Proper Complex Random Processes with Applications to Information Theory

Fredy D. Neeser, *Student Member, IEEE* and James L. Massey, *Fellow, IEEE*

Abstract— The "covariance" of complex random variables and processes, when defined consistently with the corresponding notion for real random variables, is shown to be determined by the usual (complex) covariance together with a quantity called the pseudo-covariance. A characterization of uncorrelatedness and wide-sense stationarity in terms of covariance and pseudo-covariance is given. Complex random variables and processes with a vanishing pseudo-covariance are called *proper*. It is shown that properness is preserved under affine transformations and that the complex-multivariate Gaussian density assumes a natural form only for proper random variables. The maximum-entropy theorem is generalized to the complex-multivariate case. The differential entropy of a complex random vector with a fixed correlation matrix is shown to be maximum, if and only if the random vector is proper, Gaussian and zero-mean. The notion of *circular stationarity* is introduced. For the class of proper complex random processes, a discrete Fourier transform correspondence is derived relating circular stationarity in the time domain to uncorrelatedness in the frequency domain. As an application of the theory, the capacity of a discrete-time channel with complex inputs, proper complex additive white Gaussian noise, and a finite complex unit-sample response is determined. This derivation is considerably simpler than an earlier derivation for the *real* discrete-time Gaussian channel with intersymbol interference, whose capacity is obtained as a by-product of the results for the complex channel.

Index Terms—Proper complex random processes, circular stationarity, intersymbol interference, capacity.

I. Introduction

THE PURPOSE of this paper is to provide a rounded treatment of certain complex random variables and processes, which we will call *proper*, and to show their usefulness in statistical communication theory. It will be shown, for instance, that the probability density function of a complex Gaussian random vector assumes the anticipated 'natural' form only for proper random vectors. The convenience of proper complex random variables will be demonstrated by the computation of capacity for the complex baseband equivalent of linear bandpass communication channels with memory and additive white Gaussian noise (AWGN).

Linear bandpass channels are usually represented in baseband by an equivalent two-dimensional channel with two

quadrature inputs and outputs [1], [2]. In general, for a passband channel with memory, the quadrature components interfere so that the two-dimensional equivalent baseband channel does not reduce to a pair of independent quadrature channels, as in the memoryless case. To simplify notation, most communication engineers describe the equivalent baseband channel in terms of complex signals and complex impulse responses. Formulations of linear systems for complex-valued signals are also increasingly employed in adaptive signal processing, see e.g., [3]. Somewhat paradoxically, one finds in the literature very few treatments of complex random variables and processes. In fact, many investigators resort to the two-dimensional real representation of systems with complex signals whenever a probabilistic treatment is needed. Notable exceptions are Doob [4], who gives considerable attention to complex Gaussian random processes, and Wooding [5], who first derived the complex-multivariate Gaussian density by assuming certain covariance relations, which are equivalent to properness in our terminology.

The organization of this paper is as follows. In Section II, we characterize second-order statistical properties such as uncorrelatedness and wide-sense stationarity of complex random variables and processes. We show that to specify the four covariances arising between the real and imaginary parts of two complex random variables X and Y, one needs both the conventional covariance $c_{XY} \triangleq E[(X - m_X)(Y - m_Y)^*]$ and the unconventional quantity $\tilde{c}_{XY} \triangleq E[(X - m_X)(Y - m_Y)]$, which we will call the *pseudo-covariance*. Complex random variables and processes with a vanishing pseudo-covariance will be called *proper*. In Section III, we justify the terminology "proper" by demonstrating several natural results for the class of proper complex random variables and processes that do not hold in general. For instance, the probability density function and the entropy of a proper complex Gaussian random vector are specified solely by the vector of means and the matrix of (conventional) covariances. It is also shown that for bandpass communication channels with real wide-sense stationary noise, the complex noise at the demodulator output is proper. In Section IV, we prove a general discrete Fourier transform correspondence between *circular stationarity* in the time-domain and uncorrelatedness in the frequency-domain for sequences of proper complex random variables. An application of this correspondence is provided in Section V, where an earlier derivation of capacity for discrete-time Gaussian channels with memory [23] is considerably simplified by first generalizing to complex channels.

II. PRELIMINARIES

A. Complex Random Variables

A complex random variable X is defined as a random variable of the form

$$X = X_c + jX_s, \qquad j = \sqrt{-1},$$

where the real and imaginary parts, X_c and X_s, are real random variables [4, p. 7]. The subscripts "c" and "s," borrowed from [1] and [2], suggest the cosine and sine components of an equivalent baseband signal. The expectation of a real random variable is naturally generalized to the complex case [6, p. 472] as $E[X] \triangleq E[X_c] + jE[X_s]$. The statistical properties of $X = X_c + jX_s$ are determined by the joint probability density function (pdf) $p_{X_c X_s}(x_c, x_s)$ of X_c and X_s, provided of course that the pdf exists. For convenience, we introduce the notation $p_X(x_c + jx_s) \triangleq p_{X_c X_s}(x_c, x_s)$.

Let F be a complex-valued function whose domain includes the range $X(\Omega)$ of the complex random variable X, where Ω is the sample space. The expectation of $F(X)$ can be expressed in terms of two expectations of real functions in the real random variables X_c and X_s as

$$E[F(X)]$$
$$\triangleq E[\mathrm{Re}\{F(X_c + jX_s)\}] + jE[\mathrm{Im}\{F(X_c + jX_s)\}].$$

Equivalently,

$$E[F(X)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x_c + jx_s)p_X(x_c + jx_s)\,dx_c\,dx_s.$$

To specify the "covariance" of the two complex random vectors $\underline{X} = \underline{X}_c + j\underline{X}_s$ and $\underline{Y} = \underline{Y}_c + j\underline{Y}_s$, the four covariance matrices

$$\mathrm{Cov}\,[\underline{X}_c, \underline{Y}_c]; \quad \mathrm{Cov}\,[\underline{X}_c, \underline{Y}_s];$$
$$\mathrm{Cov}\,[\underline{X}_s, \underline{Y}_c]; \quad \mathrm{Cov}\,[\underline{X}_s, \underline{Y}_s] \tag{1}$$

are needed, where the covariance of two real random vectors $\underline{U}$ and $\underline{V}$ is defined as

$$\mathrm{Cov}\,[\underline{U}, \underline{V}] \triangleq E[(\underline{U} - E[\underline{U}])(\underline{V} - E[\underline{V}])^T)]. \tag{2}$$

The *covariance matrix*

$$\varLambda_{\underline{X}\underline{Y}} \triangleq E[(\underline{X} - m_{\underline{X}})(\underline{Y} - m_{\underline{Y}})^*], \tag{3}$$

where $m_{\underline{X}} \triangleq E[\underline{X}]$, $m_{\underline{Y}} \triangleq E[\underline{Y}]$ and "*" denotes conjugate-transpose, is widely used in the literature. We define also the *pseudo-covariance matrix*

$$\tilde{\varLambda}_{\underline{X}\underline{Y}} \triangleq E[(\underline{X} - m_{\underline{X}})(\underline{Y} - m_{\underline{Y}})^T], \tag{4}$$

which will play a key role in what follows. To simplify the notation for (pseudo-)autocovariance matrices we will write $\varLambda_{\underline{X}}$ (or $\tilde{\varLambda}_{\underline{X}}$) instead of $\varLambda_{\underline{X}\underline{X}}$ (or $\tilde{\varLambda}_{\underline{X}\underline{X}}$). The "covariance" of two complex random vectors can be specified alternatively

by the complex covariance and the pseudo-covariance since it follows from (2)–(4) that

$$\mathrm{Cov}\,[\underline{X}_c, \underline{Y}_c] = \tfrac{1}{2}\mathrm{Re}\,\{\varLambda_{\underline{X}\underline{Y}} + \tilde{\varLambda}_{\underline{X}\underline{Y}}\};$$
$$\mathrm{Cov}\,[\underline{X}_s, \underline{Y}_s] = \tfrac{1}{2}\mathrm{Re}\,\{\varLambda_{\underline{X}\underline{Y}} - \tilde{\varLambda}_{\underline{X}\underline{Y}}\};$$
$$\mathrm{Cov}\,[\underline{X}_s, \underline{Y}_c] = \tfrac{1}{2}\mathrm{Im}\,\{\varLambda_{\underline{X}\underline{Y}} + \tilde{\varLambda}_{\underline{X}\underline{Y}}\};$$
$$\mathrm{Cov}\,[\underline{X}_c, \underline{Y}_s] = -\tfrac{1}{2}\mathrm{Im}\,\{\varLambda_{\underline{X}\underline{Y}} - \tilde{\varLambda}_{\underline{X}\underline{Y}}\}. \tag{5}$$

The natural definition of the *uncorrelatedness* of $\underline{X}$ and $\underline{Y}$ is that all four covariances in (1) vanish. From (5), we now obtain the following simple result.

Lemma 1: The complex random vectors $\underline{X}$ and $\underline{Y}$ are uncorrelated if and only if $\varLambda_{\underline{X}\underline{Y}} = 0$ and $\tilde{\varLambda}_{\underline{X}\underline{Y}} = 0$, i.e., if and only if both the covariance matrix and the pseudo-covariance matrix vanish.

B. Complex Random Processes

A continuous-time (or discrete-time) complex random process is defined as a random process of the form $X(t) \triangleq X_c(t) + jX_s(t)$ (or $X[k] \triangleq X_c[k] + jX_s[k]$), where $X_c(t)$ and $X_s(t)$ (or $X_c[k]$ and $X_s[k]$) are a pair of real continuous-time (or discrete-time) random processes. By definition, a complex random process is wide-sense stationary (w.s.s.) if its real and imaginary parts are jointly w.s.s.. The following result [7, p. 121] characterizes wide-sense stationarity in terms of the mean $m_X(t) \triangleq E[X(t)]$, the *autocorrelation function*

$$r_X(\tau, t) \triangleq E[X(t + \tau)X^*(t)],$$

and the *pseudo-autocorrelation function*

$$\tilde{r}_X(\tau, t) \triangleq E[X(t + \tau)X(t)]$$

of the complex random process $X(\cdot)$.

Lemma 2: A complex random process $X(\cdot)$ is w.s.s., if and only if $m_X(t)$, $r_X(\tau, t)$, and $\tilde{r}_X(\tau, t)$ are independent of t.

The corresponding result for discrete-time processes is obvious.

III. PROPER COMPLEX RANDOM VARIABLES AND PROCESSES

A. Proper Complex Random Variables

Definition 1: A complex random vector $\underline{Z} = \underline{Z}_c + j\underline{Z}_s$ will be called *proper* if its pseudocovariance $\tilde{\varLambda}_{\underline{Z}}$ vanishes. The complex random vectors $\underline{Z}_1$ and $\underline{Z}_2$ will be called *jointly proper* if the composite random vector having $\underline{Z}_1$ and $\underline{Z}_2$ as subvectors is proper.

Note that any subvector of a proper random vector is also proper. However, two individually proper random vectors are not necessarily also jointly proper. Defining

$$\varLambda_{cc} \triangleq \mathrm{Cov}\,[\underline{Z}_c, \underline{Z}_c]; \quad \varLambda_{cs} \triangleq \mathrm{Cov}\,[\underline{Z}_c, \underline{Z}_s];$$
$$\varLambda_{sc} \triangleq \mathrm{Cov}\,[\underline{Z}_s, \underline{Z}_c]; \quad \varLambda_{ss} \triangleq \mathrm{Cov}\,[\underline{Z}_s, \underline{Z}_s] \tag{6}$$

and using the fact that $\varLambda_{cs} = \varLambda_{sc}^T$, the pseudo-covariance of $\underline{Z}$ can be written as

$$\tilde{\varLambda}_{\underline{Z}} = \varLambda_{cc} - \varLambda_{ss} + j(\varLambda_{sc} + \varLambda_{sc}^T). \tag{7}$$

Thus, the vanishing of $\tilde{A}_{\underline{Z}}$ is equivalent to the conditions that

$$A_{cc} = A_{ss} \quad \text{and} \quad A_{sc} = -A_{sc}^{T}, \tag{8}$$

i.e., $\tilde{A}_{\underline{Z}}$ vanishes, if and only if $\underline{Z}_c$ and $\underline{Z}_s$ have identical autocovariance matrices and their crosscovariance matrix is skew-symmetric. Note that the skew-symmetry of A_{sc} implies that A_{sc} has a zero main diagonal, which means that the real and imaginary part of each component Z_k of $\underline{Z}$ are uncorrelated. The vanishing of $\tilde{A}_{\underline{Z}}$ does not, however, imply that the real part of Z_k and the imaginary part of Z_l are uncorrelated for $k \neq l$. It should be pointed out that a *real* random vector is a proper complex random vector, if and only if it is constant (with probability 1), since $A_{ss} = 0$ and (8) imply $A_{cc} = 0$.

The appropriateness of the term "proper" in connection with complex random vectors is supported by the following lemma dealing with closure under affine transformations as well as by a number of other results to follow.

Lemma 3: Let $\underline{Z}$ be a proper complex n-dimensional random vector, i.e., $\tilde{A}_{\underline{Z}} = 0$. Then any random vector obtained from $\underline{Z}$ by a linear or affine transformation, i.e., any random vector $\underline{Y}$ of the form $\underline{Y} = A\underline{Z} + \underline{b}$, where $A \in \mathbb{C}^{m \times n}$ and $\underline{b} \in \mathbb{C}^m$ are constant, is also proper.

Proof: Since

$$m_{\underline{Y}} = Am_{\underline{Z}} + \underline{b} \text{ and } \underline{Y} - m_{\underline{Y}} = A(\underline{Z} - m_{\underline{Z}}),$$

we have

$$\tilde{A}_{\underline{Y}} = E[(\underline{Y} - m_{\underline{Y}})(\underline{Y} - m_{\underline{Y}})^T] = A\tilde{A}_{\underline{Z}}A^T = 0. \quad \square$$

Note that $\underline{Y}$ and $\underline{Z}$ as in Lemma 3 are automatically jointly proper, since the vector having $\underline{Y}$ and $\underline{Z}$ as subvectors is obtained by the affine transformation

$$\begin{bmatrix} \underline{Y} \\ \underline{Z} \end{bmatrix} = \hat{A}\underline{Z} + \begin{bmatrix} \underline{b} \\ \underline{0}_n \end{bmatrix}, \qquad \hat{A} \triangleq \begin{bmatrix} A \\ I_n \end{bmatrix}.$$

Lemma 4: Let $\underline{Z}_1$ and $\underline{Z}_2$ be two independent complex random vectors and let $\underline{Z}_2$ be proper. Then the linear combination $\underline{Y} = a_1\underline{Z}_1 + a_2\underline{Z}_2$, where a_1 and a_2 are complex numbers and $a_1 \neq 0$, is proper, if and only if $\underline{Z}_1$ is also proper.

Proof: The independence of $\underline{Z}_1$ and $\underline{Z}_2$ and the properness of $\underline{Z}_2$ imply

$$\tilde{A}_{\underline{Y}} = a_1^2\tilde{A}_{\underline{Z}_1} + a_2^2\tilde{A}_{\underline{Z}_2} = a_1^2\tilde{A}_{\underline{Z}_1}.$$

Thus, $\tilde{A}_{\underline{Y}}$ vanishes, if and only if $\tilde{A}_{\underline{Z}_1}$ vanishes. $\quad \square$

Lemma 1 immediately implies the following result.

Lemma 5: Two jointly proper, complex random vectors $\underline{Z}_1$ and $\underline{Z}_2$ are uncorrelated, if and only if their covariance matrix $A_{\underline{Z}_1\underline{Z}_2}$ vanishes.

A *complex Gaussian random vector* $\underline{Z}$ is defined as a vector with jointly Gaussian real and imaginary parts. Following Feller [6, p. 86], we consider Gaussian distributions to include degenerate distributions concentrated on a lower-dimensional manifold. In such degenerate cases, the $2n \times 2n$-covariance matrix

$$\Phi \triangleq \text{Cov}\left[\begin{bmatrix} \underline{Z}_c \\ \underline{Z}_s \end{bmatrix}, \begin{bmatrix} \underline{Z}_c \\ \underline{Z}_s \end{bmatrix}\right] = \begin{bmatrix} A_{cc} & A_{cs} \\ A_{sc} & A_{ss} \end{bmatrix} \tag{9}$$

is singular and the pdf does not exist unless one admits generalized functions.

Note that two jointly proper Gaussian random vectors $\underline{Z}_1$ and $\underline{Z}_2$ are independent, if and only if $A_{\underline{Z}_1\underline{Z}_2} = 0$, which follows from Lemma 5 and the fact that uncorrelatedness and independence are equivalent for Gaussian random variables.

Wooding [5] first derived the pdf of a complex Gaussian random vector satisfying the conditions (8), i.e., of a *proper* complex Gaussian random vector. Goodman [8] gave an alternative derivation based on the observation that the multiplication of certain orthogonal 2×2-matrices is isomorphic to the multiplication of related complex numbers. The complex-multivariate Gaussian pdf is also found in [9]-[11]. The results can be stated as follows.

Theorem 1: Let $\underline{Z}$ be a proper complex n-dimensional Gaussian random vector with mean $\underline{m}$ and nonsingular covariance matrix $A \triangleq E[(\underline{Z} - \underline{m})(\underline{Z} - \underline{m})^*]$. Then the pdf of $\underline{Z}$ is given by

$$p_{\underline{Z}}(\underline{z}) \triangleq p_{\underline{Z}_c\underline{Z}_s}(\underline{z}_c, \underline{z}_s)$$
$$= \frac{1}{\pi^n \det(A)}\exp\left\{-(\underline{z} - \underline{m})^*A^{-1}(\underline{z} - \underline{m})\right\}. \tag{10}$$

Conversely, let the pdf of a random vector $\underline{Z}$ be given by (10), where A is Hermitian and positive definite. Then $\underline{Z}$ is proper complex and Gaussian with covariance matrix A and mean $\underline{m}$. Moreover, for a proper complex $\underline{Z}$

$$A = 2(A_{cc} + jA_{sc})$$
$$A^{-1} = \frac{1}{2}\Delta^{-1}(I - jA_{sc}A_{cc}^{-1})$$
$$\Delta \triangleq A_{cc} + A_{sc}A_{cc}^{-1}A_{sc}$$
$$\det(A) = 2^n\sqrt{\det(A_{cc})\det(\Delta)}. \tag{11}$$

Note that the pdf (10) is completely specified by the vector of means and the conventional covariance matrix. The fact that the function (10) integrates to one over $\underline{z}_c$ and $\underline{z}_s$ for any positive-definite Hermitian matrix A was proved by Bellman without connection to pdf's [12, ch. 6, section 10]. A proof of Theorem 1 is included in the Appendix since it provides insight and some intermediate results will be used in the sequel.

As an application of Theorem 1, we generalize the maximum-entropy theorem [13, theorem 7.4.1], [15, theorem 9.6.5] to the complex-multivariate case. The result will be used in Section V to compute the capacity of a channel with proper complex Gaussian noise. Consider a complex Gaussian random vector $\underline{Z} = \underline{Z}_c + j\underline{Z}_s$, whose differential entropy is appropriately defined as the joint differential entropy of its real and imaginary part, i.e., $h(\underline{Z}) \triangleq h(\underline{Z}_c\underline{Z}_s)$.

Theorem 2: Let $\underline{Z}$ be a complex, continuous, n-dimensional random vector with nonsingular correlation matrix $R_{\underline{Z}} \triangleq E[\underline{Z}\underline{Z}^*]$. Then,

$$h(\underline{Z}) \leq \log[(\pi e)^n \det(R_{\underline{Z}})]$$

with equality, if and only if $\underline{Z}$ is *proper* and Gaussian with zero mean.

It is somewhat surprising that $\underline{Z}$ must be *proper* in order to maximize entropy. Note also that no real random vector

maximizes entropy for a given correlation matrix $R_{\underline{Z}}$ when complex random vectors are allowed. The proof of the analog to Theorem 2 for real random variables [13, p. 336], [15, p. 234] is easily generalized to a proof of Theorem 2 by using the Gaussian density (10). We give a different proof of Theorem 2 for a scalar complex random variable $Z = Z_c + jZ_s$ with the constraint $E[|Z|^2] = S$ that better illustrates the role of properness. According to the maximum-entropy theorem for real random vectors, the real random vector $\underline{W} \triangleq [Z_c, Z_s]^T$, for which $R_{\underline{W}} \triangleq E[\underline{W}\,\underline{W}^T]$ is nonsingular, satisfies

$$h(Z) = h(\underline{W}) \leq \tfrac{1}{2} \log\left[(2\pi e)^2 \det(R_{\underline{W}})\right]$$

with equality, if and only if $\underline{W}$ is zero-mean Gaussian. By hypothesis, $E[|Z|^2] = E[Z_c^2] + E[Z_s^2] = S$ and thus

$$\det(R_{\underline{W}}) = E[Z_c^2]E[Z_s^2] - (E[Z_cZ_s])^2$$
$$\leq E[Z_c^2]E[Z_s^2] \leq S^2/4,$$

where equality holds at both places, if and only if $E[Z_cZ_s] = 0$ and $E[Z_c^2] = E[Z_s^2]$. Therefore, $h(Z) = h(Z_cZ_s) = h(\underline{W}) \leq \log[\pi eS]$ with equality, if and only if Z is proper and Gaussian with zero mean.

Note that $h(\underline{Z}) = h(\underline{Z}_c\underline{Z}_s) = h(\underline{Z}_c) + h(\underline{Z}_s|\underline{Z}_c)$ for a complex random vector $\underline{Z} = \underline{Z}_c + j\underline{Z}_s$ and, when $\underline{Z}$ is Gaussian,

$$h(\underline{Z}_c) = \tfrac{1}{2}\log\left[(2\pi e)^n \det(\Lambda_{cc})\right], \qquad (12)$$

where Λ_{cc} is as in Theorem 1. It follows from (12), the proof of Theorem 1 (see (A.9) in the Appendix), and from Theorem 2 that, for a proper complex Gaussian random vector $\underline{Z}$,

$$h(\underline{Z}_s|\underline{Z}_c) = \tfrac{1}{2}\log\left[(2\pi e)^n \det(\Delta)\right].$$

The matrix Δ is defined in (11) and known as the *Schur-complement* [14, p. 46] of Λ_{cc} in the matrix Φ of (9).

For a complex Gaussian random vector $\underline{Z}$ with zero mean and covariance matrix $\Lambda = \Lambda_c + j\Lambda_s$, Theorem 2 implies the following nontrivial result in matrix theory, as suggested by a reviewer.

Corollary 1: For any symmetric matrix $\Lambda_c \in \mathbb{R}^{n \times n}$ and skew-symmetric matrix $\Lambda_s \in \mathbb{R}^{n \times n}$ such that $\Lambda_c + j\Lambda_s$ is positive definite,

$$\max_{\substack{\begin{bmatrix} \Lambda_{cc} & \Lambda_{sc}^T \\ \Lambda_{sc} & \Lambda_{ss} \end{bmatrix} \geq 0 \\ \Lambda_{cc}+\Lambda_{ss}=\Lambda_c \\ \Lambda_{sc}-\Lambda_{sc}^T=\Lambda_s \\ \Lambda_{cc}^T=\Lambda_{cc}, \Lambda_{ss}^T=\Lambda_{ss}}} \det\begin{bmatrix} \Lambda_{cc} & \Lambda_{sc}^T \\ \Lambda_{sc} & \Lambda_{ss} \end{bmatrix}$$

$$= \frac{1}{2^{2n}}\det(\Lambda_c)\det(\Lambda_c + \Lambda_s\Lambda_c^{-1}\Lambda_s),$$

which is achieved, if and only if $\Lambda_{cc} = \Lambda_{ss} = \tfrac{1}{2}\Lambda_c$ and $\Lambda_{sc} = \tfrac{1}{2}\Lambda_s$.

The simple proof is left to the reader.

As one might expect, the differential entropy of real and complex random variables is affected differently by *scaling*. For any matrix $A = A_c + jA_s \in \mathbb{C}^{n \times n}$, we can represent $\underline{Y} = A\underline{X}$ by $\begin{bmatrix} \underline{Y}_c \\ \underline{Y}_s \end{bmatrix} = B\begin{bmatrix} \underline{X}_c \\ \underline{X}_s \end{bmatrix}$, where $B \triangleq \begin{bmatrix} A_c & -A_s \\ A_s & A_c \end{bmatrix}$. The

scaling property for real random vectors [15, p. 234] and the fact that $\det(B) = |\det(A)|^2$ [8, p. 156] now imply

$$h(A\underline{X}) = h(\underline{X}_c\underline{X}_s) + \log|\det(B)|$$
$$= h(\underline{X}) + 2\log|\det(A)|. \qquad (13)$$

For a complex, nondegenerate scalar random variable X, (13) yields

$$e^{h(aX)} = |a|^2 e^{h(X)}, \qquad a \in \mathbb{C},$$

which is plausible since the entropy power of a random variable can be interpreted as the effective size of its support set and the support set of X is an area.

B. Proper Complex Random Processes

The *covariance function* of a complex random process is defined as

$$c_Z(\tau, t) \triangleq E[(Z(t+\tau) - m_Z(t+\tau))(Z(t) - m_Z(t))^*], \qquad (14)$$

for continuous-time processes and as

$$c_Z[k, n] \triangleq E[(Z[n+k] - m_Z[n+k])(Z[n] - m_Z[n])^*], \qquad (15)$$

for discrete-time processes, where obvious notation has been used for the means. Analogously, we will define the *pseudo-covariance function* of a complex random process as

$$\tilde{c}_Z(\tau, t) \triangleq E[(Z(t+\tau) - m_Z(t+\tau))(Z(t) - m_Z(t))] \qquad (16)$$

for continuous-time processes and as

$$\tilde{c}_Z[k, n] \triangleq E[(Z[n+k] - m_Z[n+k])(Z[n] - m_Z[n])] \qquad (17)$$

for discrete-time processes.

Definition 2: A complex random process will be called *proper* if its pseudo-covariance function vanishes identically.

Using similar arguments as in Section III-A one can show that any linear or affine transformation of a proper complex random process is proper and that a linear combination of independent proper complex random processes is also proper. Moreover, any vector of samples taken from a proper complex random process is also proper.

Proper complex random processes arise in equivalent baseband representations of bandpass communication systems, as we show next. Consider the (real) additive noise channel together with the receiver front-end shown in Fig. 1. The real process $X_0(\cdot)$ is assumed to be w.s.s. and bandlimited to frequencies ω such that $||\omega| - \omega_0| \leq 2\pi W$, where $\omega_0 > 2\pi W$, and the real noise process $Z_0(\cdot)$ is w.s.s. with zero mean and power spectral density $S_{Z_0}(\omega)$. The channel output $Y_0(\cdot)$, another real process, is converted to baseband by a complex demodulator and an ideal lowpass filter $g(\tau)$ with frequency response

$$G(\omega) = \begin{cases} 1, & \text{if } |\omega| \leq 2\pi W, \\ 0, & \text{otherwise.} \end{cases} \qquad (18)$$

Let the complex random processes $X(t)$ and $Z(t)$ denote the response of the receiver front-end, which is a time-varying linear system, to the real random processes $X_0(t)$
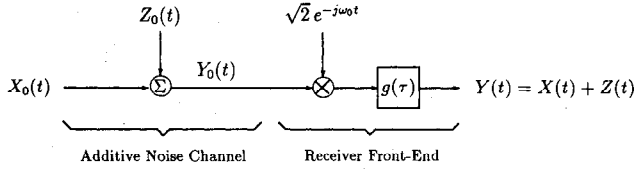
Fig. 1. Additive noise channel and receiver front-end.

and $Z_0(t)$, respectively. Wozencraft and Jacobs have shown that $Y(t) = X(t) + Z(t)$ provides sufficient statistics for an optimum receiver [1, p. 496]. Of particular interest here are the properties of the demodulated noise $Z(\cdot)$ that were proved in [1, p. 498] and can be summarized in our terminology as follows.

Theorem 3: Let the real w.s.s. process $Z_0(\cdot)$ with zero mean and power spectral density $S_{Z_0}(\omega)$ be the input to a complex demodulator with angular frequency ω_0 followed by an ideal low-pass filter (18). Then, if $\omega_0 > 2\pi W$, the complex random process

$$Z(t) \triangleq \sqrt{2} \int_{-\infty}^{\infty} Z_0(u) \, e^{-j\omega_0 u} g(t-u) \, du$$

at the lowpass filter output is w.s.s., *proper*, zero mean, and has the autocorrelation function

$$r_Z(\tau) \triangleq E[Z(t+\tau)Z^*(t)]$$
$$= \frac{1}{\pi} \int_{-2\pi W}^{2\pi W} S_{Z_0}(\omega + \omega_0) e^{j\omega\tau} \, d\omega. \qquad (19)$$

In particular, if $Z_0(\cdot)$ is white noise with power spectral density $S_{Z_0}(\omega) = \mathcal{N}_0/2$, then

$$r_Z(\tau) = \mathcal{N}_0 \frac{\sin 2\pi W\tau}{\pi \tau}. \qquad (20)$$

Since $Z(\cdot)$ is w.s.s., proper, and zero mean, $c_Z(\tau) = r_Z(\tau)$ and the pseudo-covariance vanishes, i.e.,

$$\tilde{c}_Z(\tau) = \tilde{r}_Z(\tau) \equiv 0. \qquad (21)$$

Property (21) is equivalent to the symmetry relations[1]

$$r_{Z_c Z_c}(\tau) = r_{Z_s Z_s}(\tau) \quad \text{and} \quad r_{Z_s Z_c}(\tau) = -r_{Z_s Z_c}(-\tau), \qquad (22)$$

i.e., the real and imaginary part of $Z(\cdot)$ have the same autocorrelation function and an odd crosscorrelation function. Equivalent symmetry relations were found by Dugundji and Zakai for a real process $X(\cdot)$ and its Hilbert transform $\hat{X}(\cdot)$ [16]–[18]. The process $Z(\cdot) \triangleq X(\cdot) + j\hat{X}(\cdot)$ was called the "pre-envelope" of $X(\cdot)$ or an "analytic signal" and satisfies (21). However, the requirement that the imaginary part be the Hilbert transform of the real part is more stringent than the symmetry relations (22) and the concept of the pre-envelope is not appropriate for single random variables, as opposed to properness.

It should be mentioned that complex random processes with nonzero mean are usually not of interest, since a "complex envelope" $X(\cdot)$ with nonzero-mean corresponds

to a nonstationary bandpass process. To see this, let $X_0(t) = \text{Re}\{X(t)\sqrt{2} \, e^{j\omega_0 t}\}$ and note that $E[X_0(t)] = \text{Re}\{E[X(t)]\sqrt{2} \, e^{j\omega_0 t}\} \neq$ constant if $E[X(t)] \neq 0$.

We are particularly interested in the class of proper complex *Gaussian* random processes. Doob has given conditions that in our terminology are the necessary and sufficient conditions for the existence of such processes [4, theorem 3.1]. Theorem 3 shows that demodulated Gaussian noise belongs to this class. The *proper complex AWGN channel* will be defined as a channel of the form $Y(\cdot) = X(\cdot) + Z(\cdot)$, where $X(\cdot)$ and $Z(\cdot)$ are independent complex processes and $Z(\cdot)$ is proper complex AWGN with power spectral density $\mathcal{N}_0$. The proper complex *white* noise idealization is supported by the following consideration: If we choose a large bandwidth W and a carrier frequency $\omega_0 > 2\pi W$ in Theorem 3, then the correlation function $r_Z(\tau)$ closely approximates $\mathcal{N}_0 \, \delta(\tau)$.

IV. CIRCULAR STATIONARITY

In this section, upper-case and lower-case letters denote frequency-domain and time-domain variables, respectively. For convenience, a length-N sequence $x[0], x[1], \cdots, x[N-1]$ will be written as $x[0, N-1]$. All indices in square brackets are understood to be taken modulo the integer N.

Definition 3: A sequence of complex random variable $z[0, N-1]$ will be called *circularly wide-sense stationary (c.w.s.s.)*, if $E[z[n]] = m_z$ is independent of n and if

$$E[z[n]z^*[i]] = r_z[n-i] \quad \text{and} \quad E[z[n]z[i]] = \tilde{r}_z[n-i] \qquad (23)$$

holds for $0 \leq i, n < N$, i.e., if the correlation of two samples depends only on their time difference modulo N. We will call $r_z[0, N-1]$ and $\tilde{r}_z[0, N-1]$ the *circular correlation sequence* and *circular pseudo-correlation sequence*, respectively, of the c.w.s.s. sequence $z[0, N-1]$. Analogously, a sequence of real random variables $x[0, N-1]$ will be called c.w.s.s., if $E[x[n]] = m_x$ is independent of n and if

$$E[x[n]x[i]] = r_x[n-i], \qquad 0 \leq i, n < N.$$

A proper complex, non-trivial c.w.s.s. sequence $z[0, N-1]$ can be generated as the circular convolution of a proper complex white noise sequence $w[0, N-1]$ with some weighting sequence $h[0, N-1]$, i.e., $z[n] = \sum_{k=0}^{N-1} h[n-k]w[k]$, where $m_w = 0$, $r_w[i] = \mathcal{N}_0 \, \delta[i]$, and $\delta[i] = 1$, $i = 0$, and $\delta[i] = 0$, $i \neq 0$. A simple calculation shows that $z[0, N-1]$ is c.w.s.s. with $m_z = 0$ and circular correlation function

$$r_z[i] = E[z[n+i]z^*[n]] = \mathcal{N}_0 \sum_{k=0}^{N-1} h[k+i]h^*[k].$$

We now show that circular stationarity of a proper complex time-domain sequence corresponds to uncorrelatedness of the components of its *discrete Fourier transform* (DFT). This fact will be used in Section V to simplify finding the capacity of the complex discrete-time Gaussian channel with memory.

Recall that the DFT of a complex sequence $z[0, N-1]$ is the sequence $Z[0, N-1]$ given by

$$Z[k] \triangleq \sum_{n=0}^{N-1} z[n] \, \Omega_N^{-kn}, \qquad 0 \leq k < N, \qquad (24)$$

---

[1] We define $r_{UV}(\tau, t) \triangleq E[U(t+\tau)V(t)]$ for any real processes $U(\cdot)$ and $V(\cdot)$ and write $r_{UV}(\tau)$ when $U(\cdot)$ and $V(\cdot)$ are jointly w.s.s.

where $\Omega_N \triangleq e^{j2\pi/N}$ is a primitive Nth root of unity. The time-domain sequence $z[0, N-1]$ can be recovered from the frequency-domain sequence $Z[0, N-1]$ by the *inverse discrete Fourier transform* (IDFT)

$$z[n] \triangleq \frac{1}{N}\sum_{k=0}^{N-1} Z[k]\, \Omega_N^{kn}, \qquad 0 \le n < N. \tag{25}$$

Note also that

$$\sum_{n=0}^{N-1}\Omega_N^{nk} = N\delta[k]. \tag{26}$$

If $z[0, N-1]$ is a sequence of complex random variables, then so also is $Z[0, N-1]$. Clearly, $z[0, N-1]$ is zero-mean, if and only if $Z[0, N-1]$ is zero-mean. Moreover, by Lemma 3 and the invertibility of the DFT, $z[0, N-1]$ is proper if and only if $Z[0, N-1]$ is proper.

Theorem 4: Let $z[0, N-1]$ and its DFT $Z[0, N-1]$ be proper complex sequences with zero mean. Then the *time-domain sequence $z[0, N-1]$ is c.w.s.s., if and only if the frequency-domain sequence $Z[0, N-1]$ is uncorrelated*, i.e., if and only if

$$E[Z[k]Z^*[l]] = NR_z[k]\delta[k-l], \tag{27}$$

where $R_z[0, N-1]$ is the DFT of the circular correlation sequence $r_z[0, N-1]$, i.e.,

$$R_z[k] = \sum_{n=0}^{N-1} r_z[n]\, \Omega_N^{-kn}; \quad r_z[n] = \frac{1}{N}\sum_{k=0}^{N-1} R_z[k]\, \Omega_N^{kn}. \tag{28}$$

Proof: Suppose that the proper complex random sequence $z[0, N-1]$ is c.w.s.s. with circular correlation sequence $r_z[0, N-1]$ and let $R_z[0, N-1]$ be the DFT of $r_z[0, N-1]$. Then,

$$E[Z[k]Z^*[l]] = \sum_{i=0}^{N-1}\Omega_N^{li}\sum_{n=0}^{N-1} E[z[n]z^*[i]]\, \Omega_N^{-kn}$$

$$= \sum_{i=0}^{N-1}\Omega_N^{li}\sum_{n=0}^{N-1} r_z[n-i]\, \Omega_N^{-kn}$$

$$= R_z[k]\sum_{i=0}^{N-1}\Omega_N^{(l-k)i} = NR_z[k]\delta[k-l],$$

where the second, third and last equality follow from (23), the shifting property of the DFT [19, p. 92] and from (26), respectively. It now follows from the properness of $Z[0, N-1]$ and from the fact that $Z[0, N-1]$ has zero mean that the components of $Z[0, N-1]$ are uncorrelated. Conversely, suppose that $Z[0, N-1]$ satisfies (27) and let $r_z[0, N-1]$ be the IDFT of $R_z[0, N-1]$. Then,

$$E[z[n]z^*[i]] = \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} E[Z[k]Z^*[l]]\, \Omega_N^{kn-li}$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} R_z[k]\sum_{l=0}^{N-1}\delta[k-l]\, \Omega_N^{kn-li}$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} R_z[k]\, \Omega_N^{k(n-i)} = r_z[n-i].$$

The circular pseudo-correlation sequence of $z[0, N-1]$ vanishes because $z[0, N-1]$ is proper with zero mean. It follows that $z[0, N-1]$ is c.w.s.s.. ☐

Note when $z[0, N-1]$ is c.w.s.s. that in general $E[|Z[k]|^2]$ depends on k, i.e., $Z[0, N-1]$ is not c.w.s.s.. Since uncorrelatedness and independence are equivalent for Gaussian random variables, we immediately have

Corollary 2: Let $z[0, N-1]$ and its DFT $Z[0, N-1]$ be proper complex *Gaussian* sequences with zero-mean. Then the *time-domain sequence $z[0, N-1]$ is c.w.s.s., if and only if the components of the *frequency-domain sequence $Z[0, N-1]$ are *independent*.

It was recently shown by Hirt and Massey that the *real* DFT of a sequence of real, i.i.d., zero-mean Gaussian random variables is another sequence of real, i.i.d., zero-mean Gaussian random variables [23, lemmas 1, 2]. Similarly, it was shown that the inverse *real* DFT of a (nonstationary) frequency-domain sequence with real, independent, zero-mean Gaussian components is a sequence of real, correlated, zero-mean Gaussian random variables [23, lemma 3]. Note that the Lemmas 1–3 of [23] are special cases of the analog to Corollary 2 for real random variables. Thus, to establish a correspondence between circular stationarity and uncorrelatedness, the *real* DFT is needed in the case of real random variables, while the *ordinary* DFT is adequate for proper complex random variables.

Note that Theorem 4 is easily generalized to K-channel systems, $K \ge 2$, if the length-N sequences $z[0, N-1]$ and $Z[0, N-1]$ are replaced by length-N sequences $\underline{z}[0, N-1]$ and $\underline{Z}[0, N-1]$ of K-dimensional vectors, the correlation sequences $r_z[0, N-1]$ and $R_z[0, N-1]$ are replaced by $K \times K$-matrix sequences $\mathbf{r}_z[0, N-1]$ and $\mathbf{R}_z[0, N-1]$, and the DFT of a vector (or matrix) sequence is defined to be the vector (or matrix) of DFT's. This generalization is useful in the study of multi-user channels with finite memory.

V. CAPACITY OF THE DISCRETE-TIME GAUSSIAN CHANNEL WITH MEMORY

Hirt and Massey recently computed the capacity C of the *discrete-time Gaussian channel* (DTGC) with finite intersymbol interference (ISI) and an average symbol-energy constraint by introducing a hypothetical channel model, the *N-circular Gaussian channel* (NGCG) [23]. Using the real DFT, they showed the equivalence of the NCGC to a set of N parallel, decoupled memoryless channels. The per-symbol capacity C_N of the NCGC was then obtained using the "water-filling theorem" [13, theorem 7.5.1]. Moreover, they proved that the DTGC and the NCGC are asymptotically equivalent in the sense that

$$C = \lim_{N\to\infty} C_N. \tag{29}$$

Verdú [20] also used a circular convolution approach to determine the capacity region of the symbol-asynchronous Gaussian multiple-access channel.

As an application of proper complex random variables, Hirt and Massey's derivation is simplified in this section

and their results are generalized to channels with proper complex AWGN and a complex unit-sample response. A similar approach can be used to simplify the computation of capacity of Gaussian multiple-access channels with memory [21] as well as their complex generalizations.

The notation is as in Section IV. We first consider a *real* DTGC whose channel filter has a real unit-sample response $(h_0, h_1, \cdots, h_M)$ and assume further that $h_0 \neq 0$ and $h_M \neq 0$. Consider now that one has available two instances of this DTGC, viz.,

$$y_{c_n} = \sum_{m=0}^{M} h_m x_{c_{n-m}} + w_{c_n}, \qquad -\infty < n < \infty, \quad (30)$$

and

$$y_{s_n} = \sum_{m=0}^{M} h_m x_{s_{n-m}} + w_{s_n}, \qquad -\infty < n < \infty, \quad (31)$$

where $\{w_{c_n}\}$ and $\{w_{s_n}\}$ are independent zero-mean white Gaussian noise (WGN) sequences each sample of which has variance $\mathcal{N}_0/2$ and where the real inputs are subject to the symbol-energy constraints

$$E[x_{c_n}^2] \leq E_s/2 \text{ and } E[x_{s_n}^2] \leq E_s/2, \quad -\infty < n < \infty. \quad (32)$$

This pair of real DTGC's can be represented by the (one-dimensional) complex (or two-dimensional real) channel

$$y_n = \sum_{m=0}^{M} h_m x_{n-m} + w_n, \qquad -\infty < n < \infty, \quad (33)$$

where $x_n \triangleq x_{c_n} + j x_{s_n}$, $w_n \triangleq w_{c_n} + j w_{s_n}$, and $y_n \triangleq y_{c_n} + j y_{s_n}$. Since $\{w_{c_n}\}$ and $\{w_{s_n}\}$ have the same autocorrelation function, a vanishing crosscorrelation function and zero means, it follows that $\{w_n\}$ is a proper complex WGN sequence. Moreover, $E[w_n] = 0$ and $E[|w_n|^2] = \mathcal{N}_0$, for all n. It will be shown next that the constraints (32) can be replaced by the weaker condition $E[|x_n|^2] \leq E_s$. Clearly, capacity can be achieved on the channel (33) by independent sequences $\{x_{c_n}\}$ and $\{x_{s_n}\}$, since there is no "crosstalk" between the real and imaginary component channels and the real and imaginary noise sequences are independent. If the capacity-achieving input distribution also satisfies (32), then $C^{2D} = 2C^{1D}$, where C^{1D} and C^{2D} are the capacities of the real channel (30) [or (31)] and the complex (or two-dimensional real) channel (33), respectively. We can now get additional generality by allowing the unit-sample response in (33) to be complex. The resulting channel (33) will be called the *complex* DTGC. Following [23], we define the *complex* NCGC[2]

$$y[n] = \sum_{i=0}^{N-1} h[i] x[n - i] + w[n], \qquad 0 \leq n < N, \quad (34)$$

where $N > M$, where the sequence $h[0, N - 1]$ is obtained by padding $h_0, h_1, \cdots, h_M$ with zeros as

$$h[i] \triangleq \begin{cases} h_i, & \text{if } 0 \leq i \leq M, \\ 0, & \text{if } M < i < N, \end{cases}$$

[2] As in Section IV, all indices in square brackets are understood to be taken modulo N.

and where $w[\cdot]$ is proper complex c.w.s.s. Gaussian noise with zero mean and circular correlation sequence

$$r_w[i] = E[w[n + i] w^*[n]] = \mathcal{N}_0 \, \delta[i], \qquad 0 \leq i < N.$$

For brevity, $w[0, N-1]$ will be called a proper complex WGN sequence. Moreover, the input data $x[n]$ are subject to the constraint

$$E[|x[n]|^2] \leq E_s, \qquad 0 \leq n < N. \quad (35)$$

It can be easily shown that the complex DTGC and the complex NCGC are asymptotically equivalent in the sense of (29) by essentially the same argument as given in [23].

Theorem 5: The per-symbol capacity of the complex NCGC (34) is given by

$$C_N^{2D} = \frac{1}{N} \sum_{k=0}^{N-1} \log \left[\max \left(\beta |H[k]|^2 / \mathcal{N}_0, 1 \right) \right], \quad (36)$$

where $H[0, N - 1]$ is the DFT of $h[0, N - 1]$ and where the parameter β is determined from the condition

$$\sum_{k=0}^{N-1} \epsilon[k] = N E_s, \quad (37)$$

in which the *spectral energy distribution* $\epsilon[0, N - 1]$ depends on β through

$$\epsilon[k] = \max \left(\beta - \mathcal{N}_0 / |H[k]|^2, 0 \right), \qquad 0 \leq k < N. \quad (38)$$

Moreover, capacity is achieved if and only if the input sequence $x[0, N - 1]$ is proper, Gaussian, and c.w.s.s. with zero mean and its circular correlation sequence is the IDFT of $\epsilon[0, N - 1]$, i.e.,

$$r_x[i] \triangleq E[x[n + i] x^*[n]] = \frac{1}{N} \sum_{k=0}^{N-1} \epsilon[k] \, \Omega_N^{ik}, \quad 0 \leq i < N. \quad (39)$$

Proof: It was proved in [23] that the capacity of the (real) NCGC equals the supremum of the average mutual information between the input and output sequence over all pdf's satisfying a weaker *block-energy constraint*. Analogously, it will be shown for the complex NCGC that

$$C_N^{2D} = I_N^{2D}, \quad (40)$$

where

$$I_N^{2D} \triangleq \sup_{p_N} \frac{1}{N} I(x[0, N - 1]; y[0, N - 1]),$$

where the supremum is over all pdf's p_N for $x[0, N - 1]$ satisfying

$$\sum_{n=0}^{N-1} E[|x[n]|^2] \leq N E_s. \quad (41)$$

Taking the DFT of (34) yields a set of parallel, *memoryless Gaussian channels* (MGC's)

$$Y[k] = H[k] X[k] + W[k], \qquad 0 \leq k < N, \quad (42)$$

where $H[0, N-1]$ is the DFT of $h[0, N-1]$ and where the $W[k]$ are i.i.d. proper complex Gaussian random variables with zero mean and variance[3] $N\mathcal{N}_0$ by Corollary 2 and (27). Using Parseval's relation, the constraint (41) becomes

$$\sum_{k=0}^{N-1} E[|X[k]|^2] \leq N^2 E_s \qquad (43)$$

in the frequency domain. By (41) and the invertibility of the DFT,

$$I_N^{2D} = \sup_{q_N} \frac{1}{N} I(X[0, N-1]; Y[0, N-1]), \qquad (44)$$

where the supremum is over all pdf's q_N for $X[0, N-1]$ satisfying (43). As in the real case, it can be easily shown that

$$I(X[0, N-1]; Y[0, N-1]) \leq \sum_{k=0}^{N-1} I(X[k]; Y[k]) \qquad (45)$$

with equality, if and only if the outputs $Y[k]$ are independent [13, p. 321]. Using (45), (44) can be written as a supremum over the allowed spectral energy distributions $\epsilon[0, N-1]$, viz.,

$$I_N^{2D} = \sup_{\substack{\epsilon[0, N-1]: \\ \sum_{i=0}^{N-1} \epsilon[i] \leq NE_s}} \frac{1}{N} \sum_{k=0}^{N-1} C[k], \qquad (46)$$

where

$$C[k] \triangleq \sup_{\substack{q[k]: \\ E[|X[k]|^2] \leq N\epsilon[k]}} I(X[k]; Y[k]) \qquad (47)$$

is the capacity of the kth MGC (42). The average energy at the output of this MGC is bounded by

$$E[|Y[k]|^2] \leq N(|H[k]|^2 \epsilon[k] + \mathcal{N}_0) = S[k], \qquad (48)$$

with equality, if and only if $E[|X[k]|^2] = N\epsilon[k]$. Note that the condition for equality applies to all MGC's k such that $H[k] \neq 0$. By Theorem 2,

$$I(X[k]; Y[k]) = h(Y[k]) - h(Y[k]|X[k])$$
$$\leq \log[S[k]/(N\mathcal{N}_0)] \qquad (49)$$

with equality, if and only if $Y[k]$ is proper and Gaussian with zero mean and variance $S[k]$. According to Cramèr's Theorem [6], which states that the sum of two independent random variables is Gaussian, if and only if each of the two random variables is itself Gaussian, and by Lemma 4, equality holds in (49) if and only if $X[k]$ is proper and Gaussian with zero mean and variance $N\epsilon[k]$. Again, the condition for equality applies for all k such that $H[k] \neq 0$. Therefore,

$$C[k] = \log[1 + |H[k]|^2 \epsilon[k]/\mathcal{N}_0]. \qquad (50)$$

The solution to the water-filling problem can be adopted from [13, theorem 7.5.1] without change and yields (36)-(38). Since capacity is achieved only for proper Gaussian inputs $X[k]$ and since the noise samples $W[k]$ are independent, the necessary and sufficient condition for equality in (45) is equivalent to the

independence of the *inputs* $X[k]$. Thus, capacity is achieved, if and only if the inputs $X[k]$ are independent, proper, and Gaussian with zero mean and variance $N\epsilon[k]$. Invoking Theorem 4 once more shows that capacity is achieved, if and only if the input sequence $x[0, N-1]$ is proper, Gaussian, and c.w.s.s. with zero mean and the circular correlation sequence (39). Since $r_x[0] = E_s$, the symbol-energy constraint (35) is also satisfied, confirming (40). $\square$

We now return to the special case of the real NCGC

$$v[n] = \sum_{i=0}^{N-1} h[i]u[n-i] + z[n], \qquad 0 \leq n < N, \qquad (51)$$

treated by [23], where the unit-sample response and all random variables are real. The noise sequence $z[0, N-1]$ is assumed to be white Gaussian with zero mean and

$$E[(z[n])^2] = \mathcal{N}_0/2, \qquad 0 \leq n < N,$$

and the inputs are subject to

$$E[(u[n])^2] \leq E_s/2, \qquad 0 \leq n < N. \qquad (52)$$

Corollary 3: The per-symbol capacity of the real NCGC (51) is given by

$$C_N^{1D} = C_N^{2D}/2, \qquad (53)$$

where C_N^{2D} is obtained as in Theorem 5. Moreover, capacity is achieved, if and only if the input sequence $u[0, N-1]$ is Gaussian and c.w.s.s. with zero mean and its circular correlation sequence is given by

$$r_u[i] \triangleq E[u[n+i]u[n]]$$
$$= \frac{1}{N} \sum_{k=0}^{N-1} \frac{\epsilon[k]}{2} \cos(2\pi ik/N), \qquad 0 \leq i < N. \qquad (54)$$

Note that β is related to the parameter θ used in [23] by $\beta = \theta\mathcal{N}_0$. Note also that our $\epsilon[k]$ is defined to be twice the "$\epsilon[k]$" defined in [23].

Proof of Corollary 3: The proof of Theorem 5 remains valid in the case of a real sequence $h[0, N-1]$ and the complex NCGC (34) reduces to a pair of independent real NCGC's of the form (51). Since capacity is achieved by zero-mean proper complex inputs and since $r_x[0] = E_s$ we obtain $E[x_c^2[n]] = E[x_s^2[n]] = E_s/2, 0 \leq n < N$, i.e., on each of the real component NCGC's the constraint (52) is satisfied with equality. Thus, (53) is proved. Moreover, since $h[0, N-1]$ is real, $H[0]$ is also real and $H[k] = H^*[N-k], 1 \leq k < N$ [19, p. 110]. Therefore, (38) yields $\epsilon[k] = \epsilon[N-k], 1 \leq k < N$, and (39) gives

$$r_x[i] = \frac{1}{N} \left[\epsilon[0] + \sum_{k=1}^{N-1} \epsilon[k] \, \Omega_N^{ik} \right]$$
$$= \frac{1}{N} \left[\epsilon[0] + \frac{1}{2} \left(\sum_{k=1}^{N-1} \epsilon[k] \, \Omega_N^{ik} \sum_{k=1}^{N-1} \epsilon[N-k] \Omega_N^{ik} \right) \right]$$
$$= \frac{1}{N} \sum_{k=0}^{N-1} \epsilon[k] \cos(2\pi ik/N), \qquad 0 \leq i < N.$$

---

[3] In accordance with (3), the *variance* of a scalar complex random variable is defined as $\text{var}[X] \triangleq E[|X - m_X|^2]$.

Since the capacity-achieving inputs are proper, Gaussian and zero-mean and since the circular correlation sequence is real, the sequences $x_c[0, N-1]$ and $x_s[0, N-1]$ are independent and $r_u[i] = r_{x_c}[i] = r_{x_s}[i] = r_x[i]/2$, $0 \leq i < N$. $\qquad\square$

VI. SUMMARY

Second-order statistical properties have been characterized for complex random variables and processes. Proper complex random variables and processes, which are characterized by a vanishing pseudo-covariance, were shown to have the desirable feature that their second-order statistics are specified completely by their mean and their covariance. It was demonstrated that the complex-multivariate Gaussian density takes on a natural form only for proper random variables. The differential entropy of a complex random vector with a fixed correlation matrix was shown to be maximum, if and only if the random vector is zero-mean Gaussian and *proper*. The notion of circular stationarity was introduced and, for the class of proper complex random processes, a DFT correspondence was derived relating circular stationarity in the time domain to uncorrelatedness in the frequency domain. The derivation of the capacity of the discrete-time Gaussian channel with memory was simplified and the results were generalized to channels with proper complex AWGN and a complex unit-sample response.

APPENDIX

To prove Theorem 1, the following result on quadratic forms is needed.

Lemma A.1: Let M_{cc}, M_{ss}, M_{sc}, and M_{cs} be real $n \times n$-matrices, where M_{cc} and M_{ss} are symmetric and $M_{cs}^T = M_{sc}$. Define the Hermitian $n \times n$-matrix

$$M = M_c + jM_s \triangleq M_{cc} + M_{ss} + j(M_{sc} - M_{sc}^T)$$

and the symmetric $2n \times 2n$-matrix

$$\Psi \triangleq 2 \begin{bmatrix} M_{cc} & M_{cs} \\ M_{sc} & M_{ss} \end{bmatrix}. \qquad (A.1)$$

Then the quadratic forms

$$\mathcal{E} \triangleq \underline{z}^* M \underline{z}$$

and

$$\mathcal{E}' \triangleq [\underline{z}_c^T \underline{z}_s^T] \Psi \begin{bmatrix} \underline{z}_c \\ \underline{z}_s \end{bmatrix}, \qquad (A.2)$$

are equal for all $\underline{z} \triangleq \underline{z}_c + j\underline{z}_s$, if and only if

$$M_{cc} = M_{ss} \text{ and } M_{sc} = -M_{sc}^T. \qquad (A.3)$$

Moreover, under the conditions (A.3) M is positive (semi)definite, if and only if Ψ is positive (semi)definite.

Proof: Since $\mathcal{E}$ is a Hermitian form, it is real for all $\underline{z}$. Hence,

$$\begin{aligned}
\mathcal{E} &= \operatorname{Re}\{\underline{z}^* M \underline{z}\} = \underline{z}_c^T M_c \underline{z}_c + \underline{z}_s^T M_c \underline{z}_s + \underline{z}_s^T M_s \underline{z}_c - \underline{z}_c^T M_s \underline{z}_s \\
&= [\underline{z}_c^T \underline{z}_s^T] \begin{bmatrix} M_c & -M_s \\ M_s & M_c \end{bmatrix} \begin{bmatrix} \underline{z}_c \\ \underline{z}_s \end{bmatrix} \\
&= [\underline{z}_c^T \underline{z}_s^T] \begin{bmatrix} M_{cc} + M_{ss} & -M_{sc} + M_{sc}^T \\ M_{sc} - M_{sc}^T & M_{cc} + M_{ss} \end{bmatrix} \begin{bmatrix} \underline{z}_c \\ \underline{z}_s \end{bmatrix}, \qquad (A.4)
\end{aligned}$$

by definition of M. Comparing (A.4) to (A.2) shows that (A.3) gives the necessary and sufficient conditions for the two quadratic forms to be identical. But $\mathcal{E} \equiv \mathcal{E}'$ shows that M is positive (semi)definite if and only if this is also true for Ψ. $\square$

Proof of Theorem 1: We first prove the direct part for $\underline{m} = \underline{0}$. Recalling that any covariance matrix Λ is positive semidefinite, we see that $\det(\Lambda) \neq 0$ implies that Λ is in fact positive definite. Thus, Φ defined by (9) is positive definite by Lemma A.1. Since $\underline{Z}_c$ and $\underline{Z}_s$ are jointly Gaussian,

$$p_{\underline{Z}_c \underline{Z}_s}(\underline{z}_c, \underline{z}_s) = \frac{1}{(2\pi)^n \sqrt{\det \Phi}} \exp\left\{-\frac{1}{2}[\underline{z}_c^T \underline{z}_s^T] \Phi^{-1} \begin{bmatrix} \underline{z}_c \\ \underline{z}_s \end{bmatrix}\right\}. \qquad (A.5)$$

We now show that the exponents of (10) and (A.5) are equal. Using a standard result for inverting block matrices [22, p. 656] and the properness of $\underline{Z}$, which implies $\Lambda_{cc} = \Lambda_{ss}$ and $\Lambda_{cs} = \Lambda_{sc}^T = -\Lambda_{sc}$, we obtain

$$\Phi^{-1} = \begin{bmatrix} \Delta^{-1} & \Lambda_{cc}^{-1} \Lambda_{sc} \Delta^{-1} \\ -\Delta^{-1} \Lambda_{sc} \Lambda_{cc}^{-1} & \Delta^{-1} \end{bmatrix}, \qquad (A.6)$$

where Δ, defined in (11), is symmetric. Note that Φ^{-1} is nonsingular since Φ is nonsingular, which implies that Λ_{cc}^{-1} and Δ^{-1} exist. Moreover, Φ^{-1} is symmetric, since the inverse of a symmetric matrix is symmetric. Next, we show that the upper-right block of Φ^{-1} is skew-symmetric. Observing that

$$\Delta \Lambda_{cc}^{-1} \Lambda_{sc} = \Lambda_{sc} + \Lambda_{sc} \Lambda_{cc}^{-1} \Lambda_{sc} \Lambda_{cc}^{-1} \Lambda_{sc} = \Lambda_{sc} \Lambda_{cc}^{-1} \Delta,$$

we obtain

$$\begin{aligned}
\Lambda_{cc}^{-1} \Lambda_{sc} \Delta^{-1} &= \Delta^{-1} \Lambda_{sc} \Lambda_{cc}^{-1} = (\Lambda_{cc}^{-1} \Lambda_{sc}^T \Delta^{-1})^T \\
&= -(\Lambda_{cc}^{-1} \Lambda_{sc} \Delta^{-1})^T,
\end{aligned}$$

i.e., the upper-right block, and thus also the lower-left block, is skew-symmetric. Thus, Φ^{-1} has the same properties as Ψ in (A.1), namely symmetry, equal diagonal blocks, and skew-symmetry of off-diagonal blocks. Therefore, Lemma A.1 applies for $\Psi \triangleq \frac{1}{2} \Phi^{-1}$ and $M \triangleq \frac{1}{2} \Lambda^{-1}(I - j\Lambda_{sc}\Lambda_{cc}^{-1})$. By the properness of $\underline{Z}$, Λ is given as in (11). Multiplying out $M\Lambda$ yields the identity matrix. Therefore, $M = \Lambda^{-1}$ and the exponents of (10) and (A.5) are equal. It remains to show that $2^n \sqrt{\det(\Phi)} = \det(\Lambda)$. Note that the determinant of a Hermitian matrix is always real. Using a well-known result for the determinant of block matrices (cf. [22, p. 650] or [14, p. 46]) and the fact that $\Lambda_{cs} = -\Lambda_{sc}$, we obtain

$$\det(\Phi) = \det(\Lambda_{cc}) \det(\Delta). \qquad (A.7)$$

To compute $\det(\Lambda)$, note that $\Lambda^T = 2(\Lambda_{cc} - j\Lambda_{sc}) = 2(I - j\Lambda_{sc}\Lambda_{cc}^{-1})\Lambda_{cc}$. Therefore, $\Lambda^{-1} = \frac{1}{4}\Delta^{-1}\Lambda^T\Lambda_{cc}^{-1}$. But

$$\det(\Lambda\Lambda^{-1}) = \det\left(\frac{1}{4}\Lambda\,\Delta^{-1}\Lambda^T\Lambda_{cc}^{-1}\right)$$
$$= \frac{[\det(\Lambda)]^2}{2^{2n}\det(\Delta)\det(\Lambda_{cc})} = 1. \qquad (A.8)$$

Combining (A.8) and (A.7) yields

$$\det(\Lambda) = 2^n\sqrt{\det(\Lambda_{cc})\det(\Delta)} = 2^n\sqrt{\det(\Phi)}. \qquad (A.9)$$

Now let $\underline{Z}$ have nonzero-mean $\underline{m}$. Then $\underline{Z} - \underline{m}$ is zero-mean Gaussian and has the pdf (10).

We now turn to the converse part. Since Λ is positive definite, so also is $M = \Lambda^{-1}$. According to Lemma A.1, there exists a unique symmetric, positive-definite matrix Ψ such that in the case $\underline{m} = \underline{0}$ one has $\mathcal{E} = \mathcal{E}'$ for all $\underline{z} \triangleq \underline{z}_c + j\underline{z}_s$. In the words of Feller [6, p. 84], Ψ induces a normal density in $2n$ dimensions. Thus, $[\underline{Z}_c^T, \underline{Z}_s^T]^T$ is Gaussian with mean $[\underline{m}_c^T, \underline{m}_s^T]^T$ and covariance matrix $\Phi = \frac{1}{2}\Psi^{-1}$. By Lemma A.1, Ψ has equal diagonal blocks and skew-symmetric off-diagonal blocks, and by the argument in the direct part of the proof, the matrix Φ shares the same properties. This implies the properness of $\underline{Z}$ and the claim follows. □

ACKNOWLEDGMENT

REFERENCES

[1] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.

[2] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1983.

[3] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice Hall, 1986.

[4] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1953.

[5] R. A. Wooding, "The multivariate distribution of complex normal variables," *Biometrika*, vol. 43, pp. 212–215, 1956.

[6] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. II. New York: Wiley, 1966.

[7] W. A. Gardner, *Introduction to Random Processes*, 2nd ed. New York: McGraw-Hill, 1990.

[8] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution," *Ann. Math. Statist.*, vol. 34, pp. 152–176, 1963.

[9] E. J. Kelly and I. S. Reed, "Some properties of stationary Gaussian processes," MIT Lincoln Lab, Tech. Rep. TR-157, June 1957.

[10] R. Arens, "Complex processes for envelopes of normal noise," *IRE Trans. Inform. Theory*, vol. IT-3, pp. 204–207, Sept. 1957.

[11] I. S. Reed, "On a moment theorem for complex Gaussian processes," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 194–195, Apr. 1962.

[12] R. Bellman, *Introduction to Matrix Analysis*, 2nd ed. New York: McGraw-Hill, 1970.

[13] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[14] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. New York: Academic Press, 1985.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[16] J. Dugundji, "Envelopes and pre-envelopes of real waveforms," *IRE Trans. Inform. Theory*, vol. IT-4, pp. 53–57, Mar. 1958.

[17] M. Zakai, "Second-order properties of the pre-envelope and envelope processes," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 556–557, Dec. 1960.

[18] ———, "The representation of narrow-band processes," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 323–325, July 1962.

[19] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1975.

[20] S. Verdú, "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 35, pp. 733–751, July 1989.

[21] R. S. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling," *IEEE Trans. Inform. Theory*, vol. 39, pp. 773–785, May 1993..

[22] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice Hall, 1980.

[23] W. Hirt and J. L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 34, pp. 380–388, May 1988.