X}$ and $\mathbf{Y}$ are two *real* RV's, the best mean square estimation (MSE) of $\mathbf{Y}$ in terms of $\mathbf{X}$ is the conditional expectation or the regression $\mathbf{r}(\mathbf{x}) = E(\mathbf{Y}|\mathbf{x})$ (see [3, p. 393]). It is possible to show that the same result is valid for complex RV's, provided that a correct definition of the conditional expectation is introduced. This can be done as follows.

Suppose that $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are two *complex* RV's written as $\mathbf{Z}_1 = \mathbf{X}_1 + j\mathbf{Y}_1$ and $\mathbf{Z}_2 = \mathbf{X}_2 + j\mathbf{Y}_2$, which introduces the real and imaginary parts. The probability distribution of the pair $\mathbf{Z}_1$, $\mathbf{Z}_2$ is, in fact, that of the four real vectors $\mathbf{X}_i$ and $\mathbf{Y}_j$. Suppose that there is a PDF $p(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2)$. The conditional PDF of $\mathbf{Z}_1$ conditional to $\mathbf{z}_2$ is defined by the classical rule applied to real quantities. Similarly, the conditional expectation is defined by

$$E(\mathbf{Z}_1|\mathbf{z}_2) = E(\mathbf{X}_1|\mathbf{x}_2, \mathbf{y}_2) + jE(\mathbf{Y}_1|\mathbf{x}_2, \mathbf{y}_2). \tag{28}$$

This especially means that if $\mathbf{Z}_1$ is real, while $\mathbf{Z}_2$ is complex, the conditional expectation is also real.

Let us now calculate the conditional PDF $p(\mathbf{z}_1|\mathbf{z}_2)$, or $p(\mathbf{x}_1, \mathbf{y}_1|\mathbf{x}_2, \mathbf{y}_2)$, when the pair of RV's $\mathbf{Z}_1$ and $\mathbf{Z}_2$ is jointly *normal* with zero mean value. By using the notations previously introduced, its PDF can be noted as $N_C(\mathbf{0}, \mathbf{0}; \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_{12}; \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_{12})$, where the matrices $\mathbf{\Gamma}_i$ and $\mathbf{C}_j$ are defined as in (1), whereas

$$\mathbf{\Gamma}_{12} \triangleq E(\mathbf{Z}_1\mathbf{Z}_2^H); \quad \mathbf{C}_{12} \triangleq E(\mathbf{Z}_1\mathbf{Z}_2^T). \tag{29}$$

It is clear that the conditional PDF $p(\mathbf{z}_1|\mathbf{z}_2)$ is still normal because this result is valid for real quantities, and the PDF is defined from the real RV's $\mathbf{X}_i$ and $\mathbf{Y}_j$. We can then say that the PDF is in the form $N_C[\mathbf{m}(\mathbf{z}_2); \mathbf{\Gamma}, \mathbf{C}]$, and the three matrices must be calculated from the matrices $\mathbf{\Gamma}_i$ and $\mathbf{C}_j$ defining the PDF $p(\mathbf{z}_1, \mathbf{z}_2)$.

The mean value $\mathbf{m}(\mathbf{z}_2)$ is the conditional expectation $E(\mathbf{Z}_1|\mathbf{z}_2)$ introduced above. In order to calculate this expectation, we note that in the normal case, the conditional expectation appearing in (28) is linear in $\mathbf{x}_2$ and $\mathbf{y}_2$. As a result, $E(\mathbf{Z}_1|\mathbf{z}_2)$ is widely linear [9] in $\mathbf{z}_2$ or

$$E(\mathbf{Z}_1|\mathbf{z}_2) = \mathbf{r}_1(\mathbf{z}_2) = \mathbf{A}\mathbf{z}_2 + \mathbf{B}\mathbf{z}_2^*. \tag{30}$$

In order to calculate the matrices $\mathbf{A}$ and $\mathbf{B}$, we use the point indicated previously and not shown here that the regression is the best mean square estimation of $\mathbf{Z}_1$ in terms of $\mathbf{Z}_2$. This is characterized by the fact that the innovation $\tilde{\mathbf{Z}}_1 = \mathbf{Z}_1 - \mathbf{r}_1(\mathbf{Z}_2)$, where $\mathbf{r}_1(.)$ is defined by (30), is uncorrelated with $\mathbf{Z}_1$ and $\mathbf{Z}_1^*$. This yields the two orthogonality relations

$$\mathbf{A}\mathbf{\Gamma}_2 + \mathbf{B}\mathbf{C}_2^* = \mathbf{\Gamma}_{12}; \quad \mathbf{A}\mathbf{C}_2 + \mathbf{B}\mathbf{\Gamma}_2^* = \mathbf{C}_{12} \tag{31}$$

allowing the calculation of $\mathbf{A}$ and $\mathbf{B}$. From simple algebraic manipulation, we deduce

$$\mathbf{A} = (\mathbf{\Gamma}_{12} - \mathbf{C}_{12}\mathbf{\Gamma}_2^{-*}\mathbf{C}_2^H)\mathbf{P}_2^{-*}; \ \mathbf{B} = (\mathbf{C}_{12} - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_2^{-1}\mathbf{C}_2)\mathbf{P}_2^{-1} \tag{32}$$

where $\mathbf{P}_2$ is the matrix defined by (4) applied to the vector $\mathbf{Z}_2$.

The matrix $\mathbf{\Gamma}$ appearing in the conditional PDF is the covariance matrix of the vector $\mathbf{Z}_1 - E(\mathbf{Z}_1|\mathbf{z}_2)$ conditional to $\mathbf{z}_2$. However, this vector is the innovation $\tilde{\mathbf{Z}}_1$, which is uncorrelated with $\mathbf{Z}_2$ and $\mathbf{Z}_2^*$. As a consequence, its real and imaginary parts are uncorrelated with $\mathbf{X}_2$ and $\mathbf{Y}_2$, and as these quantities are real and normal, noncorrelation implies independence. This means that $\mathbf{\Gamma}$ is simply the *a priori* covariance matrix of $\tilde{\mathbf{Z}}_1$. By using the expectation given by (30), we obtain

$$\mathbf{\Gamma} = \mathbf{\Gamma}_1 - \mathbf{A}\mathbf{\Gamma}_{12}^H - \mathbf{B}\mathbf{C}_{12}^H. \tag{33}$$

The same reasoning can be applied to the calculation of $\mathbf{C}$, which is the relation matrix of $\tilde{\mathbf{Z}}_1$ conditional to $\mathbf{z}_2$. This yields

$$\mathbf{C} = \mathbf{C}_1 - \mathbf{A}\mathbf{C}_{12}^T - \mathbf{B}\mathbf{\Gamma}_{12}^T. \tag{34}$$

This can be summarized as follows. The conditional PDF of the RV $\mathbf{Z}_1$ conditional to $\mathbf{z}_2$, noted $p(\mathbf{z}_1|\mathbf{z}_2)$, is $N[\mathbf{m}(\mathbf{z}_2); \mathbf{\Gamma}, \mathbf{C}]$. In this expression $\mathbf{m}(\mathbf{z}_2)$ is given by (30), where $\mathbf{A}$ and $\mathbf{B}$ are expressed by (32). Furthermore, $\mathbf{\Gamma}$ and $\mathbf{C}$ are given by (33) and (34), respectively.

Let us finally consider the circular case. It is characterized by the fact that all the matrices $\mathbf{C}$ are zero. This considerably simplifies the structure, and the PDF $p(\mathbf{z}_1|\mathbf{z}_2)$ is normal circular with the mean value $\mathbf{\Gamma}_{12}\mathbf{\Gamma}_2^{-1}\mathbf{z}_2$ and the covariance matrix $\mathbf{\Gamma} = \mathbf{\Gamma}_1 - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_2^{-1}\mathbf{\Gamma}_{21}$. This is exactly the same expression as in the real case.

References

[1] S. Kay, *Modern Spectral Analysis, Theory and Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1989.

[2] R. A. Monzigo and T. W. Miller, *Introduction to Adaptive Arrays.* New York: Wiley, 1980.

[3] B. Picinbono, *Random Signals and Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[4] ——, "On circularity," *IEEE Trans. Signal Processing,* vol. 42, pp. 3473–3482, 1994.

[5] K. Miller, *Multidimensional Gaussian Distributions.* New York: Wiley, 1964.

[6] F. Neeser and J. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inform. Theory,* vol. 39, pp. 1293–1302, 1993.

[7] A. van den Bos, "The multivariate complex normal distribution—A generalization," *IEEE Trans. Inform. Theory,* vol. 41, pp. 537–539, 1995.

[8] P. Amblard and P. Duvaut, "Filtrage adapté dans le cas gaussien complexe non circulaire," in *Proc. 15th Symp. GRETSI,* 1995, pp. 141–144.

[9] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Processing,* vol. 43, pp. 2030–2033, 1995.

Non-Gaussian Multivariate Adaptive AR Estimation Using the Super Exponential Algorithm

Massimiliano Martone

Abstract— **We formulate as a deconvolution problem the causal/noncausal non-Gaussian multichannel autoregressive (AR) parameter estimation problem. The super exponential algorithm presented in a recent paper by Shalvi and Weinstein is generalized to the vector case. We present an adaptive implementation that is very attractive since it is higher order statistics (HOS) based but does not present the high computational complexity of methods proposed up to now.**

I. Introduction

Multichannel time series analysis is widely applied in multisensor signal processing, parallel image processing, multichannel power spectrum estimation, and multichannel digital communication systems. The direct extension of the single channel estimation methods to the multichannel case involves complex matrix operations and becomes unattractive when real-time algorithms are needed to track time-varying parameters. In [5] and [6], the derivation of algorithms to solve this problem was addressed, but those approaches were limited by the use of second-order statistics to causal, minimum-phase models. In a recent paper [1], an attractive approach to deconvolution (which was called the *super exponential algorithm*) was presented: The method is generally valid for non-Gaussian inputs and nonminimum phase systems. In [1], the algorithm was applied to a typical communication problem: the equalization (deconvolution) of a linear system driven by a non-Gaussian source. In this work, we present the generalization of the algorithm to the multichannel case and apply it to the estimation of a causal/noncausal autoregressive (AR) multichannel model. In addition, we propose an adaptive implementation of the algorithm. An important application of the method is the estimation of baseband radio frequency propagation channels in a multisensor antenna receiver for wireless communications when the channels afflicted by multipath are modeled as AR. The interference effect among transmitters is, in fact, well modeled as a multichannel system (see [4]). In a communication system, the channel estimation procedure is important and directly affects the overall system performance. The main motivation of this work is the computational complexity of known algorithms for multichannel system identification based on higher order statistics (HOS) [3], [9] and the limited applicability presented by second-order statistics (SOS) methods. In particular, the approach presented in [3] is not adaptive and not attractive from the computational point of view. On the other hand, the algorithm proposed here is based on a more computationally efficient procedure whose performance is very promising. In Section II, we formulate the problem of the AR estimation as a deconvolution problem and give the necessary equations to derive the algorithm. In Section III, we derive the set of linear equations necessary to perform estimation and describe the iterative procedure to obtain the AR parameters. In Section IV, an adaptive implementation is proposed. In Section V, we show the results of some computer simulations.

Manuscript received December 13, 1994; revised May 3, 1996. The associate editor coordinating the review of this paper and approving it for publication was Dr. Jitendra Tugnait.

The author is with the Watkins Johnson Company, Telecommunications Group, 700 Quince Orchard Rd., Gaithersburg, MD 20878-1794 USA.

Publisher Item Identifier S 1053-587X(96)07140-1.

II. Formulation of the Problem

We assume that the output vector sequence of dimensionality $\tilde{r}$ is generated by a stationary causal/noncausal, non-Gaussian process described by ($(.)^T$ designates transposition, vectors and matrices are bold, notations $[\mathbf{M}]_{l,\,m}$ and $[\mathbf{v}]_k$ stand for the l, m element of matrix $\mathbf{M}$ and the kth element of vector $\mathbf{v}$, respectively)

$$\sum_{i=p_1}^{p_2} \tilde{\mathbf{A}}(i)\mathbf{y}(n-i) = \mathbf{x}(n) \tag{1}$$

where $\mathbf{x}(n)$ is a $\tilde{r}$-variate input process zero mean, spatially and temporally i.i.d., non-Gaussian, and $\mathbf{y}(n)$ is the $\tilde{r}$-variate output vector process. $\tilde{\mathbf{A}}(i)$, $i = p_1, p_1+1, \cdots, p_2$ are matrices of dimension $\tilde{r}\times\tilde{r}$, $\tilde{\mathbf{A}}(p_1) = \mathbf{I}_{\tilde{r}\times\tilde{r}}$ (we indicate with $\mathbf{I}_{n\times n}$ the identity matrix of dimensions $n\times n$), and $\tilde{\mathbf{A}}(p_2)$ has full rank. The observed output process is $\mathbf{y}_n(n) = \mathbf{y}(n) + \mathbf{v}(n)$, where $\mathbf{v}(n)$ is a Gaussian vector process independent of $\mathbf{x}(n)$. We assume, however, for the derivation of the algorithm that the observation noise is not present. Let $\mathbf{A}(i) = \tilde{\mathbf{A}}(i-p_1)$, $p = p_2 - p_1$. In the z domain, we have $\tilde{\mathcal{A}}(z) = \mathcal{A}(z)z^{-p_1}$, where $\mathcal{A}(z) = \mathbf{I}_{\tilde{r}\times\tilde{r}} + \sum_{i=1}^{p}\mathbf{A}(i)z^{-i}$ and the roots of $\det[\tilde{\mathcal{A}}(z)]$ do not lie on the unit circle. It is assumed that $\tilde{\mathcal{A}}(z)$ is *irreducible.* The transfer function of the AR model in the z domain is $\mathcal{H}(z) = [\tilde{\mathcal{A}}(z)]^{-1}$. The input/output representation of the system under the stated assumptions can be expressed (the noise is not considered) as follows:

$$\mathbf{y}(n) = \sum_k \mathbf{H}(k)\mathbf{x}(n-k) \tag{2}$$

where $\mathbf{x}(n) = [x_1(n), x_2(n), \cdots, x_{\tilde{r}}(n)]^T$, $\mathbf{y}(n) = [y_1(n), y_2(n), \cdots, y_{\tilde{r}}(n)]^T$, and $\mathbf{H}(k)$; $k = \cdots, -1, 0, 1, \cdots$