COMPOSITE REAL PRINCIPAL COMPONENT ANALYSIS OF COMPLEX SIGNALS

Christoph Hellings, Patrick Gögler, and Wolfgang Utschick

Fachgebiet Methoden der Signalverarbeitung, Technische Universität München, 80290 München, Germany
Phone: +49 89 289 28516, Fax: +49 89 289 28522, {hellings, patrick.goegler, utschick}@tum.de

ABSTRACT

Principal component analysis (PCA) is a tool for dimensionality reduction, feature extraction, and data compression, which is applied to both real-valued and complex-valued data sets. For complex data, a modified version of PCA based on widely linear transformations was shown to be beneficial if the considered random variables are improper, i.e., in the case of correlations or power imbalances between real and imaginary parts. This widely linear approach is formulated in an augmented complex representation in the existing literature. In this paper, we propose a composite real PCA, which instead transforms the complex data into a set of real-valued principal components. This alternative approach is superior in dimensionality reduction due to the finer granularity that is possible when counting dimensions in the real-valued representation. Moreover, it can be used to obtain the same results as the augmented complex version at a lower computational complexity.

***Index Terms*—** Composite real representation, dimensionality reduction, improper signals, noncircular, principal component analysis (PCA).

1. INTRODUCTION

Complex random variables are often implicitly assumed to be *proper* [1], which is the case if the real and imaginary parts are uncorrelated and have the same variance. However, there are many situations of practical importance in which this requirement is not fulfilled (see, e.g., [2, 3]). Then, the random variables are called *improper*, and special attention is needed when developing methods to process or analyze such signals.

Methods for complex-valued signal processing are often derived based on existing approaches for real-valued signals. When transferring such a method to a complex scenario, linear operations are usually translated to linear operations in the complex domain. However, whenever at least one improper signal is present, complex linear operations might not exploit the full potential of the system, and an extension to so-called *widely linear processing* can be beneficial (e.g., [2]).

This also applies to principal component analysis (PCA), which has become a standard tool for dimensionality reduction, feature extraction, and data compression. Even though it was originally derived for real-valued data sets or sets of real-valued random variables in [4, 5], it is straightforward to apply PCA as a (strictly) linear rank reduction technique for complex random vectors (e.g., [2, 6, 7]). A modified linear rank reduction for improper signals was proposed in [8]. However, a better concentration of the signal variance in the first few principal components is achieved if a widely linear transformation is used to perform the rank reduction [2, 6]. For extensions to this widely linear PCA, see, e.g., [9].

A mapping is called widely linear if it is a linear function of the signal and of the complex conjugate of the signal [10]. As a convenient way to write such mappings, a so-called *augmented complex representation* has been favored by many researchers (e.g., [2, 3, 11]). However, since a widely linear mapping is equivalent to a linear function of the real and imaginary parts of the complex signal, an alternative description of widely linear mappings can be given by introducing the so-called *composite real representation* (e.g., [11,12]), where the real and imaginary parts are stacked in a real-valued vector with twice the dimension of the complex signal. Then, a signal processing method for a complex scenario can be obtained by applying an existing algorithm for real-valued settings to the composite real representation. This possibility has recently been exploited in many research papers (e.g., in the context of signal processing for communications in [12–15]).

In this paper, we propose a new PCA method by applying the traditional real-valued PCA to composite real representations of complex signals. After summarizing fundamentals of the augmented complex and composite real representations in Section 2 and revisiting various existing linear and widely linear versions of PCA in Section 3, we establish this new composite real PCA in Section 4.

Depending on the data set under consideration, the composite real PCA either achieves the same performance as the widely linear PCA using the augmented complex formulation (see Section 5) or it even outperforms the augmented complex counterpart in terms of dimensionality reduction potential. As discussed in detail by means of a numerical example in Section 6, the latter is the case whenever the composite real dimensionality reduction leads to an odd number of real-valued principal components. Especially if a maximally improper (e.g., real-valued) signal contributes to the complex signal under consideration, the composite real PCA is beneficial. An additional advantage of the composite real PCA in terms of computational complexity is discussed in Section 7.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

2. MATHEMATICAL PRELIMINARIES

Let $\underset{\sim}{x} = x_\mathrm{R} + \mathrm{j}\, x_\mathrm{I} \in \mathbb{C}^N$ with $x_\mathrm{R}, x_\mathrm{I} \in \mathbb{R}^N$ be a complex random vector (tildes $\underset{\sim}{\bullet}$ denote complex quantities). Then,

$$\underline{\underset{\sim}{x}} = \begin{bmatrix} \underset{\sim}{x} \\ \underset{\sim}{x}^* \end{bmatrix} \in \mathbb{C}^{2N} \qquad \text{and} \qquad \check{x} = \begin{bmatrix} x_\mathrm{R} \\ x_\mathrm{I} \end{bmatrix} \in \mathbb{R}^{2N} \tag{1}$$

are its augmented complex and its composite real representation, respectively. For the ease of notation, we assume all random vectors to be zero-mean without loss of generality.

The second-order statistical properties of a complex random vector $\underset{\sim}{x}$ are characterized by the conventional covariance matrix $\underset{\sim}{C}_{\underset{\sim}{x}} = \mathrm{E}[\underset{\sim}{x}\underset{\sim}{x}^\mathrm{H}]$ and the so-called *pseudocovariance matrix* $\tilde{\underset{\sim}{C}}_{\underset{\sim}{x}} = \mathrm{E}[\underset{\sim}{x}\underset{\sim}{x}^\mathrm{T}]$ [1]. If the pseudocovariance vanishes, i.e., $\tilde{\underset{\sim}{C}}_{\underset{\sim}{x}} = \mathbf{0}$, the random vector is called proper.

As an alternative to considering these two matrices, a complete second-order characterization is also obtained by considering either the *augmented covariance matrix* [2]

$$\underline{\underset{\sim}{C}}_{\underline{\underset{\sim}{x}}} = \mathrm{E}\left[\underline{\underset{\sim}{x}}\,\underline{\underset{\sim}{x}}^\mathrm{H}\right] = \begin{bmatrix} \underset{\sim}{C}_{\underset{\sim}{x}} & \tilde{\underset{\sim}{C}}_{\underset{\sim}{x}} \\ \tilde{\underset{\sim}{C}}^*_{\underset{\sim}{x}} & \underset{\sim}{C}^*_{\underset{\sim}{x}} \end{bmatrix} \in \mathbb{C}^{2N\times 2N} \tag{2}$$

or the *composite real covariance matrix* (e.g., [2, 12])

$$C_{\check{x}} = \mathrm{E}\left[\check{x}\check{x}^\mathrm{T}\right] = \begin{bmatrix} C_{x_\mathrm{R}} & C_{x_\mathrm{R},x_\mathrm{I}} \\ C^\mathrm{T}_{x_\mathrm{R},x_\mathrm{I}} & C_{x_\mathrm{I}} \end{bmatrix} \in \mathbb{R}^{2N\times 2N} \tag{3}$$

where $C_{x_\mathrm{R},x_\mathrm{I}} = \mathrm{E}[x_\mathrm{R}x_\mathrm{I}^\mathrm{T}]$ is the cross-covariance matrix between the real and imaginary parts x_R and x_I.

A complex mapping $\underset{\sim}{f} : \underset{\sim}{x} \mapsto \underset{\sim}{f}(\underset{\sim}{x})$ is called widely linear if it can be written in the form (e.g., [2, 11])

$$\underset{\sim}{f}(\underset{\sim}{x}) = \underset{\sim}{A}_\mathrm{L}\underset{\sim}{x} + \underset{\sim}{A}_\mathrm{CL}\underset{\sim}{x}^* \tag{4}$$

where the subscripts of the factors $\underset{\sim}{A}_\mathrm{L}$ and $\underset{\sim}{A}_\mathrm{CL}$ stand for *linear* and *conjugate linear*, respectively. This is equivalent to a linear real-valued mapping $\check{f} : \check{x} \mapsto \check{f}(\check{x}) = A\check{x}$ applied to the composite real representation of $\underset{\sim}{x}$, where [12]

$$A = \begin{bmatrix} \Re\left(\underset{\sim}{A}_\mathrm{L}\right) & -\Im\left(\underset{\sim}{A}_\mathrm{L}\right) \\ \Im\left(\underset{\sim}{A}_\mathrm{L}\right) & \Re\left(\underset{\sim}{A}_\mathrm{L}\right) \end{bmatrix} + \begin{bmatrix} \Re\left(\underset{\sim}{A}_\mathrm{CL}\right) & \Im\left(\underset{\sim}{A}_\mathrm{CL}\right) \\ \Im\left(\underset{\sim}{A}_\mathrm{CL}\right) & -\Re\left(\underset{\sim}{A}_\mathrm{CL}\right) \end{bmatrix}. \tag{5}$$

3. LINEAR AND WIDELY LINEAR PCA

In [4], PCA was formulated as a minimization of the mean squared error between the original set of random variables and their orthogonal projections onto a subspace. The following considerations are instead based on the method of [5], which is known to lead to equivalent results (e.g., [16, Ch. 6]).

Given a vector $\underset{\sim}{x} = [\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N]^\mathrm{T}$ of correlated random variables, the goal of PCA is to find a mapping $\underset{\sim}{f}$ that extracts a vector of uncorrelated random variables $[\underset{\sim}{\zeta}_1, \ldots, \underset{\sim}{\zeta}_N]^\mathrm{T} = \underset{\sim}{\zeta} = \underset{\sim}{f}(\underset{\sim}{x})$ such that as much randomness as possible is covered by the first ℓ variables $\underset{\sim}{\zeta}_1, \ldots, \underset{\sim}{\zeta}_\ell$ for any $\ell \leq N$. The various versions of PCA that are considered in this section differ in the assumptions on the vector $\underset{\sim}{x}$ and the mapping $\underset{\sim}{f}$.

3.1. Linear PCA for real-valued random vectors

Let us first consider a real-valued random vector $x \in \mathbb{R}^N$. The classical PCA for this setting assumes a linear mapping $f : x \mapsto Ax$ with $A \in \mathbb{R}^{N\times N}$ [4, 5]. Nonlinear extensions exist in the literature (e.g., [16]), but are beyond the scope of this paper. The linear version can be written as

$$\max_{A} \sum_{k=1}^{\ell} \mathrm{E}\left[\xi_k^2\right] \quad \text{s.t.} \quad \mathrm{E}\left[\xi_i\xi_j\right] = 0 \;\; \forall (i,j) : i \neq j \tag{6}$$

for all $\ell \leq N$, where ξ_k is the kth entry of $\boldsymbol{\xi} = Ax$. Conceptually, this is a multiobjective optimization (different objective function for each choice of ℓ), but it turns out that there exists a solution that is simultaneously optimal for all $\ell \leq N$.

Let $U\Lambda U^\mathrm{T} = C_x$ be the ordered eigenvalue decomposition of the covariance matrix of x such that the entries λ_i of the diagonal matrix Λ are in descending order. The optimal linear mapping is then found by choosing $A = U^\mathrm{T}$ (e.g., [16]). Due to $C_{\boldsymbol{\xi}} = AC_xA^\mathrm{T} = \Lambda$, this yields $\mathrm{E}[\xi_i\xi_j] = \lambda_i$ if $i = j$ and zero otherwise. The sum of the first ℓ variances $\sum_{k=1}^{\ell} \lambda_k$ is maximized for any choice of ℓ due to the ordering of the eigenvalue matrix Λ. Optimality of this solution can formally be shown, e.g., by successive optimization of the rows of A (e.g., [16]) or via majorization theory (e.g., [2]).

3.2. Linear PCA for complex random vectors

A straightforward extension to complex random vectors $\underset{\sim}{x}$ is obtained by restricting the mapping $\underset{\sim}{f}$ to be (strictly) linear (e.g., [2, 6, 7]), i.e., to demand $\underset{\sim}{A}_\mathrm{CL} = \mathbf{0}$ in (4). We then have

$$\max_{\underset{\sim}{A}_\mathrm{L}} \sum_{k=1}^{\ell} \mathrm{E}\left[|\underset{\sim}{\zeta}_k|^2\right] \quad \text{s.t.} \quad \mathrm{E}\left[\underset{\sim}{\zeta}_i\underset{\sim}{\zeta}_j^*\right] = 0 \;\; \forall (i,j) : i \neq j \tag{7}$$

for all $\ell \leq N$, where $\underset{\sim}{\zeta}_k$ is the kth component of $\underset{\sim}{\zeta} = \underset{\sim}{A}_\mathrm{L}\underset{\sim}{x}$. The optimal solution is $\underset{\sim}{A}_\mathrm{L} = \underset{\sim}{U}^\mathrm{H}$ where $\underset{\sim}{U}$ is the unitary modal matrix from the ordered eigenvalue decomposition $\underset{\sim}{U}\Phi\underset{\sim}{U}^\mathrm{H} = \underset{\sim}{C}_{\underset{\sim}{x}}$ with decreasing diagonal elements φ_k [2].

3.3. Widely linear PCA for complex random vectors

A relaxed version of the optimization in the previous Subsection is obtained by allowing $\underset{\sim}{f}$ to be widely linear. In [2, 6], it was shown that for an improper complex random vector $\underset{\sim}{x}$, a better concentration of the variance in the first principal components can be achieved with the widely linear formulation

$$\max_{\underset{\sim}{A}_\mathrm{L}, \underset{\sim}{A}_\mathrm{CL}} \sum_{k=1}^{\ell} \mathrm{E}\left[|\underset{\sim}{\zeta}_k|^2\right] \quad \text{s.t.} \quad \mathrm{E}\left[\underline{\underset{\sim}{\zeta}}_i\underline{\underset{\sim}{\zeta}}_j^\mathrm{H}\right] = \mathbf{0} \;\; \forall (i,j) : i \neq j \tag{8}$$

for all $\ell \leq N$, where $\underset{\sim}{\zeta}_k$ is the kth entry of $\underset{\sim}{\zeta} = \underset{\sim}{A}_\mathrm{L}\underset{\sim}{x} + \underset{\sim}{A}_\mathrm{CL}\underset{\sim}{x}^*$. Due to the constraint, $\mathrm{E}[\underset{\sim}{\zeta}_i\underset{\sim}{\zeta}_j^*] = 0 = \mathrm{E}[\underset{\sim}{\zeta}_i\underset{\sim}{\zeta}_j]$ for $i \neq j$.

The optimal solution can be obtained via the augmented eigenvalue decomposition $\underline{\underset{\sim}{U}}\,\underline{\Psi}\,\underline{\underset{\sim}{U}}^\mathrm{H} = \underline{\underset{\sim}{C}}_{\underline{\underset{\sim}{x}}}$ where

$$\underline{\underset{\sim}{U}} = \begin{bmatrix} \underset{\sim}{U}_\mathrm{L} & \underset{\sim}{U}_\mathrm{CL} \\ \underset{\sim}{U}^*_\mathrm{CL} & \underset{\sim}{U}^*_\mathrm{L} \end{bmatrix} \text{ and } \underline{\Psi} = \frac{1}{2}\begin{bmatrix} \Psi_1 + \Psi_2 & \Psi_1 - \Psi_2 \\ \Psi_1 - \Psi_2 & \Psi_1 + \Psi_2 \end{bmatrix}. \tag{9}$$

Here, $\boldsymbol{\Psi}_1 = \mathrm{diag}_{k=1}^{N}(\psi_{2k-1})$ and $\boldsymbol{\Psi}_2 = \mathrm{diag}_{k=1}^{N}(\psi_{2k})$, where $\psi_1, \ldots, \psi_{2N}$ are the eigenvalues of the augmented covariance matrix $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$ sorted in descending order [2, 6]. $\underline{\boldsymbol{U}}$ is a so-called widely unitary matrix satisfying $\underline{\boldsymbol{U}}\,\underline{\boldsymbol{U}}^{\mathrm{H}} = \underline{\boldsymbol{U}}^{\mathrm{H}}\underline{\boldsymbol{U}} = \mathbf{I}_{2N}$. The optimal solution of (8) is then given by $\underset{\sim}{\boldsymbol{A}}_{\mathrm{L}} = \underset{\sim}{\boldsymbol{U}}_{\mathrm{L}}^{\mathrm{H}}$ and $\underset{\sim}{\boldsymbol{A}}_{\mathrm{CL}} = \underset{\sim}{\boldsymbol{U}}_{\mathrm{CL}}^{\mathrm{T}}$, and the variance and pseudovariance of the kth principal component are given by [2, 6]

$$\mathrm{E}\left[|\underset{\sim}{\zeta}_k|^2\right] = \frac{\psi_{2k-1} + \psi_{2k}}{2}, \quad \mathrm{E}\left[\underset{\sim}{\zeta}_k^2\right] = \frac{\psi_{2k-1} - \psi_{2k}}{2}. \quad (10)$$

Note that the augmented eigenvalue matrix $\underline{\boldsymbol{\Psi}}$ is generally not diagonal [2], which seems a bit counterintuitive at first glance. We will later see that the composite real formulation allows a more intuitive derivation of the widely linear PCA.

4. COMPOSITE REAL PCA

After this summary of existing versions of PCA, we now propose a composite real PCA of complex random vectors. Even though the basic idea of the composite real PCA is rather simple, the subsequent analysis shows that this new approach has significant advantages compared to the existing ones.

Consider the composite real representation $\check{\boldsymbol{x}} \in \mathbb{R}^{2N}$ of a complex random vector $\underset{\sim}{\boldsymbol{x}} \in \mathbb{C}^N$ as given in (1). By applying the real-valued PCA from Section 3.1 to $\check{\boldsymbol{x}}$, we obtain

$$\max_{\boldsymbol{A}} \sum_{k=1}^{L} \mathrm{E}\left[\xi_k^2\right] \quad \text{s.t.} \quad \mathrm{E}\left[\xi_i \xi_j\right] = 0 \;\; \forall (i,j) : i \neq j \quad (11)$$

for all $L \leq 2N$, where ξ_k is the kth entry of $\boldsymbol{\xi} = \boldsymbol{A}\check{\boldsymbol{x}} \in \mathbb{R}^{2N}$. Note that the vector $\boldsymbol{\xi}$ now contains $2N$ elements, i.e., we have transformed the N-dimensional complex random vector $\underset{\sim}{\boldsymbol{x}}$ into a set of $2N$ real-valued principal components $\xi_1, \ldots, \xi_{2N}$ instead of into N complex principal components.

From Section 3.1, it is clear that the solution is given by $\boldsymbol{A} = \boldsymbol{U}^{\mathrm{T}} \in \mathbb{R}^{2N \times 2N}$, where the orthogonal matrix $\boldsymbol{U}$ is the modal matrix from the ordered eigenvalue decomposition $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{C}_{\check{\boldsymbol{x}}}$. The variances $\mathrm{E}[\xi_1^2], \ldots, \mathrm{E}[\xi_{2N}^2]$ of the real-valued principal components are equal to the eigenvalues $\lambda_1, \ldots, \lambda_{2N}$, which are arranged in descending order.

Note that we have used a conventional eigenvalue decomposition of a real symmetric matrix, which is conceptually easier than the augmented eigenvalue decomposition in the augmented complex approach from [2, 6] (see Section 3.3).

5. COMPARISON TO COMPLEX VERSIONS OF PCA

We compare the real-valued principal components to the complex principal components obtained using the widely linear PCA (Section 3.3) or the complex linear PCA (Section 3.2).

5.1. Comparison to the widely linear PCA

As each of the N complex principal components can be represented as $\underset{\sim}{\zeta}_k = \zeta_{k,\mathrm{R}} + \mathrm{j}\,\zeta_{k,\mathrm{I}}$, we arrange the $2N$ real-valued principal components in pairs for the sake of comparison. Thus, let $\underset{\sim}{\xi}_k = \xi_{2k-1} + \mathrm{j}\,\xi_{2k}$ for $k = 1, \ldots, N$. Since the real-valued principal components are uncorrelated, we have

$$\mathrm{E}\left[|\underset{\sim}{\xi}_k|^2\right] = \mathrm{E}\left[\xi_{2k-1}^2\right] + \mathrm{E}\left[\xi_{2k}^2\right] = \lambda_{2k-1} + \lambda_{2k} \quad (12)$$

$$\begin{aligned}\mathrm{E}\left[\underset{\sim}{\xi}_k^2\right] &= \mathrm{E}\left[\xi_{2k-1}^2\right] + 2\,\mathrm{j}\,\mathrm{E}\left[\xi_{2k-1}\xi_{2k}\right] + \mathrm{j}^2\,\mathrm{E}\left[\xi_{2k}^2\right] \\ &= \lambda_{2k-1} - \lambda_{2k} \end{aligned} \quad (13)$$

and $\mathrm{E}[\underset{\sim}{\xi}_i \underset{\sim}{\xi}_j^*] = 0 = \mathrm{E}[\underset{\sim}{\xi}_i \underset{\sim}{\xi}_j]$ for $i \neq j$.

According to [2], the ordered eigenvalues $\lambda_1, \ldots, \lambda_{2N}$ of the composite real covariance matrix $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ and the ordered eigenvalues $\psi_1, \ldots, \psi_{2N}$ of the augmented covariance matrix $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$ fulfill $[\psi_1, \ldots, \psi_{2N}] = 2[\lambda_1, \ldots, \lambda_{2N}]$ for any complex random vector $\underset{\sim}{\boldsymbol{x}}$. Thus, comparing (10) to (12) and (13) shows that $\underset{\sim}{\xi}_1, \ldots, \underset{\sim}{\xi}_N$ and the complex principal components $\underset{\sim}{\zeta}_1, \ldots, \underset{\sim}{\zeta}_N$ share the same second-order statistical properties.

Thus, by restricting L in (11) to even numbers, we obtain an alternative derivation of the widely linear PCA from Section 3.3, without making use of an augmented eigenvalue decomposition. Note that the equivalence of the two approaches is plausible since a complex widely linear transformation corresponds to a real linear transformation, i.e., equally powerful mappings $\underset{\sim}{\boldsymbol{f}}$ are allowed in both cases. This is different from complex linear approaches to rank reduction (e.g., [7, 8]), which allow only a smaller class of transformations.

However, since there is no need to restrict L to being even, the composite real PCA is more general than the augmented complex widely linear PCA. In Section 6, we show that the increased flexibility obtained by this generalization can lead to an improved dimensionality reduction capability. Before doing so, we compare the composite real PCA also to the complex linear PCA for the sake of a complete picture.

5.2. Comparison to the (strictly) linear PCA

Due to the equivalence of the widely linear PCA in the augmented complex representation and the composite real PCA with even values of L, the following derivation delivers an alternative proof of the observations in [2, 6]. Therein, it was shown that the widely linear version outperforms the (strictly) linear PCA in case of improper complex random vectors.

Any composite real covariance matrix can be written as

$$\begin{aligned}\boldsymbol{C}_{\check{\boldsymbol{x}}} &= \grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} \quad (14) \\ &= \frac{1}{2}\begin{bmatrix}\Re(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}) & -\Im(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}) \\ \Im(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}) & \Re(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}})\end{bmatrix} + \frac{1}{2}\begin{bmatrix}\Re(\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}) & \Im(\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}) \\ \Im(\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}) & -\Re(\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}})\end{bmatrix}\end{aligned}$$

[12], where $\grave{\boldsymbol{P}}$ is a so-called *block-skew-circulant* ($\mathcal{BSC}_2$) matrix (denoted by a grave accent), and $\acute{\boldsymbol{N}}$ is a so-called *block-Hankel-skew-circulant* ($\mathcal{BHSC}_2$) matrix (denoted by an acute accent). It can be shown (using [12, Lemma 20]) that $\grave{\boldsymbol{P}}$ has the same eigenvalues as the scaled complex covariance matrix $\frac{1}{2}\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$, but the multiplicity of each eigenvalue is doubled.

If the complex random vector $\underset{\sim}{\boldsymbol{x}}$ is proper, the $\mathcal{BHSC}_2$

component $\acute{\boldsymbol{N}}$, which is related to the pseudocovariance matrix $\widetilde{\underaccent{\tilde}{\boldsymbol{C}}}_{\underaccent{\tilde}{\boldsymbol{x}}}$, vanishes, so that $\boldsymbol{C}_{\check{\boldsymbol{x}}} = \grave{\boldsymbol{P}}$. Then, the ordered eigenvalues $\lambda_1, \ldots, \lambda_{2N}$ fulfill $\lambda_{2k-1} = \lambda_{2k} = \frac{1}{2}\varphi_k$, where φ_k is the kth largest eigenvalue of $\underaccent{\tilde}{\boldsymbol{C}}_{\underaccent{\tilde}{\boldsymbol{x}}}$. As a consequence, (12) simplifies to $\mathrm{E}[|\xi_k|^2] = \varphi_k$ for proper random vectors $\underaccent{\tilde}{\boldsymbol{x}}$, which is equal to the variance of the kth complex principal component $\underaccent{\tilde}{\zeta}_k$ obtained with the (strictly) linear PCA from Section 3.2.

To study the case where $\underaccent{\tilde}{\boldsymbol{x}}$ is improper, we make use of

$$\grave{\boldsymbol{J}} = \begin{bmatrix} \mathbf{0} & -\mathbf{I}_N \\ \mathbf{I}_N & \mathbf{0} \end{bmatrix}, \quad \grave{\boldsymbol{J}}^{-1} = \grave{\boldsymbol{J}}^{\mathrm{T}} \tag{15}$$

for which we have $\grave{\boldsymbol{P}} = \grave{\boldsymbol{J}}^{\mathrm{T}} \grave{\boldsymbol{P}} \grave{\boldsymbol{J}}$ and $\acute{\boldsymbol{N}} = -\grave{\boldsymbol{J}}^{\mathrm{T}} \acute{\boldsymbol{N}} \grave{\boldsymbol{J}}$ [12, Lemma 10]. If λ is an eigenvalue of $\boldsymbol{C}_{\check{\boldsymbol{x}}}$, and $\boldsymbol{q}$ is the corresponding eigenvector, i.e., $(\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}})\boldsymbol{q} = \boldsymbol{q}\lambda$, we have that

$$(\grave{\boldsymbol{J}}^{\mathrm{T}} \grave{\boldsymbol{P}} \grave{\boldsymbol{J}} - \grave{\boldsymbol{J}}^{\mathrm{T}} \acute{\boldsymbol{N}} \grave{\boldsymbol{J}})\boldsymbol{q} = \boldsymbol{q}\lambda \;\Leftrightarrow\; (\grave{\boldsymbol{P}} - \acute{\boldsymbol{N}})\grave{\boldsymbol{J}}\boldsymbol{q} = \grave{\boldsymbol{J}}\boldsymbol{q}\lambda \tag{16}$$

i.e., $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}}$ and $\grave{\boldsymbol{P}} - \acute{\boldsymbol{N}}$ have the same eigenvalues. Thus,

$$\begin{bmatrix} \grave{\boldsymbol{P}} & \acute{\boldsymbol{N}} \\ \acute{\boldsymbol{N}} & \grave{\boldsymbol{P}} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \mathbf{I}_N & \mathbf{I}_N \\ \mathbf{I}_N & -\mathbf{I}_N \end{bmatrix} \begin{bmatrix} \grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} & \\ & \grave{\boldsymbol{P}} - \acute{\boldsymbol{N}} \end{bmatrix} \begin{bmatrix} \mathbf{I}_N & \mathbf{I}_N \\ \mathbf{I}_N & -\mathbf{I}_N \end{bmatrix} \tag{17}$$

has the same eigenvalues as $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} = \boldsymbol{C}_{\check{\boldsymbol{x}}}$, but the multiplicity of each eigenvalue is doubled.

Note that the block-diagonal matrix $\mathrm{blockdiag}(\grave{\boldsymbol{P}}, \grave{\boldsymbol{P}})$ has the same eigenvalues as the matrix $\frac{1}{2}\underaccent{\tilde}{\boldsymbol{C}}_{\underaccent{\tilde}{\boldsymbol{x}}}$, but with quadrupled multiplicity. From [17, Ch. 9], we can conclude that the vector of eigenvalues of $\mathrm{blockdiag}(\grave{\boldsymbol{P}}, \grave{\boldsymbol{P}})$ is majorized by the vector of eigenvalues $[\lambda'_1, \ldots, \lambda'_{4N}]$ of the matrix in (17), i.e., the partial sum of the first $K \leq 4N$ eigenvalues of the matrix in (17) is larger than or equal to the respective partial sum for the block-diagonal matrix with equality for $K = 4N$.

Thus, the sorted eigenvalues $\lambda_1, \ldots, \lambda_{2N}$ of $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ fulfill

$$\begin{aligned} \sum_{k=1}^{\ell} (\lambda_{2k-1} + \lambda_{2k}) &= \frac{1}{2} \sum_{k=1}^{\ell} (\lambda'_{4k-3} + \lambda'_{4k-2} + \lambda'_{4k-1} + \lambda'_{4k}) \\ &\geq \frac{1}{2} \sum_{k=1}^{\ell} 4(\tfrac{1}{2}\varphi_k) = \sum_{k=1}^{\ell} \varphi_k \end{aligned} \tag{18}$$

where the inequality is due to the majorization. This shows that the composite real PCA covers a larger share of total variance with the first 2ℓ real-valued principal components than the complex linear PCA does with the first ℓ complex principal components. Equality holds for proper random vectors $\underaccent{\tilde}{\boldsymbol{x}}$ and in any case for $\ell = N$.

6. DIMENSIONALITY REDUCTION CAPABILITY

Up to this point, we have seen that the composite real PCA with even values of L in (11) is equivalent to the widely linear PCA and outperforms the complex linear PCA. In the following, we turn our attention to cases in which the composite real PCA even outperforms the augmented complex version of the widely linear PCA. In particular, we exploit the additional flexibility of an arbitrary (even or odd) integer L to obtain an improved dimensionality reduction. To demonstrate the improvement, we use a rather simple example, but the same effect happens in more complicated scenarios as well.

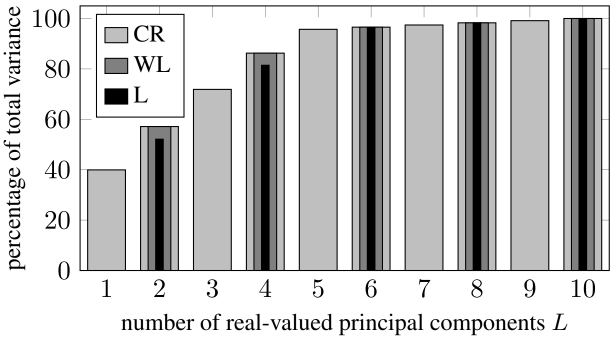

Fig. 1. Composite real PCA (CR), augmented complex widely linear PCA (WL), and complex linear PCA (L).

Consider the complex random vector

$$\underaccent{\tilde}{\boldsymbol{x}} = \sum\nolimits_{m=1}^{M} \underaccent{\tilde}{\boldsymbol{s}}_m + \underaccent{\tilde}{\boldsymbol{\eta}} \in \mathbb{C}^N \tag{19}$$

where the random vectors $\underaccent{\tilde}{\boldsymbol{s}}_1, \ldots, \underaccent{\tilde}{\boldsymbol{s}}_{M-1}$ are proper complex, $\underaccent{\tilde}{\boldsymbol{s}}_M$ is maximally improper, and $\underaccent{\tilde}{\boldsymbol{\eta}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$ is proper complex Gaussian noise. Let us assume that $\underaccent{\tilde}{\boldsymbol{s}}_m = \underaccent{\tilde}{\boldsymbol{h}}_m \underaccent{\tilde}{b}_m$, where $\underaccent{\tilde}{\boldsymbol{h}}_m \in \mathbb{C}^N$ are constant vectors, $\underaccent{\tilde}{b}_1, \ldots, \underaccent{\tilde}{b}_{M-1}$ are proper complex random variables, and $\underaccent{\tilde}{b}_M = b_M$ is a real-valued random variable. In this case, the covariance matrices of all signals $\underaccent{\tilde}{\boldsymbol{s}}_m$ have rank 1.

We intentionally keep this example as abstract as possible in order to not limit the discussion to a particular signal processing application. However, to get an idea where such a situation might occur, we could think of a communication system where several single-antenna transmitters send data to a multiantenna receiver. A received signal $\underaccent{\tilde}{\boldsymbol{x}}$ as given in (19) is then obtained if one transmitter uses a real-valued modulation alphabet (e.g., ASK) while all other users apply proper complex constellations (e.g., QPSK).

For the following numerical simulation, we assume the composite real covariance matrix $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ and the augmented complex covariance matrix $\underaccent{\tilde}{\boldsymbol{C}}_{\underaccent{\tilde}{\boldsymbol{x}}}$ to be known. In case of unknown covariance matrices, similar results can be computed based on a set of random samples of $\underaccent{\tilde}{\boldsymbol{x}}$ by using the composite real and augmented complex sample covariance matrices instead of the true covariance matrices.

In Fig. 1, we plot the share in overall variance that is covered by the first L real-valued principal components. For comparison, we add the respective value obtained with linear and widely linear complex PCA. However, the latter two methods only allow for integer numbers $\ell = \frac{L}{2}$ of complex principal components, i.e., they can only be considered for even L. We have chosen $N = 5$, $M = 3$, and $\sigma^2 = \frac{1}{4}$, and we have assumed that $\underaccent{\tilde}{b}_1$ and $\underaccent{\tilde}{b}_2$ are uniformly distributed on $\{\pm 1, \pm \mathrm{j}\}$ while b_3 is uniformly distributed on $\{\pm 1\}$. The entries of $\underset{\sim}{\boldsymbol{h}}_m$ have been sampled from an i.i.d. proper complex Gaussian distribution with zero mean and unit variance.

The plot confirms the theoretical findings from the previous section. Firstly, the composite real (CR) PCA with an even value of L is equivalent to the widely linear (WL) PCA based on the augmented complex representation with $\ell = \frac{L}{2}$. Secondly, both the composite real PCA and the widely linear PCA outperform the linear (L) version.

In addition, the composite real PCA can achieve a better dimensionality reduction than its complex counterparts. In the example, the composite real PCA covers most of the signal variance (more than 95%) with 5 real-valued principal components. To reach the same percentage by means of a complex PCA, we have to use at least 3 complex principal components, which corresponds to 6 real-valued ones.

The improved dimensionality reduction of the composite real PCA is possible in the scenario under consideration due to the fact that the composite real representation of the useful signal lies in a subspace with an odd number of dimensions. Since this does not correspond to an integer number of complex dimensions, the resolution of the complex representation is too coarse to account for this fact, and the number of required complex dimensions is the next higher integer.

7. COMPUTATIONAL COMPLEXITY

In addition to the finer granularity of the composite real PCA, there is also an advantage in terms of computational complexity. The most complex operation of PCA clearly is the eigenvalue decomposition (EVD) with complexity order $\mathcal{O}(N^w)$ for an $N \times N$ matrix, where $w > 2$. This operation has to be performed for a complex $2N \times 2N$ matrix in the augmented complex widely linear PCA. In the composite real PCA, we have to compute the EVD of a real-valued $2N \times 2N$ symmetric matrix, which has real-valued eigenvalues and eigenvectors. Therefore, complexity can be reduced by conducting all computations within the field of real numbers.

Since $2^w > 4$, the EVD of a real $2N \times 2N$ matrix with complexity order $\mathcal{O}((2N)^w)$ requires more numerical operations than the EVD of a complex $N \times N$ matrix if each complex multiplication is implemented as four real-valued ones. Thus, the complex linear PCA has lower computational complexity than the composite real PCA and is therefore preferable if the signal under consideration is known to be proper.

8. CONCLUSION

The proposed composite real principal component analysis (PCA) can be considered as a generalization of the widely linear PCA from [2,6] since it allows for a finer granularity of the dimension of the signal subspace. This makes the composite real PCA superior in dimensionality reduction tasks. When restricted to even numbers of real-valued principal components, the composite real PCA has the same performance as the widely linear PCA from [2, 6], but it has a lower computational complexity and a more intuitive derivation.

9. REFERENCES

[1] Fredy D. Neeser and James L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.

[2] Peter J. Schreier and Louis L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*, Cambridge University Press, 2010.

[3] D. P. Mandic and S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, John Wiley & Sons, 2009.

[4] Karl Pearson, "On lines and planes of closest fit to systems of points in space," *London, Edinburgh, Dublin Philosophical Mag. J. Sci.*, vol. 2, no. 11, pp. 559–572, Nov. 1901.

[5] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *J. Educational Psychology*, vol. 24, no. 6, pp. 417–441, Sept. 1933.

[6] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, Mar. 2003.

[7] I. T. Jolliffe, *Principal Component Analysis*, Springer, 2nd edition, 2002.

[8] Xi-Lin Li, T. Adali, and M. Anderson, "Noncircular principal component analysis and its application to model selection," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4516–4528, Oct. 2011.

[9] A. Papaioannou and S. Zafeiriou, "Principal component analysis with complex kernel: The widely linear model," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 25, no. 9, pp. 1719–1726, Sept. 2014.

[10] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[11] T. Adalı, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101–5125, Nov. 2011.

[12] C. Hellings and W. Utschick, "Block-skew-circulant matrices in complex-valued signal processing," *IEEE Trans. Signal Process.*, vol. 63, no. 8, pp. 2093–2107, Apr. 2015.

[13] Davide Mattera, Luigi Paura, and Fabio Sterle, "Widely linear MMSE transceiver for real-valued sequences over MIMO channel," in *Proc. 14th European Signal Process. Conf. (EUSIPCO)*, Florence, Italy, Sep. 4–8, 2006.

[14] S. Buzzi, M. Lops, and S. Sardellitti, "Widely linear reception strategies for layered space-time wireless communications," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2252–2262, June 2006.

[15] Z. K. M. Ho and E. Jorswieck, "Improper Gaussian signaling on the two-user SISO interference channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3194–3203, Sept. 2012.

[16] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

[17] A. Marshall, I. Olkin, and B. Arnold, *Inequalities: Theory of Majorization and its Applications*, Springer, 2011.