Rotationally Variant Complex Channels

Georg Tauböck

Telecommunications Research Center Vienna (FTW), Tech Gate Vienna,
Donau-City-Straße 1, A-1220 Vienna, Austria,
tauboeck@ftw.at,
WWW home page: http://www.ftw.at

Abstract. In the past, several authors studied the properties of rotationally variant complex random vectors (see e.g. [1]-[5]). It is well known that the covariance[1] matrix $C_n = \mathbf{E}\{nn^H\}$ is not sufficient for a complete second order description, the knowledge of the *pseudo-covariance* matrix $P_n = \mathbf{E}\{nn^T\}$ (the random vector n is assumed to be zero-mean) is mandatory, too. Additionally, if a channel introduces rotationally variant complex noise, the corresponding encoder and decoder have to take into account the pseudo-covariance matrix or, equivalently, work on a real/imaginary part level. Otherwise, a capacity loss has to be accepted. The purpose of this paper is to derive analytical formulas (in terms of covariance and pseudo-covariance matrix) for this capacity loss.

1 Introduction

Throughout this paper, the channel is modeled via

$$y = Ax + n, \tag{1}$$

where $y \in \mathbb{C}^r$ and $x \in \mathbb{C}^t$ denote the received and transmitted vectors, respectively. A is a deterministic $r \times t$ complex matrix, and $n \in \mathbb{C}^r$ is zero-mean complex Gaussian noise with nonsingular covariance matrix $C_n = \mathbf{E}\{nn^H\}$ and *pseudo-covariance* matrix $P_n = \mathbf{E}\{nn^T\}$. The transmitter is constrained in its total power to S,

$$\mathbf{E}\{x^H x\} \le S, \tag{2}$$

or equivalently, since $x^H x = \mathrm{tr}\left(xx^H\right)$, and expectation and trace commute,

$$\mathrm{tr}\left(\mathbf{E}\{xx^H\}\right) \le S. \tag{3}$$

In principle, all encoding-decoding schemes for this channel can be classified into two types, the ones which utilize the knowledge of P_n and the others which neglect it, or - to be more precise - erroneously assume that P_n is the zero matrix. These two types of transmission have different capacities in general, $C_{\text{neglect } P_n}$ and $C_{\text{utilize } P_n}$, which motivates the calculation of the capacity loss

[1] The superscripts ^{H} and ^{T} denote *Hermitian transposed* and *transposed*, respectively, $\mathbf{E}\{\cdot\}$ denotes the *expectation operator*.

$\Delta C = C_{\text{neglect } P_n} - C_{\text{utilize } P_n}$. In the following, it will turn out that this can be done analytically, and furthermore, that the final formula is astonishingly simple, i.e., it depends only on the singular values (c.f., [8]) of a transformed version of the pseudo-covariance matrix P_n.

Before we proceed with this specific problem, we present some preliminaries about complex random vectors.

2 Preliminaries

A complex random vector $x \in \mathbb{C}^n$ is said to be Gaussian if the real random vector $\hat{x} \in \mathbb{R}^{2n}$ consisting of its real and imaginary parts, $\hat{x} = \begin{bmatrix} \Re\{x\} \\ \Im\{x\} \end{bmatrix}$, is Gaussian. Thus, to specify the distribution of a zero-mean complex Gaussian random vector x, it is necessary to specify the covariance of $\hat{x}$, namely,

$$C_{\hat{x}} = \mathbf{E}\{\hat{x}\hat{x}^H\} = \mathbf{E}\{\hat{x}\hat{x}^T\} \in \mathbb{R}^{2n \times 2n}. \tag{4}$$

Using the covariance matrix $C_x = \mathbf{E}\{xx^H\}$ and the pseudo-covariance matrix $P_x = \mathbf{E}\{xx^T\}$ one finds - after some algebra -

$$\begin{aligned} C_{\hat{x}} &= \begin{bmatrix} \mathbf{E}\{\Re\{x\}\Re\{x^T\}\} & \mathbf{E}\{\Re\{x\}\Im\{x^T\}\} \\ \mathbf{E}\{\Im\{x\}\Re\{x^T\}\} & \mathbf{E}\{\Im\{x\}\Im\{x^T\}\} \end{bmatrix} \\ &= \underbrace{\frac{1}{2}\begin{bmatrix} \Re\{C_x\} & -\Im\{C_x\} \\ \Im\{C_x\} & \Re\{C_x\} \end{bmatrix}}_{\hat{C}_x} + \underbrace{\frac{1}{2}\begin{bmatrix} \Re\{P_x\} & \Im\{P_x\} \\ \Im\{P_x\} & -\Re\{P_x\} \end{bmatrix}}_{\tilde{P}_x}, \end{aligned} \tag{5}$$

from which it follows that the knowlege of C_x and P_x also determines a zero-mean complex Gaussian random vector completely.

Lemma 1.

$$\begin{aligned} \Re\{C_x\} &= \Re\{C_x^T\} \\ \Im\{C_x\} &= -\Im\{C_x^T\} \\ \Re\{P_x\} &= \Re\{P_x^T\} \\ \Im\{P_x\} &= \Im\{P_x^T\} \end{aligned} \tag{6}$$

Proof. The real part of an Hermitian matrix is symmetric, the imaginary part of an Hermitian matrix is anti-symmetric, and real and imaginary parts of a symmetric matrix are symmetric. □

We will use Lemma 1 to show an important property of $\tilde{P}_x$ which will be needed later on:

Theorem 1. *There exists a diagonal matrix $\Lambda_x \in \mathbb{R}^{2n}$ (consisting of the eigenvalues of $\tilde{P}_x$) and an orthonormal matrix $Q_x \in \mathbb{R}^{2n}$, s.t.*

$$\tilde{P}_x = Q_x \Lambda_x Q_x^T, \tag{7}$$

$$\lambda \in \textit{diag}\{\Lambda_x\} \Leftrightarrow -\lambda \in \textit{diag}\{\Lambda_x\} \quad \textit{(including multiplicities).} \tag{8}$$

Proof. According to Lemma 1, $\tilde{P}_x$ is a symmetric matrix. Applying the *spectral theorem*, c.f. [8], yields (7). Let I_n denote the $n \times n$ identitiy matrix. We have,

$$\begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} \tilde{P}_x \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} \begin{bmatrix} \Re\{P_x\} & \Im\{P_x\} \\ \Im\{P_x\} & -\Re\{P_x\} \end{bmatrix} \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} = -\tilde{P}_x, \tag{9}$$

and therefore,

$$\begin{aligned} \det\left(\tilde{P}_x - \lambda I_{2n}\right) = \det\left(-\tilde{P}_x - \lambda I_{2n}\right) = (-1)^{2n} \det\left(\tilde{P}_x + \lambda I_{2n}\right) = \\ = \det\left(\tilde{P}_x + \lambda I_{2n}\right), \end{aligned} \tag{10}$$

which implies (8). □

Definition 1. *A complex random vector is called* rotationally invariant *(c.f., [1]) or* proper *(c.f., [5]) or* circularly symmetric *(c.f., [6]) if its pseudo-covariance matrix vanishes.*

For any $z \in \mathbb{C}^n$ and $A \in \mathbb{C}^{n \times m}$ define

$$\hat{z} = \begin{bmatrix} \Re\{z\} \\ \Im\{z\} \end{bmatrix} \quad \text{and} \quad \hat{A} = \begin{bmatrix} \Re\{A\} & -\Im\{A\} \\ \Im\{A\} & \Re\{A\} \end{bmatrix}, \text{ s.t.,}$$

for a rotationally invariant random vector, $C_{\hat{x}} = \frac{1}{2}\hat{C}_x$.

Lemma 2. *The mappings* $z \mapsto \hat{z} = \begin{bmatrix} \Re\{z\} \\ \Im\{z\} \end{bmatrix}$ *and* $A \mapsto \hat{A} = \begin{bmatrix} \Re\{A\} & -\Im\{A\} \\ \Im\{A\} & \Re\{A\} \end{bmatrix}$ *have the following properties:*

$$\begin{gathered} C = AB \Leftrightarrow \hat{C} = \hat{A}\hat{B} \\ C = A + B \Leftrightarrow \hat{C} = \hat{A} + \hat{B} \\ C = A^H \Leftrightarrow \hat{C} = \hat{A}^T \\ C = A^{-1} \Leftrightarrow \hat{C} = \hat{A}^{-1} \\ \det\left(\hat{A}\right) = \left|\det(A)\right|^2 = \det\left(AA^H\right) \\ z = x + y \Leftrightarrow \hat{z} = \hat{x} + \hat{y} \\ y = Ax \Leftrightarrow \hat{y} = \hat{A}\hat{x} \\ \Re\{x^H y\} = \hat{x}^T \hat{y} \\ U \in \mathbb{C}^{n \times n} \textit{ is unitary} \Leftrightarrow \hat{U} \in \mathbb{R}^{2n \times 2n} \textit{ is orthonormal} \\ C \in \mathbb{C}^{n \times n} \textit{ is non-negative definite} \Rightarrow \hat{C} \in \mathbb{R}^{2n \times 2n} \textit{ is non-negative definite} \end{gathered} \tag{11}$$

Proof. All properties are immediate or almost immediate (c.f., also [6]). □

We can now state the following theorem about the eigenvalues of $\tilde{P}_x$ (see, Theorem 1):

Theorem 2. *The positive eigenvalues of* $\tilde{P}_x$ *are the non-zero singular values of* P_x.

Proof. On one hand, we have

$$P_x P_x^H = \Re\{P_x\}\Re\{P_x\} + \Im\{P_x\}\Im\{P_x\} + j\left(\Im\{P_x\}\Re\{P_x\} - \Re\{P_x\}\Im\{P_x\}\right), \quad (12)$$

where Lemma 1 was used, on the otherhand,

$$\tilde{P}_x \tilde{P}_x - \lambda I_{2n} = \quad (13)$$
$$= \begin{bmatrix} \Re\{P_x\}\Re\{P_x\} + \Im\{P_x\}\Im\{P_x\} - \lambda I_n & \Re\{P_x\}\Im\{P_x\} - \Im\{P_x\}\Re\{P_x\} \\ -\Re\{P_x\}\Im\{P_x\} + \Im\{P_x\}\Re\{P_x\} & \Re\{P_x\}\Re\{P_x\} + \Im\{P_x\}\Im\{P_x\} - \lambda I_n \end{bmatrix}.$$

Therefore (λ is real),

$$\tilde{P}_x \tilde{P}_x - \lambda I_{2n} = \widehat{P_x P_x^H - \lambda I_n}, \quad (14)$$

and furthermore, applying Lemma 2,

$$\det\left(\tilde{P}_x \tilde{P}_x - \lambda I_{2n}\right) = |\det\left(P_x P_x^H - \lambda I_n\right)|^2. \quad (15)$$

The eigenvalues of $\tilde{P}_x \tilde{P}_x$ are the squared eigenvalues of $\tilde{P}_x$ and are equal to the eigenvalues $P_x P_x^H$, which are the squared singular values of P_x. □

3 Differential Entropy of Complex Random Vectors

The *differential entropy* $h(x)$ of a complex random vector x is defined as the differential entropy $h(\hat{x})$ of the corresponding real random vector $\hat{x}$. It is well known, c.f., [5] or [6], that rotationally invariant complex Gaussians are entropy maximizers:

Theorem 3. *Suppose the complex random vector $x \in \mathbb{C}^n$ is zero-mean and has a nonsingular covariance matrix C_x. Then the entropy of x satisfies*[2]

$$h(x) \leq \log\det\left(\pi e C_x\right) \quad (16)$$

with equality if and only if x is rotationally invariant and Gaussian.

Proof. [5], [6]. □

Theorem 3 tells us that for a zero-mean rotationally variant (P_x does not vanish) complex Gaussian random vector x, $h(x) < \log\det\left(\pi e C_x\right)$. Unfortunatley it does not tell us by how much the quantity $h(x)$ is degraded when the pseudo-covariance matrix P_x is non-vanishing. The following theorem gives us the answer:

Theorem 4. *Suppose the complex random vector $x \in \mathbb{C}^n$ is zero-mean Gaussian and has a nonsingular covariance matrix C_x and a pseudo-covariance matrix P_x.*

---

[2] Throughout this paper $\log = \log_2$.

Let $C_x = L_x L_x^H$ (L_x lower triangular) be the Cholesky decomposition (c.f., [8]) of C_x. Then the entropy of x satisfies

$$h(x) = \log\det(\pi e C_x) + \frac{1}{2}\sum_{\lambda_i}\log\left(1-\lambda_i^2\right), \tag{17}$$

where λ_i are the singular values of $P_{L_x^{-1}x} = L_x^{-1}P_x L_x^{-T}$, the pseudo-covariance matrix of the complex random vector $L_x^{-1}x$.

Proof. Perform a Cholesky decomposition (c.f., [8]) $C_x = L_x L_x^H$ (L_x lower triangular) of C_x and consider the complex random vector $y = L_x^{-1}x$ (L_x is non-singular). Obviously, y is Gaussian with covariance matrix $C_y = I_n$ and pseudo-covariance matrix $P_y = P_{L_x^{-1}x}$. This yields ($\hat{x} = \hat{L}_x\hat{y}$, according to Lemma 2)

$$\begin{aligned} h(x) = h(\hat{x}) = & \qquad\qquad (18) \\ &= \frac{1}{2}\log\det(2\pi e C_{\hat{x}}) = \frac{1}{2}\log\det\left(2\pi e C_{\hat{L}_x\hat{y}}\right) = \frac{1}{2}\log\det\left(2\pi e \hat{L}_x C_{\hat{y}}\hat{L}_x^T\right) = \\ &= \frac{1}{2}\log\left(\det(2\pi e C_{\hat{y}})\det\left(\hat{L}_x\hat{L}_x^T\right)\right) = \frac{1}{2}\log\left(\det(2\pi e C_{\hat{y}})\det\hat{C}_x\right) = \\ &= \frac{1}{2}\log\left(\det(2\pi e C_{\hat{y}})(\det C_x)^2\right) = \frac{1}{2}\log\left((\pi e)^{2n}(\det C_x)^2\det(2C_{\hat{y}})\right) = \\ &= \log\det(\pi e C_x) + \frac{1}{2}\log\det(2C_{\hat{y}}) = \log\det(\pi e C_x) + \frac{1}{2}\log\det\left(I_{2n}+\tilde{P}_y\right), \end{aligned}$$

and applying Theorem 1,

$$\begin{aligned} h(x) &= \log\det(\pi e C_x) + \frac{1}{2}\log\det(I_{2n}+\Lambda_y) = \\ &= \log\det(\pi e C_x) + \frac{1}{2}\sum_{\lambda_i>0}\log\left(1-\lambda_i^2\right), \end{aligned}$$

where the sum is over all positive eigenvalues (counted with multiplicity) of $\tilde{P}_{L_x^{-1}x}$, or equivalently (Theorem 2), over all (non-zero) singular values of $P_{L_x^{-1}x}$.
□

Using the inequality $\ln x \leq x-1$, one easily finds a lower and upper bound for the entropy:

Corollary 1. *Suppose we have the same situation as in Theorem 4. Let d_{max} denote the largest eigenvalue of $P_y P_y^H$ ($y = L_x^{-1}x$). Then,*

$$\log\det(\pi e C_x) - \frac{tr\left(P_y P_y^H\right)}{2\ln 2\,(1-d_{max})} \leq h(x) \leq \log\det(\pi e C_x) - \frac{tr\left(P_y P_y^H\right)}{2\ln 2}. \tag{19}$$

Observe that the deviation of the entropy from the "ideal" rotationally invariant entropy is approximately determined by the trace and the largest eigenvalue of $P_y P_y^H$.

Theorem 4 together with the *maximum-entropy theorem* for real random vectors (c.f., [7]) implies Theorem 3. In this sense, it is a generalization, since it is an equality, not an inequality. It also follows from the theorem that the singular values of $P_{L_x^{-1}x}$ have to be smaller than (or equal to) 1. The converse is true as well:

Theorem 5. *Let $C \in \mathbb{C}^{n\times n}$ be an Hermitian positive definite matrix and $C = LL^H$ be its Cholesky decomposition. C and a matrix $P \in \mathbb{C}^{n\times n}$ are the covariance matrix and the pseudo-covariance matrix of a zero-mean complex random vector, respectively, if and only if P is symmetric and the singular values of $L^{-1}PL^{-T}$ are smaller or equal to 1.*

Proof. We only sketch the proof. Consider

$$\hat{C} + \tilde{P} = \hat{L}\hat{L}^T + \tilde{P} = \hat{L}\left(I_{2n} + \hat{L}^{-1}\tilde{P}\hat{L}^{-T}\right)\hat{L}^T \tag{20}$$

and apply[3] Theorem 1 and Theorem 2 to $\hat{L}^{-1}\tilde{P}\hat{L}^{-T}$. Note that $\hat{C} + \tilde{P}$ has to be non-negative definite. □

4 Capacity Loss

In this section, we return to our channel model as presented in the introduction. We intend to calculate the capacity loss if only the information given by the covariance matrix C_n is used, i.e., to calculate the capacity difference between the case where the pseudo-covariance matrix P_n is taken into account and the case where the pseudo-covariance matrix is erroneously assumed to be vanishing. Note that many coding and signal processing algorithms make only use of the covariance matrix C_n.

The mutual information $\mathbf{I}(x;y)$ can be written as

$$\mathbf{I}(x;y) = h(y) - h(y|x) = h(y) - h(n), \tag{21}$$

and thus, maximizing $\mathbf{I}(x;y)$ is equivalent to maximizing $h(y)$. We have $h(y) = h(Ax+n)$. In the case where the pseudo-covariance matrix P_n is neglected, the maximation of $\log\det(\pi e C_y) = \log\det\left(AC_xA^H + C_n\right) + r\log(\pi e)$ yields a zero-mean rotationally invariant complex Gaussian random vector x. Let $C_n = L_nL_n^H$ be the Cholesky decomposition of C_n. With $B = L_n^{-1}A$ and the determinant identity $\det(XY+I) = \det(YX+I)$, one finds

$$\begin{aligned}\log\det\left(AC_xA^H + C_n\right) &= \log\det\left(AC_xA^H + L_nL_n^H\right) = \\ &= \log\det\left(BC_xB^H + I_n\right) + \log\det C_n =\end{aligned}$$

[3] For the converse part, one cannot apply the theorems directly, since they assume that the matrices are (pseudo-)covariance matrices. However, these matrices have the necessary structure used in the corresponding proofs, s.t. the statements about the eigen/singular values are valid for them as well.

$$
\begin{aligned}
&= \log\det\left(C_x B^H B + I_n\right) + \log\det C_n = \\
&= \log\det\left(C_x U D^{\frac{1}{2}} D^{\frac{1}{2}} U^H + I_n\right) + \log\det C_n = \\
&= \log\det\left(D^{\frac{1}{2}} U^H C_x U D^{\frac{1}{2}} + I_n\right) + \log\det C_n = \\
&= \log\det\left(D^{\frac{1}{2}} \bar{C}_x D^{\frac{1}{2}} + I_n\right) + \log\det C_n,
\end{aligned}
$$

where $D^{\frac{1}{2}}$ (diagonal) and U contain the square roots of the eigenvalues and the normalized eigenvectors of $B^H B$, respectively, and $\bar{C}_x = U^H C_x U$. Note that $\operatorname{tr}\left(\bar{C}_x\right) = \operatorname{tr}\left(C_x\right)$, s.t. $\bar{C}_x$ can be found via "water-filling". The random vector y then[4] has a covariance matrix $C_y = A C_x A^H + C_n = A U \bar{C}_x U^H A^H + C_n$ and a pseudo-covariance matrix[5] $P_y = P_n$.

In an high signal-to-noise environment, it is justifiable (see Theorem 5) to assume that it is possible to compensate for the rotationally variant noise, i.e., to maximize the mutual information with the same covariance matrix C_x by an appropriate selection of the pseudo-covariance matrix P_x. Applying Theorem 4, this yields a capacity loss of

$$
\Delta C = C_{\text{neglect } P_n} - C_{\text{utilize} P_n} = \frac{1}{2} \sum_{\lambda_i} \log\left(1 - \lambda_i^2\right), \tag{22}
$$

where λ_i are the singular values of $L_y^{-1} P_n L_y^{-T}$, L_y being the Cholesky factor of $C_y = A U \bar{C}_x U^H A^H + C_n$. Lower and upper bounds are available from Corollary 1.

5 Example

It can be shown[6] that for *Discrete Multitone (DMT)* modulation (baseband *OFDM*, [9]) as used in *ADSL* the noise at the input of the decision device is rotationally variant if the noise at the input of the receiver is colored. Applying (22), the following expression holds for the capacity loss,

$$
\Delta C = \frac{1}{2} \sum_{k=1}^{\frac{N}{2}-1} \log\left(1 - \frac{4\left(\sum_{s=1}^{N-1} R_Z(s) \sin\left(\frac{2\pi}{N} k s\right)\right)^2}{N^2 L^2 \sin^2\left(\frac{2\pi}{N} k\right) \left|H\left(e^{\frac{j 2\pi k}{N}}\right)\right|^4}\right), \tag{23}
$$

where $\frac{N}{2}$ is the number of sub-carriers, $R_Z(s)$ is the autocorrelation function of the zero-mean, discrete-time, real-valued (baseband), stationary Gaussian random noise process at the input of the receiver, $H(z)$ denotes the channel transfer function, and L is the water level (*Water Filling*). Note that L depends on the signal power.

[4] Note that x and n are independent random vectors.
[5] If x is rotationally invariant, so is Ax (c.f., [5]).
[6] Will be addressed in a separate paper.

6 Conclusions

We investigated complex matrix channels with additive complex Gaussian noise. We did not assume rotationally invariant noise and included the corresponding pseudo-covariance matrix into our considerations. For this, we proved a stronger maximum-entropy theorem for the complex multivariate case than previously known, which also takes into account the pseudo-covariance matrix. We derived lower and upper bounds for the deviation of the entropy of a rotationally variant complex Gaussian random vector from the rotationally invariant counterpart. We used these results to find an analytic expression for the capacity loss if the pseudo-covariance matrix is not taken into account, and provided some bounds as well. Finally, we gave an example of a practical, widespread transmission system (xDSL), which fits perfectly in our framework and in which the pseudo-covariance matrix is usually neglected.

References

1. **A. Lampe, R. Schober, W.H. Gerstacker**, "Iterative Multiuser Detection for Complex Modulation Schemes", *Proc. of IEEE International Symposium on Information Theory (ISIT)*, (Washington D.C.), p. 33, 2001.
2. **Y.C. Yoon, H. Leib**, "Maximizing SNR in Improper Complex Noise and Applications in CDMA", *IEEE Communications Letters*, vol. 1, pp. 5-8, Jan. 1997.
3. **S. Buzzi, M. Loops, A. M. Tulino**, "A new Class of Muliuser CDMA Receivers based on The Minimum Mean-Output-Energy Strategy", *Proc. of IEEE International Symposium on Information Theory (ISIT)*, (Naples, Italy), p. 355, June 2000.
4. **B. Picinbono, P.Chevalier**, "Widely Linear Estimation with Complex Data", *IEEE Transactions on Signal Processing*, vol. 43, pp. 2030-2033, Aug. 1995.
5. **F.D. Neeser**, "Communication Theory and Coding for Channels with Intersymbol Interference", *PhD thesis, ETH Zürich, Switzerland*, 1993.
6. **I.Emre Telatar**, "Capacity of Multi-antenna Gaussian Channels", *European Transactions on Telecommunications*, Vol. 10, No. 6, pp. 585-595, Nov/Dec 1999.
7. **T.M. Cover, Y.A. Thomas**, "Elements of Information Theory", *John Wiley & Sons, Inc.*, 1991.
8. **G.H. Golub, C.F. Van Loan**, "Matrix Computations", *North Oxford Academic Publishers Ltd, a subsidiary of Kogan Page Ltd*, p. 16, 1986.
9. **R. Fischer**, "Mehrkanal- und Mehrträgerverfahren für die schnelle digitale Übertragung im Ortsanschlussleitungsnetz", *PhD thesis, Universität Erlangen/Nürnberg, Germany*, 1996.