THE KARHUNEN-LOÈVE EXPANSION OF IMPROPER COMPLEX RANDOM SIGNALS WITH APPLICATIONS IN DETECTION

Peter J. Schreier

Electrical & Computer Engineering
Univ. of Colorado, Boulder, CO 80309-0425
www.peter-schreier.com
peter@peter-schreier.com

Louis L. Scharf

Electrical & Computer Eng. and Statistics
Colorado State University
Ft. Collins, CO 80523
scharf@engr.colostate.edu

ABSTRACT

Non-stationary complex random signals are in general improper (not circularly symmetric), which means that their complementary covariance is non-zero. Since the Karhunen-Loève expansion in its known form is only valid for proper processes, we derive the improper version of this expansion. It produces two sets of eigenvalues and an improper internal description. We use the Karhunen-Loève expansion to solve the problem of detecting non-stationary improper complex random signals in additive white Gaussian noise. Using the deflection criterion we compare the performance of conventional processing, which ignores complementary covariances, with processing that takes these into account. The performance gain can be as big as a factor of 2.

1. INTRODUCTION

Consider the following communications example. Suppose we want to detect a *real* waveform $x(t)$ that is transmitted over a channel that rotates it by some random phase ϕ and adds complex white Gaussian noise $n(t)$. The observations are then given by

$$r(t) = x(t)e^{j\phi} + n(t), \qquad (1)$$

and we shall assume mutual independence of $x(t)$, $n(t)$, and ϕ. Furthermore, denote the rotated signal by $s(t) = x(t)e^{j\phi}$. Its covariance is given by $\gamma_{ss}(t_1,t_2) = Ex(t_1)x(t_2)$. In general, $\gamma_{ss}(t_1,t_2)$ does not give a complete second-order characterization of $s(t)$. It must be complemented by the *complementary covariance* $c_{ss}(t_1,t_2) = Es(t_1)s(t_2) = Ex(t_1)x(t_2)\cdot Ee^{j2\phi}$.

There are, of course, two important special cases. If the phase ϕ is uniformly distributed, then detection is *noncoherent* and $c_{ss}(t_1,t_2) \equiv 0$. Processes that have a vanishing complementary covariance play an important role in signal processing and are called *proper* (sometimes also circularly symmetric). On the other hand, if ϕ is known, then detection is *coherent* and $c_{ss}(t_1,t_2) = \gamma_{ss}(t_1,t_2)$. This means there exists a *real* sufficient statistic, with no need to use complex algebra. However, that leaves us with everything in between these two limiting cases, e.g., phases that are known with some uncertainty (non-uniform phase distribution) or the whole range of adaptive detection techniques. There, the process $s(t)$ is *improper*. The treatment of improper complex processes requires the use of *widely linear* rather than linear transformations. Widely linear transformations also depend linearly on the conjugate of the vector or process that they are applied to [1].

In this paper, we solve the problem of detecting non-stationary improper complex random signals in white Gaussian noise. The essential tool for the treatment of non-stationary signals is the Karhunen-Loève (K–L) expansion. Because the well-known form of this expansion is only valid in the proper case, we develop the improper version of it in Section 2. In Section 3 we solve the problem of detecting an improper signal in additive white Gaussian noise. Deflection will be used to compare conventional processing, which ignores complementary covariances, with processing that takes these into account. We find that widely linear processing of improper processes promises gains up to a factor of 2.

This work was supported by the 1999 NSF Wireless Initiative under contract ECS-9979400, and by the Office of Naval Research under contract N00014-00-1-0033.

2. THE KARHUNEN-LOÈVE EXPANSION

The tools required for dealing with improper complex signals have been presented in a unified framework for the vector and WSS cases in [2]. The basic ideas carry over to a non-stationary setting. The advocated algebra is based on augmented signals $\boldsymbol{\sigma}(t) = [s(t)\ s^*(t)]^\mathsf{T}$ that carry along their conjugate. Their covariance matrix is called the *augmented covariance matrix* and is given by

$$\boldsymbol{\Gamma}_{\sigma\sigma}(t_1,t_2) = E\begin{bmatrix} s(t_1) \\ s^*(t_1)\end{bmatrix}[s^*(t_2), s(t_2)] = \begin{bmatrix} \gamma_{ss}(t_1,t_2) & c_{ss}(t_1,t_2) \\ c^*_{ss}(t_1,t_2) & \gamma^*_{ss}(t_1,t_2)\end{bmatrix}.$$

The augmented covariance matrix gives a complete second-order description of the signal. It belongs to a matrix algebra

$$\mathcal{W} = \left\{ \begin{bmatrix} h_1(t_1,t_2) & h_2(t_1,t_2) \\ h_2^*(t_1,t_2) & h_1^*(t_1,t_2) \end{bmatrix} \middle| h_1, h_2 \in L^2([0,T]^2 \to \mathbb{C}) \right\},$$

which is closed under addition, multiplication, inversion (if inverses exist), and multiplication with a real, but not with a complex scalar. The notation $L^2([0,T]^2 \to \mathbb{C})$ will stand for square-integrable functions defined on $[0,T] \times [0,T]$ that take their values in the complex field. When we apply a member of $\mathcal{W}$ to an augmented signal, we obtain a widely linear transformation [1].

Now, Mercer's Theorem and the K–L expansion for improper complex signals can be stated as follows.

Theorem. Suppose that $s(t)$, $0 \le t \le T$, is a zero-mean second-order complex random process characterized by covariance $\gamma_{ss}(t_1,t_2)$ and complementary covariance $c_{ss}(t_1,t_2)$, which are both continuous on $[0,T]^2$. Then the augmented covariance $\mathbf{\Gamma}_{\sigma\sigma}(t_1,t_2)$ can be expanded in the uniformly and absolutely convergent series

$$\mathbf{\Gamma}_{\sigma\sigma}(t_1,t_2) = \sum_n \mathbf{\Phi}_n(t_1)\mathbf{\Lambda}_n\mathbf{\Phi}_n^{\mathsf{H}}(t_2).$$

where

$$\mathbf{\Lambda}_n = \frac{1}{2}\begin{bmatrix} \lambda_n^r + \lambda_n^i & \lambda_n^r - \lambda_n^i \\ \lambda_n^r - \lambda_n^i & \lambda_n^r + \lambda_n^i \end{bmatrix}$$

involves two non-negative eigenvalues λ_n^r and λ_n^i. The matrix

$$\mathbf{\Phi}_n(t) = \begin{bmatrix} \phi_n(t) & \psi_n(t) \\ \psi_n^*(t) & \phi_n^*(t) \end{bmatrix}$$

satisfies

$$\int_0^T \mathbf{\Phi}_n(t)\mathbf{\Phi}_m^{\mathsf{H}}(t)\,dt = \mathbf{I}\,\delta_{nm}.$$

The matrices $\mathbf{\Lambda}_n$ and $\mathbf{\Phi}_n(t)$ are found as the solutions to the equation

$$\mathbf{\Phi}_n(t_1)\mathbf{\Lambda}_n = \int_0^T \mathbf{\Gamma}_{\sigma\sigma}(t_1,t_2)\mathbf{\Phi}_n(t_2)\,dt_2.$$

Then $s(t)$ can be represented by the mean-square convergent series

$$\boldsymbol{\sigma}(t) = \sum_n \mathbf{\Phi}_n(t)\boldsymbol{\sigma}_n \Leftrightarrow s(t) = \sum_n \phi_n(t)s_n + \psi_n(t)s_n^*,$$

where

$$\boldsymbol{\sigma}_n = \int_0^T \mathbf{\Phi}_n^{\mathsf{H}}(t)\boldsymbol{\sigma}(t)dt \Leftrightarrow s_n = \int_0^T (\phi_n^*(t)s(t) + \psi_n(t)s^*(t))\,dt.$$

The K–L coefficients s_n satisfy

$$E(s_n s_m^*) = \tfrac{1}{2}(\lambda_n^r + \lambda_n^i)\delta_{nm}, \quad E(s_n s_m) = \tfrac{1}{2}(\lambda_n^r - \lambda_n^i)\delta_{nm}.$$

Because the proof of the theorem is long, it is relegated to the Appendix. This theorem is the continuous-time equivalent of the finite-dimensional eigenvalue decomposition for the improper case presented as Proposition 1 in [2]. The main differences between the proper and improper version of this theorem are that there are two sets of eigenvalues $\{\lambda_n^r\}$, $\{\lambda_n^i\}$, and the K–L coefficients s_n are improper. In the proper case, $c_{ss}(t_1,t_2) \equiv 0$, the two eigenvalues λ_n^r and λ_n^i are equal, the K–L coefficients become proper, and the functions $\psi_n(t) \equiv 0 \; \forall n$. Then the K–L expansion simplifies to the known formulas.

To gain more insight into the role of the two sets of eigenvalues, let us return to the communications example of the introduction. In non-coherent detection, $c_{ss}(t_1,t_2) \equiv 0$ and the eigenvalues that belong to $s(t)$ satisfy $\lambda_n^r = \lambda_n^i = \lambda_n$. On the other hand, in the coherent case, $c_{ss}(t_1,t_2) = \gamma_{ss}(t_1,t_2)$ and the eigenvalues that belong to $s(t)$ satisfy $\lambda_n^r = 2\lambda_n$ and $\lambda_n^i = 0$. Therefore, the coherent case is the most improper case under the power constraint $\lambda_n^r + \lambda_n^i = 2\lambda_n$. These comments are clarified by noting that

$$\tfrac{1}{2}\lambda_n^r = E\{s(t)\mathrm{Re}\, s_n\}, \quad \tfrac{j}{2}\lambda_n^i = E\{s(t)\mathrm{Im}\, s_n\}.$$

Thus, λ_n^r measures the correlation between the real part of the internal representation and the continuous-time signal, and λ_n^i does so for the imaginary part. In the non-coherent version of (1), these two correlations are equal, suggesting that the information is carried equally in real and imaginary part of s_n. In the coherent version, $\lambda_n^r = 2\lambda_n$, $\lambda_n^i = 0$ shows that the information is carried exclusively in the real part of s_n, making $\mathrm{Re}\, s_n$ a *sufficient statistic* for the decision on $s(t)$. Therefore, in the coherent problem, widely linear processing amounts to only considering the real part of the internal description. The more interesting applications of widely linear filtering, however, lie either in between the coherent and non-coherent case, characterized by a non-uniform phase-distribution, or in adaptive realizations of coherent algorithms. For further discussion of applications of widely linear filtering in communications refer to [2].

3. DETECTION

The detection problem we would like to solve is a simple hypothesis test:

$$\begin{aligned} H_0 &: r(t) = n(t) \\ H_1 &: r(t) = s(t) + n(t) \end{aligned} \tag{2}$$

We observe the complex signal $r(t)$ over the time interval $0 \le t \le T$. The noise $n(t)$ is zero-mean complex white (i.e., proper) Gaussian with power spectral density N_0, and the zero-mean complex Gaussian signal $s(t)$ is described by its covariance $\gamma_{ss}(t_1,t_2)$ and its complementary covariance $c_{ss}(t_1,t_2)$.

3.1. Solution of the Detection Problem

The first step is to determine the K–L expansion for the signal $s(t)$ as shown above. The K–L coefficients s_n are improper complex Gaussian random variables with $E(s_n s_m^*) = \frac{1}{2}(\lambda_n^r + \lambda_n^i)\delta_{nm}$ and $E(s_n s_m) = \frac{1}{2}(\lambda_n^r - \lambda_n^i)\delta_{nm}$, and the resolution of the observed signal $r(t)$ onto the K–L basis functions is denoted by

$$r_n = \int_0^T (\phi_n^*(t)r(t) + \psi_n(t)r^*(t))\,dt.$$

The log-likelihood ratio is then determined by

$$\begin{aligned} L_R = & \sum_n \underbrace{\frac{\lambda_n^r\lambda_n^i + \frac{N_0}{2}(\lambda_n^r + \lambda_n^i)}{N_0(\lambda_n^r\lambda_n^i + N_0(\lambda_n^r + \lambda_n^i) + N_0^2)}}_{\frac{1}{N_0}h_{n,1}} |r_n|^2 \\ & + \underbrace{\frac{\frac{1}{2}(\lambda_n^r - \lambda_n^i)}{\lambda_n^r\lambda_n^i + N_0(\lambda_n^r + \lambda_n^i) + N_0^2}}_{\frac{1}{N_0}h_{n,2}} \operatorname{Re} r_n^2 \\ = & \frac{1}{N_0}\operatorname{Re}\left(\sum_n \hat{s}_n^* r_n\right), \end{aligned} \tag{3}$$

with

$$\hat{s}_n = h_{n,1} r_n + h_{n,2} r_n^*.$$

A derivation and more detailed discussion of this formula can be found in [4]. It is interesting to note that the widely linear detector (3) also takes on the form of an estimator–correlator. If $s(t)$ is proper, then $\lambda_n^r = \lambda_n^i$, $\psi_n(t) \equiv 0\ \forall n$, and (3) simplifies to the known result

$$L_R = \frac{1}{N_0}\sum_n \frac{\lambda_n}{\lambda_n + N_0}|r_n|^2.$$

An equivalent time-frequency formulation of (3) is derived in [3].

3.2. Performance

To evaluate the performance of the detector, we use the deflection criterion, which is defined as

$$D(L_R) = \frac{(E_1(L_R) - E_0(L_R))^2}{V_0(L_R)}.$$

Here $E_i(\cdot)$ denotes expectation under hypothesis i, and $V_0(\cdot)$ variance under hypothesis zero. The deflection can be regarded as an output signal-to-noise ratio. It can be computed to be

$$D(L_R) = \frac{1}{4N_0^2}\cdot\frac{\left(\sum_n h_{n,1}(\lambda_n^r + \lambda_n^i) + h_{n,2}(\lambda_n^r - \lambda_n^i)\right)^2}{\sum_n (h_{n,1}^2 + h_{n,2}^2)}. \tag{4}$$

Let us determine the deflection for the two limiting cases in the communications example (1). In the coherent problem, using the optimum coefficients $h_{n,1} = h_{n,2} = \frac{\lambda_n}{2\lambda_n + N_0}$, the deflection is

$$D(L_R) = \frac{2}{N_0^2}\cdot\frac{\left(\sum_n \frac{\lambda_n^2}{2\lambda_n + N_0}\right)^2}{\sum_n\left(\frac{\lambda_n}{2\lambda_n + N_0}\right)^2}. \tag{5}$$

On the other hand, if we were to completely ignore the information contained in the complementary covariance, detection would be non-coherent and the coefficients $h'_{n,1} = \frac{\lambda_n}{\lambda_n + N_0}$, $h'_{n,2} = 0$, would only give us a deflection of

$$D'(L_R) = \frac{1}{N_0^2}\cdot\frac{\left(\sum_n \frac{\lambda_n^2}{\lambda_n + N_0}\right)^2}{\sum_n\left(\frac{\lambda_n}{\lambda_n + N_0}\right)^2}. \tag{6}$$

The ratio $D(L_R)/D'(L_R)$ is at most 2, and this is asymptotically the case for either $N_0 \to 0$ or $N_0 \to \infty$. This 3 dB result is certainly not surprising for the comparison of coherent vs. non-coherent detection in the communications example of (1). However, at this point, we conjecture without proof that this factor of 2 is also the *maximum* performance gain possible in any Gauss-Gauss detector that incorporates knowledge of complementary covariances (see [4] for more details).

4. CONCLUSIONS

We have presented a version of the Karhunen-Loève expansion for improper complex random signals. It produces an improper internal description and two sets of eigenvalues. We have used the K–L expansion to solve the problem of detecting non-stationary improper random signals in additive white Gaussian noise. The maximum performance gain of widely linear vs. strictly linear processing as measured by deflection is a factor of 2 [4]. In a communications example, we have linked the two situations that display this maximum performance difference to coherent detection (perfect phase knowledge) and non-coherent detection (no phase information). Widely linear processing, however, can cover any situation in between, where *some* phase knowledge is available, as well as adaptive realizations of coherent detection [2]. A time-frequency perspective of detection of improper signals is developed in [3].

5. APPENDIX: PROOF OF MERCER'S THEOREM AND K–L EXPANSION IN THE IMPROPER CASE

Let $\mathbb{C}_*^2$ be the image of $\mathbb{R}^2$ under the unitary map with matrix

$$\mathbf{T} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix}.$$

The space of augmented signals consists of square-integrable functions $[0,T] \to \mathbb{C}^2_*$ and is denoted by $L^2([0,T] \to \mathbb{C}^2_*)$. This space is not linear (rather it is widely linear), but it is isomorphic to the space of square-integrable maps $[0,T] \to \mathbb{R}^2$. This isomorphism enables us to write down the K–L expansion for an augmented signal, using the results of [5] for vector random processes. Let the assumptions be as in the statement of the theorem. Then the augmented covariance $\mathbf{\Gamma}_{\sigma\sigma}(t_1,t_2)$ can be expanded in the series

$$\mathbf{\Gamma}_{\sigma\sigma}(t_1,t_2) = \sum_n \lambda_n \boldsymbol{\varphi}_n(t_1)\boldsymbol{\varphi}_n^{\mathsf{H}}(t_2), \tag{7}$$

where $\{\lambda_n\}$ are the non-negative scalar eigenvalues and the $\{\boldsymbol{\varphi}_n(t) = [f_n(t), f_n^*(t)]^{\mathsf{T}}\}$ are the corresponding orthonormal eigenfunctions. Each $\boldsymbol{\varphi}_n(t)$ is $L^2([0,T] \to \mathbb{C}^2_*)$. Eigenvalues and eigenfunctions are obtained as solutions to the integral equation

$$\lambda_n \boldsymbol{\varphi}_n(t_1) = \int_0^T \mathbf{\Gamma}_{\sigma\sigma}(t_1,t_2)\boldsymbol{\varphi}_n(t_2)dt_2, \quad 0 \le t_1 \le T,$$

where the eigenfunctions $\boldsymbol{\varphi}_n(t)$ form a complete orthonormal set for $L^2([0,T] \to \mathbb{C}^2_*)$,

$$\int_0^T \boldsymbol{\varphi}_n^{\mathsf{H}}(t)\boldsymbol{\varphi}_m(t)dt = 2\,\mathrm{Re}\int_0^T f_n^*(t) f_m(t)\,dt = \delta_{nm}. \tag{8}$$

Then $s(t)$ can be represented by the mean-square convergent series

$$\boldsymbol{\sigma}(t) = \sum_n x_n \boldsymbol{\varphi}_n(t) \Leftrightarrow s(t) = \sum_n x_n f_n(t),$$

where

$$x_n = \int_0^T \boldsymbol{\varphi}_n^{\mathsf{H}}(t)\boldsymbol{\sigma}(t)dt = 2\,\mathrm{Re}\int_0^T f_n^*(t)s(t)\,dt.$$

The surprising result here is that the K–L coefficients x_n are not only scalars, but they are actually *real* with correlation

$$E\,x_n x_m = \lambda_n \delta_{nm}. \tag{9}$$

The reason for this lies in (8). This equation shows that the functions $f_n(t)$ do not have to be orthogonal in $L^2([0,T] \to \mathbb{C})$ to ensure that the eigenfunctions $\boldsymbol{\varphi}_n(t)$ be orthogonal in $L^2([0,T] \to \mathbb{C}^2_*)$. In fact, it is clear that in general there are more orthogonal augmented functions in $L^2([0,T] \to \mathbb{C}^2_*)$ than there are orthogonal functions in $L^2([0,T] \to \mathbb{C})$. In other words, we were able to reduce the dimension of the internal description (real rather than complex K–L coefficients) because $L^2([0,T] \to \mathbb{C}^2_*)$ allows an increased number of orthonormal eigenfunctions compared to $L^2([0,T] \to \mathbb{C})$. This increase in eigenfunctions is not clearly visible since in both cases we have infinitely many.

From (7) it is not clear how the improper version of Mercer's theorem is connected to its proper version. To make this connection apparent, we re-write (7) as

$$\begin{aligned}
&\mathbf{\Gamma}_{\sigma\sigma}(t_1,t_2) = \\
&= \sum_n \left([\boldsymbol{\varphi}_{2n}(t_1), \boldsymbol{\varphi}_{2n+1}(t_1)]\mathbf{T}^{\mathsf{H}}\right)\left(\mathbf{T}\begin{bmatrix}\lambda_{2n} & 0\\ 0 & \lambda_{2n+1}\end{bmatrix}\mathbf{T}^{\mathsf{H}}\right) \times \\
&\qquad \times \left(\mathbf{T}\begin{bmatrix}\boldsymbol{\varphi}_{2n}^{\mathsf{H}}(t_2)\\ \boldsymbol{\varphi}_{2n+1}^{\mathsf{H}}(t_2)\end{bmatrix}\right) \\
&= \sum_n \begin{bmatrix}\phi_n(t_1) & \psi_n(t_1)\\ \psi_n^*(t_1) & \phi_n^*(t_1)\end{bmatrix}\begin{bmatrix}\frac{1}{2}(\lambda_{2n}+\lambda_{2n+1}) & \frac{1}{2}(\lambda_{2n}-\lambda_{2n+1})\\ \frac{1}{2}(\lambda_{2n}-\lambda_{2n+1}) & \frac{1}{2}(\lambda_{2n}+\lambda_{2n+1})\end{bmatrix} \times \\
&\qquad \times \begin{bmatrix}\phi_n^*(t_2) & \psi_n(t_2)\\ \psi_n^*(t_2) & \phi_n(t_2)\end{bmatrix} \\
&= \sum_n \mathbf{\Phi}_n(t_1)\mathbf{\Lambda}_n\mathbf{\Phi}_n^{\mathsf{H}}(t_2).
\end{aligned}$$

Thus, the internal representation is now given by *complex* K–L coefficients $s_n = \frac{1}{\sqrt{2}}(x_{2n} + jx_{2n+1})$ and

$$\boldsymbol{\sigma}_n = \begin{bmatrix}s_n\\ s_n^*\end{bmatrix} = \mathbf{T}\begin{bmatrix}x_{2n}\\ x_{2n+1}\end{bmatrix} = \int_0^T \mathbf{\Phi}_n^{\mathsf{H}}(t)\boldsymbol{\sigma}(t)\,dt.$$

For these coefficients we find because of (9)

$$\begin{aligned}
E(s_n s_m) =& \tfrac{1}{2}[E(x_{2n}x_{2m}) - E(x_{2n+1}x_{2m+1}) \\
&+ j(E(x_{2n}x_{2m+1}) + E(x_{2n+1}x_{2m}))] \\
=& \tfrac{1}{2}(\lambda_{2n} - \lambda_{2n+1})\delta_{nm}
\end{aligned}$$

and, similarly, $E(s_n s_m^*) = \frac{1}{2}(\lambda_{2n} + \lambda_{2n+1})\delta_{nm}$. If we now agree on the definitions $\lambda_n^r \triangleq \lambda_{2n}$, $\lambda_n^i \triangleq \lambda_{2n+1}$, the proof is complete.

6. REFERENCES

[1] B. Picinbono, P. Chevalier, "Widely Linear Estimation with Complex Data," *IEEE Trans. Signal Processing*, vol. 43, no. 8, pp. 2030 – 2033, Aug. 1995

[2] P. J. Schreier, L. L. Scharf, "Second-Order Analysis of Improper Complex Random Vectors and Processes," accepted for publication, *IEEE Trans. Signal Processing*

[3] P. J. Schreier, L. L. Scharf, "Stochastic Time-Frequency Analysis Using the Analytic Signal: Why the Complementary Distribution Matters," submitted to *IEEE Trans. Signal Processing*

[4] P. J. Schreier, L. L. Scharf, C. T. Mullis, "Detection and Estimation of Non-Stationary Improper Complex Random Signals," submitted to *IEEE Trans. Inform. Theory*

[5] E. J. Kelly, W. L. Root, "A Representation of Vector-Valued Random Processes," *Group Rept. 55-21*, MIT Lincoln Laboratory, April 1960