Block-Skew-Circulant Matrices in Complex-Valued Signal Processing

Christoph Hellings and Wolfgang Utschick

IEEE Transactions on Signal Processing

vol. 63, no. 8, Apr. 2015

Technische Universität München
Fachgebiet Methoden der Signalverarbeitung
http://www.msv.ei.tum.de

Block-Skew-Circulant Matrices in Complex-Valued Signal Processing

Christoph Hellings, *Student Member, IEEE*
and Wolfgang Utschick, *Senior Member, IEEE*

Abstract

Two main lines of approach can be identified in the recent literature on improper signals and widely linear operations. The augmented complex formulation based on the signal and its complex conjugate is considered as more insightful since it leads to convenient mathematical formulations for many considered problems. Moreover, it allows an easy distinction between proper and improper signals as well as between linear and widely linear operations. On the other hand, the composite real representation using the real and imaginary parts of the signal is closer to the actual implementation, and it allows to readily reuse results that have originally been derived for real-valued signals or proper complex signals. In this work, we aim at getting the best of both worlds by introducing mathematical tools that make the composite real representation more powerful and elegant. The proposed approach relies on a decomposition of real matrices into a block-skew-circulant and a block-Hankel-skew-circulant component. By means of various application examples from the field of signal processing for communications, we demonstrate the usefulness of the proposed framework.

Index Terms

Asymmetric complex, composite real representation, improper signals, linear algebra, noncircular, proper signals, statistical signal processing, widely linear processing.

The authors are with Fachgebiet Methoden der Signalverarbeitung, Technische Universität München, 80290 München, Germany, Telephone: +49 89 289-28516, Fax: +49 89 289-28504, e-mail: hellings@tum.de, utschick@tum.de.

I. Motivation

The successful application of widely linear filtering in practice as well as proofs of performance gains due to improper signaling in information theoretic studies have recently drawn the attention of the signal processing community to the processing of improper complex signals. An example from practice is the so-called *single-antenna interference cancellation* (*SAIC*, e.g., [1]), which is based on widely linear filtering. This technique is applied in the throughput-increasing *Voice over Adaptive Multi-user channels on One Slot* (*VAMOS*) method, which is included in 3GPP Release 9 [2]. From an information theoretic point of view, improper signals have attained considerable interest first of all due to the unexpected performance gains obtained using *interference alignment with asymmetric complex signaling* [3], i.e., using improper signals.

But what are the outstanding properties of widely linear filters and improper signals, which have led to these successes recently? And what is the relation between these two concepts? To answer these questions, let us consider the following. Any complex number $z = x + \mathrm{j}y,\ x, y \in \mathbb{R}$ can be represented as a real vector $\check{\boldsymbol{z}} = [x,\ y]^{\mathrm{T}}$, and, accordingly, a complex linear mapping $z \mapsto (a_1 + \mathrm{j}a_2)z,\ a_1, a_2 \in \mathbb{R}$ can be represented as

$$\check{\boldsymbol{z}} \mapsto \grave{\boldsymbol{A}}\check{\boldsymbol{z}} = \begin{bmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{bmatrix} \check{\boldsymbol{z}}. \tag{1}$$

But what about the reverse direction? Can any real linear mapping $\check{\boldsymbol{z}} \mapsto \boldsymbol{A}\check{\boldsymbol{z}}$ with $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ be represented as a linear mapping in the complex domain? The answer is no. This is possible only if the matrix $\boldsymbol{A}$ has the particular structure in (1) (e.g., [4], [5]). Otherwise, the corresponding mapping in the complex domain is not linear.

In a technical system, complex multiplications are implemented as four real-valued multiplications, and it seems to be natural to allow arbitrary real linear mappings $\check{\boldsymbol{z}} \mapsto \boldsymbol{A}\check{\boldsymbol{z}}$, while the restriction to complex linear mappings could be considered as artificial. Therefore, the term *widely linear* has been introduced for these more general operations, which are linear functions of the real and imaginary parts of a complex number—or, equivalently, of the complex number and its conjugate (e.g., [4]–[7]).

Now let $z = x + \mathrm{j}y$ be a complex random variable, where x and y are uncorrelated real random variables with equal variance. Such a complex random variable is called *proper* (e.g., [4], [5], [8], [9]). We can now ask the question what happens if we apply a linear transformation $z \mapsto Az$.

The answer is that the resulting random variable again has uncorrelated real and imaginary parts with equal variance, i.e., it is again a proper random variable. This changes if we instead apply a widely linear operation, which can introduce correlations between real and imaginary parts or individually change the power of these two components. The resulting random variable is then called *improper* (e.g., [4], [5], [9]). As above, excluding this possibility could be considered as an artificial restriction while allowing improper signals could—despite the name—be considered as the more natural assumption.

Nevertheless, the restriction to (strictly) linear operations and to proper signals is a common assumption in complex-valued signal processing, and there are good reasons for this. It has been shown that proper Gaussian random vectors have maximum differential entropy for a given covariance matrix [8], [10], and that physical signals such as the demodulated noise in a communication system can be assumed to be proper Gaussian [8], [10]. Moreover, we have seen that proper random vectors stay proper under linear transformations, and it follows that linear operations are the right choice as long as all signals in the system are proper (for an example of a formal proof in a particular scenario see [11]). Therefore, even though proper signals and linear processing are only special cases of general complex signals and widely linear processing, they are not arbitrary special cases, but cases of particular importance.

On the other hand, there are also many examples of improper signals, which are of practical importance. For example, many practical coding and modulation schemes in communication systems (e.g., BPSK, ASK, or GMSK) lead to transmit signals that do not fulfill the conditions for propriety. Whenever this is the case, complex linear operations might not exploit the potential of the system, and the natural extension to widely linear processing can be beneficial. The abovementioned VAMOS method takes advantage of this, by allowing simultaneous transmission of maximally improper signals to two users and by separating these two signals at the receivers—using widely linear filters. Another application of widely linear filters is the compensation of hardware imperfections—such as I/Q imbalance, which leads to improper received signals (e.g., [12]).

The above discussion makes clear that improper signals and widely linear operations are concepts that go hand in hand. Whenever a signal in the system is improper, widely linear processing should be applied, but linear processing is sufficient if all signals are proper. Knowing this, there still is the need to find out which one out of all widely linear filters is optimal for a given setting. This has led to a large number of publications on how to process improper signals and how to optimize widely linear filters (e.g., [1], [12]–[25]).

From a theoretical perspective, there is a second interesting question. If the given signals (e.g., the noise signals) are proper, and we are free to design the remaining signals as well as the filters, what is the better choice? Using only proper signals and making all filters linear? Or applying widely linear operations in order to introduce impropriety into the system? The probably best known result from this line of research is the abovementioned combination of interference alignment [26] and improper signals proposed in [3]. The large gains achieved by this scheme in a three-user interference channel (extended to four users in [27]) might have been somehow surprising at the first glance since using proper signals and linear filters was shown to be optimal for point-to-point communication [10]. However, from a different perspective, the surprise might not be that big since the restrictions to linear processing and proper signals can be considered as artificial as discussed above. Gains by dropping this restriction have recently been reported also for other interference channel scenarios [28]–[31] and for broadcast (point-to-multipoint) scenarios without nonlinear interference cancellation [11], [31], [32]. Apparently, the generalization to allow improper signals is particularly useful in interference-limited communication systems.

For both purposes, i.e., for the optimization of widely linear processing as well as for the study of potential gains due to improper signaling, it is clear that adequate mathematical tools and formulations are needed. Note that examples for both kinds of studies are included in Section VII.

Two kinds of mathematical tools have been widely applied in the recent literature on the subject. In the above discussion, a so-called *composite real representation* of vectors ($\check{\boldsymbol{z}}$) and matrices ($\grave{\boldsymbol{A}}$) was used. Apart from being closely related to the practical implementation and to our intuitive understanding of complex numbers, this representation has the following important advantage, which was exploited, e.g., in [11], [16], [17], [32]–[35] (and for the special case of scalar complex random variables in [3], [24], [28]). Whenever we know a method to solve a given problem involving real vectors, we can readily apply the method to composite real representations of general (proper or improper) complex vectors. Moreover, methods that have originally been developed for proper complex vectors can usually be easily transferred to the real-valued case and, via this intermediate step, also to general complex vectors.

On the other hand, there is a second formulation, which might be less straightforward in the first place, but has led to convenient mathematical formulations for various problems considered in recent research. This so-called *augmented complex representation* is based on the complex vector and its conjugate instead of on the real and imaginary parts. This formulation allows an easy identification of the proper signals among all complex signals as well as an easy identification of linear operations among all widely linear ones (e.g., [4], [5], [7], [9]). Moreover, a large variety of mathematical tools to work with augmented complex representations have been introduced in the literature, e.g., [4], [5], [7], [9], [29], [31], [36], [37]. However, the augmented complex representation makes it necessary to derive new algorithms since it does not feature the property that methods developed for real-valued signals can be adopted.

In this work, we aim at getting the best of both worlds by introducing mathematical tools that make the composite real representation more powerful and elegant. This shall not be misunderstood as a campaign against the augmented complex representation. Quite the contrary, we think that it is worthwhile having not only one powerful framework for dealing with impropriety and widely linear processing, but two different approaches, in order to always choose the one that is more convenient and more adequate for the problem under consideration. However, we want to argue against the impression that the augmented complex formulation generally leads to more convenient formulations than the composite real representation (see, e.g., [4], [5], [9], [29], [31], [37]). The framework based on *block-skew-circulant* and *block-Hankel-skew-circulant* matrices, which we propose in this paper, makes the composite real representation a powerful tool that can handle many problems at least as good as the augmented complex formulation. Moreover, we show examples of recent advances in communication theory that would not have been achieved without the composite real representation.

We first review the formal definitions of proper and improper signals, linear and widely linear operations, as well as augmented complex and composite real representations (Section II). Then, we introduce the definition and the properties of real block-skew-circulant ($\mathcal{BSC}$) and block-Hankel-skew-circulant ($\mathcal{BHSC}$) matrices (Section III) as well as their relation to complex matrices (Section IV). In Section V, we show how $\mathcal{BSC}$ and $\mathcal{BHSC}$ matrices can be used to describe widely linear filters, and in Section VI, we apply them to describe the statistical properties of complex random vectors. Having established the framework of block-skew-circulant matrices for complex-valued signal processing, we proceed to the presentation of application examples in Section VII: we show how various results from the recent literature fit into the proposed framework, and we point out possible further applications.

Notation: Vectors are typeset in boldface lowercase letters and matrices in boldface uppercase letters. To easily distinguish real quantities from complex quantities, we use a tilde $\underset{\sim}{\bullet}$ below complex quantities. We write $\mathbf{0}$ for the zero matrix or vector, $\mathbf{I}_L$ for the identity matrix of size L, $\bullet^{\mathrm{T}}$ for the transpose of a vector or matrix, $\bullet^{\mathrm{H}}$ for the conjugate transpose, and $\bullet^{+}$ for the pseudoinverse. We use $\Re$, $\Im$, and $\bullet^*$ for real part, imaginary part, and complex conjugation, respectively, and the shorthand notation $\check{\boldsymbol{x}}$ (pronounced as x-check or x-real) is used for a vector $[\Re(\underset{\sim}{\boldsymbol{x}})^{\mathrm{T}}, \Im(\underset{\sim}{\boldsymbol{x}})^{\mathrm{T}}]^{\mathrm{T}}$. We use $\otimes$ for the Kronecker product, and $\bullet^{\perp}$ for the orthogonal complement of a linear subspace. The space $\mathbb{S}^{2M} \subset \mathbb{R}^{2M\times 2M}$ is the space of real symmetric matrices. The order relation $\succeq$ has to be understood in the sense of positive semi-definiteness.

II. Improper Signals and Widely Linear Operations

In the last section, we have already seen scalar examples of linear and widely linear operations and of proper and improper signals. However, for the remainder of the paper, it is necessary to formally define these notions for the vector case. In accordance with the existing literature on the subject (e.g., [4]–[9], [38]), we introduce the following definitions.

Definition 1: Given a complex random vector $\underset{\sim}{\boldsymbol{x}}$, the matrix $\boldsymbol{C}_{\underset{\sim}{\boldsymbol{x}}} = \mathrm{E}\big[(\underset{\sim}{\boldsymbol{x}} - \mathrm{E}[\underset{\sim}{\boldsymbol{x}}])(\underset{\sim}{\boldsymbol{x}} - \mathrm{E}[\underset{\sim}{\boldsymbol{x}}])^{\mathrm{H}}\big]$ is the *covariance matrix* of $\underset{\sim}{\boldsymbol{x}}$, and the matrix $\tilde{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}} = \mathrm{E}\big[(\underset{\sim}{\boldsymbol{x}} - \mathrm{E}[\underset{\sim}{\boldsymbol{x}}])(\underset{\sim}{\boldsymbol{x}} - \mathrm{E}[\underset{\sim}{\boldsymbol{x}}])^{\mathrm{T}}\big]$ is the *pseudo-covariance matrix* of $\underset{\sim}{\boldsymbol{x}}$.

Alternative names for the pseudocovariance matrix are *complementary covariance matrix*, *conjugate covariance matrix*, or *relation matrix* (e.g., [5], [38]).

Definition 2: A complex random vector $\underset{\sim}{\boldsymbol{x}}$ is called *proper* if the pseudocovariance matrix $\tilde{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$ vanishes. Otherwise, it is called *improper*.

For zero-mean Gaussian distributions, propriety is equivalent to a circularly symmetric probability density function [5].

Definition 3: A complex mapping $\underset{\sim}{\boldsymbol{x}} \mapsto \underset{\sim}{\boldsymbol{f}}(\underset{\sim}{\boldsymbol{x}})$ is called *widely linear* if it can be expressed as [6]

$$\underset{\sim}{\boldsymbol{f}}(\underset{\sim}{\boldsymbol{x}}) = \underset{\sim}{\boldsymbol{A}}_{\mathrm{L}}\underset{\sim}{\boldsymbol{x}} + \underset{\sim}{\boldsymbol{A}}_{\mathrm{CL}}\underset{\sim}{\boldsymbol{x}}^* \tag{2}$$

where the complex matrices $\underset{\sim}{\boldsymbol{A}}_{\mathrm{L}}$ and $\underset{\sim}{\boldsymbol{A}}_{\mathrm{CL}}$ are used as factors for the linear part and the conjugate linear part, respectively.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TSP.2015.2395992

Therefore, widely linear operations are also called *linear-conjugate-linear* operations (e.g., [5]). Alternatively, widely linear mappings can be expressed as

$$\underset{\sim}{\boldsymbol{f}}(\underset{\sim}{\boldsymbol{x}}) = \underset{\sim}{\boldsymbol{A}}_{\mathrm{R}} \Re(\underset{\sim}{\boldsymbol{x}}) + \underset{\sim}{\boldsymbol{A}}_{\mathrm{I}} \Im(\underset{\sim}{\boldsymbol{x}}) \tag{3}$$

with complex matrices $\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}$ and $\underset{\sim}{\boldsymbol{A}}_{\mathrm{I}}$.

Definition 4: The *augmented complex* representation of a widely linear mapping is

$$\underline{\boldsymbol{f}}(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{A}}\,\underline{\boldsymbol{x}} = \begin{bmatrix} \underset{\sim}{\boldsymbol{A}}_{\mathrm{L}} & \underset{\sim}{\boldsymbol{A}}_{\mathrm{CL}} \\ \underset{\sim}{\boldsymbol{A}}^{*}_{\mathrm{CL}} & \underset{\sim}{\boldsymbol{A}}^{*}_{\mathrm{L}} \end{bmatrix} \begin{bmatrix} \underset{\sim}{\boldsymbol{x}} \\ \underset{\sim}{\boldsymbol{x}}^{*} \end{bmatrix}. \tag{4}$$

In the special case of a linear mapping, $\underset{\sim}{\boldsymbol{A}}_{\mathrm{CL}}$ vanishes, and $\underline{\boldsymbol{A}}$ becomes block-diagonal. Similarly, the augmented complex covariance matrix

$$\underset{\sim}{\boldsymbol{C}}_{\underline{\boldsymbol{x}}} = \mathrm{E}\big[(\underline{\boldsymbol{x}} - \mathrm{E}[\underline{\boldsymbol{x}}])(\underline{\boldsymbol{x}} - \mathrm{E}[\underline{\boldsymbol{x}}])^{\mathrm{H}}\big] = \begin{bmatrix} \underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}} & \tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}} \\ \tilde{\underset{\sim}{\boldsymbol{C}}}^{*}_{\underset{\sim}{\boldsymbol{x}}} & \underset{\sim}{\boldsymbol{C}}^{*}_{\underset{\sim}{\boldsymbol{x}}} \end{bmatrix} \tag{5}$$

is block-diagonal if $\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}} = \boldsymbol{0}$, i.e., if $\underset{\sim}{\boldsymbol{x}}$ is proper.

Definition 5: The *composite real* representation of a widely linear mapping is

$$\check{\boldsymbol{f}}(\check{\boldsymbol{x}}) = \boldsymbol{A}_{\mathrm{WL}}\check{\boldsymbol{x}} = \begin{bmatrix} \Re(\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}) & \Re(\underset{\sim}{\boldsymbol{A}}_{\mathrm{I}}) \\ \Im(\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}) & \Im(\underset{\sim}{\boldsymbol{A}}_{\mathrm{I}}) \end{bmatrix} \begin{bmatrix} \Re(\underset{\sim}{\boldsymbol{x}}) \\ \Im(\underset{\sim}{\boldsymbol{x}}) \end{bmatrix}. \tag{6}$$

In the composite real representation, identifying the special case of a complex linear mapping is not as simple as in the augmented complex version. Here, we have to verify whether $\Re(\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}) = \Im(\underset{\sim}{\boldsymbol{A}}_{\mathrm{I}})$ and $\Im(\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}) = -\Re(\underset{\sim}{\boldsymbol{A}}_{\mathrm{I}})$ such that we have the special block structure

$$\check{\boldsymbol{f}}(\check{\boldsymbol{x}}) = \begin{bmatrix} \Re(\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}) & -\Im(\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}) \\ \Im(\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}) & \Re(\underset{\sim}{\boldsymbol{A}}_{\mathrm{R}}) \end{bmatrix} \begin{bmatrix} \Re(\underset{\sim}{\boldsymbol{x}}) \\ \Im(\underset{\sim}{\boldsymbol{x}}) \end{bmatrix}. \tag{7}$$

In Section V, we introduce a method to easily identify complex linear mappings in the composite real representation.

As shown in Section VI, the framework introduced in this paper additionally allows easy identification of proper complex signals. However, at the first glance, the covariance matrix of the composite real representation[1]

$$\boldsymbol{C}_{\check{\boldsymbol{x}}} = \mathrm{E}\big[(\check{\boldsymbol{x}} - \mathrm{E}[\check{\boldsymbol{x}}])(\check{\boldsymbol{x}} - \mathrm{E}[\check{\boldsymbol{x}}])^{\mathrm{T}}\big] = \begin{bmatrix} \boldsymbol{C}_{\Re\underset{\sim}{\boldsymbol{x}}} & \boldsymbol{C}_{\Re\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{x}}} \\ \boldsymbol{C}^{\mathrm{T}}_{\Re\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{x}}} & \boldsymbol{C}_{\Im\underset{\sim}{\boldsymbol{x}}} \end{bmatrix} \tag{8}$$

[1]This matrix may not be confused with the composite real equivalent $\grave{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$ [see (23)] of the (complex) covariance matrix $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$.

does not directly reveal whether or not a random vector $\boldsymbol{x}$ is proper: we have to check the conditions

$$\boldsymbol{C}_{\Re \underset{\sim}{\boldsymbol{x}}} = \boldsymbol{C}_{\Im \underset{\sim}{\boldsymbol{x}}} \quad \text{and} \quad \boldsymbol{C}^{\mathrm{T}}_{\Re \underset{\sim}{\boldsymbol{x}} \Im \underset{\sim}{\boldsymbol{x}}} = -\boldsymbol{C}_{\Re \underset{\sim}{\boldsymbol{x}} \Im \underset{\sim}{\boldsymbol{x}}} \tag{9}$$

to test for propriety based on the composite real formulation (e.g., [5]), i.e., we have to verify whether the covariance matrix of the composite real random vector has the special structure

$$\boldsymbol{C}_{\check{\boldsymbol{x}}} = \begin{bmatrix} \boldsymbol{C}_{\Re \underset{\sim}{\boldsymbol{x}}} & -\boldsymbol{C}^{\mathrm{T}}_{\Re \underset{\sim}{\boldsymbol{x}} \Im \underset{\sim}{\boldsymbol{x}}} \\ \boldsymbol{C}^{\mathrm{T}}_{\Re \underset{\sim}{\boldsymbol{x}} \Im \underset{\sim}{\boldsymbol{x}}} & \boldsymbol{C}_{\Re \underset{\sim}{\boldsymbol{x}}} \end{bmatrix}. \tag{10}$$

Note that the block structures in (10) and (7) are the same.

III. Block-Skew-Circulant Matrices

It can be seen from (7) and (10) that there is a particular block structure that plays an important role for complex-valued signal processing. As explained below, matrices with this special structure can be called *block-skew-circulant* matrices.

While there exists a significant number of works dealing with skew-circulant matrices (e.g., [39]–[41]) and block-circulant matrices (e.g., [42] and the references therein), only few results on block-skew-circulant matrices can be found in the existing literature. The rare examples include the statements in [43], [44] that are referenced later in this section. Moreover, block-skew-circulant matrices were recently used as a tool in [45], but without a detailed study of their properties.

Therefore, this section is meant as a collection of formulae for block-skew-circulant and block-Hankel-skew-circulant matrices. Most of the properties presented in this section are quite easy to prove. Nevertheless, we formally state them as lemmas in order to allow easy referencing later on. Even though many of the properties introduced in this section analogously hold for complex matrices, we restrict ourselves to real-valued matrices since this is sufficient for the intended application.

A. Fundamentals

A skew-circulant matrix is the special case of a Toeplitz matrix where each row is a cyclically shifted copy of the preceding row with a sign change for the elements on one side of the main diagonal (e.g., [39]).[2] With matrix blocks $\boldsymbol{A}_i \in \mathbb{R}^{K\times L}$ instead of scalar elements a_i, we obtain an N-block-skew-circulant ($\mathcal{BSC}_N$) matrix.[3]

Definition 6: A matrix with the block structure

$$\grave{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A}_1 & -\boldsymbol{A}_2 & \dots & -\boldsymbol{A}_N \\ \boldsymbol{A}_N & \boldsymbol{A}_1 & \dots & -\boldsymbol{A}_{N-1} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{A}_2 & \boldsymbol{A}_3 & \dots & \boldsymbol{A}_1 \end{bmatrix} \tag{11}$$

is called N-block-skew-circulant ($\mathcal{BSC}_N$). The set of real $\mathcal{BSC}$ matrices with N blocks of size $K \times L$ is denoted by $\mathcal{BSC}_N^{K\times L} \subset \mathbb{R}^{NK\times NL}$, where the superscript of $\mathcal{BSC}_N$ may be omitted if the block size becomes clear from the context.

With a block-Hankel structure instead of a block-Toeplitz structure, we can define N-block-Hankel-skew-circulant ($\mathcal{BHSC}_N$) matrices,[4] which are a generalization of skew-left-circulant matrices (e.g., [39]) to the case with block elements.

Definition 7: A matrix with the block structure

$$\acute{\boldsymbol{B}} = \begin{bmatrix} \boldsymbol{B}_1 & \dots & \boldsymbol{B}_{N-1} & \boldsymbol{B}_N \\ \boldsymbol{B}_2 & \dots & \boldsymbol{B}_N & -\boldsymbol{B}_1 \\ \vdots & \iddots & & \vdots \\ \boldsymbol{B}_N & \dots & -\boldsymbol{B}_{N-2} & -\boldsymbol{B}_{N-1} \end{bmatrix} \tag{12}$$

is called N-block-Hankel-skew-circulant ($\mathcal{BHSC}_N$). The set of real $\mathcal{BHSC}$ matrices with N blocks of size $K \times L$ is denoted by $\mathcal{BHSC}_N^{K\times L} \subset \mathbb{R}^{NK\times NL}$, where the superscript of $\mathcal{BHSC}_N$ may be omitted if the block size becomes clear from the context.

Throughout the paper, the notations $\grave{\boldsymbol{A}}$ and $\acute{\boldsymbol{B}}$ are used to indicate the block-Toeplitz and block-Hankel structures, respectively. We have chosen these notations since the shape of the grave accent ` reminds us of northwest-to-southeast diagonals, which are constant in a Toeplitz matrix. Accordingly, the acute accent ´ resembles the southwest-to-northeast diagonals of a

[2]Note that [46] uses a different nomenclature where *skew-circulant* refers to a circulant Hankel matrix. Here, *skew* refers to the sign change instead.

[3]Or block-Toeplitz-skew-circulant, block-skew-right-circulant.

[4]Or block-skew-left-circulant.

Hankel matrix. Depending on the reader's preference, $\grave{\boldsymbol{A}}$ could, e.g., be pronounced as A-grave, A-right, or A-BSC, while $\acute{\boldsymbol{B}}$ could, e.g., be pronounced as B-acute, B-left, or B-BHSC.

For the case $N = 2$, the block structures reduce to

$$\grave{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A}_1 & -\boldsymbol{A}_2 \\ \boldsymbol{A}_2 & \boldsymbol{A}_1 \end{bmatrix} \qquad \text{and} \qquad \acute{\boldsymbol{B}} = \begin{bmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_2 \\ \boldsymbol{B}_2 & -\boldsymbol{B}_1 \end{bmatrix}. \tag{13}$$

The first important property to note is that the set of block-skew-circulant matrices is closed under transposition (cf. [43]) and under vector space operations. The same is true for the set of block-Hankel-skew-circulant matrices.

Lemma 1: Let $\grave{\boldsymbol{A}} \in \mathcal{BSC}_N^{K\times L}$, $\grave{\boldsymbol{A}}' \in \mathcal{BSC}_N^{K\times L}$, $\acute{\boldsymbol{B}} \in \mathcal{BHSC}_N^{K\times L}$, $\acute{\boldsymbol{B}}' \in \mathcal{BHSC}_N^{K\times L}$, and $\alpha, \beta \in \mathbb{R}$. Then,

1) $\grave{\boldsymbol{A}}^{\mathrm{T}} \in \mathcal{BSC}_N^{L\times K}$,
2) $\acute{\boldsymbol{B}}^{\mathrm{T}} \in \mathcal{BHSC}_N^{L\times K}$,
3) $\alpha\grave{\boldsymbol{A}} + \beta\grave{\boldsymbol{A}}' \in \mathcal{BSC}_N^{K\times L}$,
4) $\alpha\acute{\boldsymbol{B}} + \beta\acute{\boldsymbol{B}}' \in \mathcal{BHSC}_N^{K\times L}$.

Proof: Immediate after inserting the block structures. ■

This means that $\mathcal{BSC}_N^{K\times L}$ and $\mathcal{BHSC}_N^{K\times L}$ are linear subspaces of $\mathbb{R}^{NK\times NL}$, see item 3) and 4).

B. Matrix Products

Inspired by the backward identity matrix used in the study of Toeplitz and Hankel circulants with scalar elements in [47], let us define the following permutation matrix, which is an orthogonal $\mathcal{BHSC}_N$ matrix:

$$\acute{\boldsymbol{\Pi}}_{N,M} = \begin{bmatrix} & & \mathbf{I}_M \\ & \cdot^{\cdot^{\cdot}} & \\ \mathbf{I}_M & & \end{bmatrix} \in \mathcal{BHSC}_N^{M\times M}. \tag{14}$$

Lemma 2: $\acute{\boldsymbol{B}} \in \mathcal{BHSC}_N^{K\times L}$ if and only if there exist $\grave{\boldsymbol{C}}, \grave{\boldsymbol{C}}' \in \mathcal{BSC}_N^{K\times L}$ such that $\acute{\boldsymbol{B}} = \grave{\boldsymbol{C}}\acute{\boldsymbol{\Pi}}_{N,L} = \acute{\boldsymbol{\Pi}}_{N,K}\grave{\boldsymbol{C}}'$.

Proof: Reversing the order of the block-rows (or block-columns) turns the structure (11) into the structure (12). ■

The permutation matrix $\acute{\boldsymbol{\Pi}}_{N,M}$ is useful to prove the following lemma about matrix multiplication of block-skew-circulant matrices and block-Hankel-skew-circulant matrices.

Lemma 3: Let $\grave{\boldsymbol{A}} \in \mathcal{BSC}_N^{K\times L}$, $\grave{\boldsymbol{A}}' \in \mathcal{BSC}_N^{L\times M}$, $\acute{\boldsymbol{B}} \in \mathcal{BHSC}_N^{K\times L}$, and $\acute{\boldsymbol{B}}' \in \mathcal{BHSC}_N^{L\times M}$. Then

1) $\grave{\boldsymbol{A}}\grave{\boldsymbol{A}}' \in \mathcal{BSC}_N^{K\times M}$,
2) $\acute{\boldsymbol{B}}\acute{\boldsymbol{B}}' \in \mathcal{BSC}_N^{K\times M}$,
3) $\grave{\boldsymbol{A}}\acute{\boldsymbol{B}}' \in \mathcal{BHSC}_N^{K\times M}$,
4) $\acute{\boldsymbol{B}}\grave{\boldsymbol{A}}' \in \mathcal{BHSC}_N^{K\times M}$.

Proof: Item 1) is easy to verify after inserting the $\mathcal{BSC}$ block structure and calculating the matrix product. Using 1) and Lemma 2, we have

2) $\acute{\boldsymbol{B}}\acute{\boldsymbol{B}}' = \grave{\boldsymbol{C}}\acute{\boldsymbol{\Pi}}_{N,L}\acute{\boldsymbol{\Pi}}_{N,L}\grave{\boldsymbol{C}}' = \grave{\boldsymbol{C}}\grave{\boldsymbol{C}}' \in \mathcal{BSC}_N^{K\times M}$.
3) $\grave{\boldsymbol{A}}\acute{\boldsymbol{B}}' = \grave{\boldsymbol{A}}(\grave{\boldsymbol{C}}\acute{\boldsymbol{\Pi}}_{N,L}) = (\grave{\boldsymbol{A}}\grave{\boldsymbol{C}})\acute{\boldsymbol{\Pi}}_{N,L} \in \mathcal{BHSC}_N^{K\times M}$ since $\grave{\boldsymbol{A}}\grave{\boldsymbol{C}} \in \mathcal{BSC}_N^{K\times M}$.
4) $\acute{\boldsymbol{B}}\grave{\boldsymbol{A}}' = (\acute{\boldsymbol{\Pi}}_{N,K}\grave{\boldsymbol{C}}')\grave{\boldsymbol{A}}' = \acute{\boldsymbol{\Pi}}_{N,K}(\grave{\boldsymbol{C}}'\grave{\boldsymbol{A}}') \in \mathcal{BHSC}_N^{K\times M}$ since $\grave{\boldsymbol{C}}'\grave{\boldsymbol{A}}' \in \mathcal{BSC}_N^{K\times M}$. ■

From Lemma 3, we see that the set of block-skew-circulant matrices is closed under multiplication (as already mentioned in [43]) whereas the product of two block-Hankel-skew-circulant matrices is block-skew-circulant.

C. Decomposition into Orthogonal Components

From now on, we restrict ourselves to the case of $N = 2$ blocks, which is sufficient for the intended application.

Lemma 4: Let $\acute{\boldsymbol{B}} \in \mathcal{BHSC}_2^{K\times K}$. Then $\mathrm{tr}[\acute{\boldsymbol{B}}] = 0$.

Proof: $\mathrm{tr}[\acute{\boldsymbol{B}}] = \mathrm{tr}[\boldsymbol{B}_1] + \mathrm{tr}[-\boldsymbol{B}_1] = 0$. ■

Lemma 4 can be generalized to arbitrary even numbers of blocks, but the following properties, which play a crucial role in the remainder of the paper, are specific for $N = 2$.

First, recall that $\mathbb{R}^{2K\times 2L}$ is a $4KL$-dimensional vector space with an inner product defined as $\langle \boldsymbol{A}, \boldsymbol{B}\rangle = \mathrm{tr}[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}]$ for $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{2K\times 2L}$ [48, Section 5.7]. Also recall that two elements of a vector space are orthogonal to each other if their inner product equals zero.

Lemma 5: $(\mathcal{BSC}_2^{K\times L})^{\perp} = \mathcal{BHSC}_2^{K\times L}$, i.e., the space of 2-block-skew-circulant matrices is the orthogonal complement of the space of 2-block-Hankel-skew-circulant matrices in $\mathbb{R}^{2K\times 2L}$.

Proof: For any $\grave{\boldsymbol{A}} \in \mathcal{BSC}_2^{K\times L}$, $\acute{\boldsymbol{B}} \in \mathcal{BHSC}_2^{K\times L}$, $\grave{\boldsymbol{A}}^{\mathrm{T}}\acute{\boldsymbol{B}} \in \mathcal{BHSC}_2^{K\times K}$ due to Lemma 3, so that $\mathrm{tr}[\grave{\boldsymbol{A}}^{\mathrm{T}}\acute{\boldsymbol{B}}] = 0$ due to Lemma 4. This establishes orthogonality. $\mathcal{BHSC}_2^{K\times L}$ and $\mathcal{BSC}_2^{K\times L}$ are both $2KL$-dimensional since two blocks of size $K \times L$ can be chosen freely in both cases. This sums up to $4KL$, which is the dimensionality of $\mathbb{R}^{2K\times 2L}$. ■

As a consequence, any matrix $\boldsymbol{C} \in \mathbb{R}^{2K\times 2L}$ can be uniquely decomposed into a $\mathcal{BSC}_2$ component and a $\mathcal{BHSC}_2$ component:[5]

$$\boldsymbol{C} = \grave{\boldsymbol{C}} + \acute{\boldsymbol{C}} = \mathbb{P}_{\mathcal{BSC}_2}(\boldsymbol{C}) + \mathbb{P}_{\mathcal{BHSC}_2}(\boldsymbol{C}) \tag{15}$$

where the projection operators $\mathbb{P}_\bullet$ are given in the following lemma.

Lemma 6: The orthogonal projections to $\mathcal{BSC}_2^{K\times L}$ and $\mathcal{BHSC}_2^{K\times L}$ are given by

$$\mathbb{P}_{\mathcal{BSC}_2}\left(\begin{bmatrix} \boldsymbol{C}_1 & \boldsymbol{C}_2 \\ \boldsymbol{C}_3 & \boldsymbol{C}_4 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} \boldsymbol{C}_1+\boldsymbol{C}_4 & \boldsymbol{C}_2-\boldsymbol{C}_3 \\ \boldsymbol{C}_3-\boldsymbol{C}_2 & \boldsymbol{C}_1+\boldsymbol{C}_4 \end{bmatrix}, \tag{16}$$

$$\mathbb{P}_{\mathcal{BHSC}_2}\left(\begin{bmatrix} \boldsymbol{C}_1 & \boldsymbol{C}_2 \\ \boldsymbol{C}_3 & \boldsymbol{C}_4 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} \boldsymbol{C}_1-\boldsymbol{C}_4 & \boldsymbol{C}_2+\boldsymbol{C}_3 \\ \boldsymbol{C}_2+\boldsymbol{C}_3 & \boldsymbol{C}_4-\boldsymbol{C}_1 \end{bmatrix}. \tag{17}$$

Proof: The approximation error

$$\boldsymbol{C} - \mathbb{P}_{\mathcal{BSC}_2}(\boldsymbol{C}) = \frac{1}{2}\begin{bmatrix} \boldsymbol{C}_1-\boldsymbol{C}_4 & \boldsymbol{C}_2+\boldsymbol{C}_3 \\ \boldsymbol{C}_2+\boldsymbol{C}_3 & \boldsymbol{C}_4-\boldsymbol{C}_1 \end{bmatrix} \in \mathcal{BHSC}_2^{K\times L} \tag{18}$$

is orthogonal to all $\grave{\boldsymbol{C}} \in \mathcal{BSC}_2^{K\times L}$ due to Lemma 5. The approximation error

$$\boldsymbol{C} - \mathbb{P}_{\mathcal{BHSC}_2}(\boldsymbol{C}) = \frac{1}{2}\begin{bmatrix} \boldsymbol{C}_1+\boldsymbol{C}_4 & \boldsymbol{C}_2-\boldsymbol{C}_3 \\ \boldsymbol{C}_3-\boldsymbol{C}_2 & \boldsymbol{C}_1+\boldsymbol{C}_4 \end{bmatrix} \in \mathcal{BSC}_2^{K\times L} \tag{19}$$

is orthogonal to all $\acute{\boldsymbol{C}} \in \mathcal{BHSC}_2^{K\times L}$ due to Lemma 5. ■

The decomposition into two orthogonal components given in (15) allows us to provide converses to some of the statements given in Lemmas 1 and 3. In the following lemma, we state only some examples that will be useful later in this paper, but we remark that further statements of similar kind are possible.

Lemma 7: Let $\grave{\boldsymbol{A}} \in \mathcal{BSC}_2^{K\times L}$, $\boldsymbol{C} \in \mathbb{R}^{2K\times 2L}$, and $\boldsymbol{C}' \in \mathbb{R}^{2L\times 2M}$. Then,

1) $\grave{\boldsymbol{A}} + \boldsymbol{C} \in \mathcal{BSC}_2^{K\times L}$ only if $\boldsymbol{C} \in \mathcal{BSC}_2^{K\times L}$.
2) If $\mathrm{null}[\grave{\boldsymbol{A}}] = \{\boldsymbol{0}\}$, $\grave{\boldsymbol{A}}\boldsymbol{C}' \in \mathcal{BSC}_2^{K\times M}$ only if $\boldsymbol{C}' \in \mathcal{BSC}_2^{L\times M}$.
3) If $\mathrm{null}[\grave{\boldsymbol{A}}] = \{\boldsymbol{0}\}$, $\grave{\boldsymbol{A}}\boldsymbol{C}' \in \mathcal{BHSC}_2^{K\times M}$ only if $\boldsymbol{C}' \in \mathcal{BHSC}_2^{L\times M}$.

[5]In a different context, a decomposition of general block-Toeplitz matrices into $\mathcal{BSC}$ and block-circulant components was proposed in [45]. The decomposition into $\mathcal{BSC}$ and $\mathcal{BHSC}$ components proposed here is on the one hand more general since it is not restricted to block-Toeplitz matrices, but on the other hand more limited since it only applies to block structures with 2×2 blocks. Another difference is that the decomposition proposed here is unique.

Proof: Let $\boldsymbol{C}$ and $\boldsymbol{C}'$ be decomposed as in (15).

1) $\mathbb{P}_{\mathcal{BHSC}_2}(\grave{\boldsymbol{A}} + \grave{\boldsymbol{C}} + \acute{\boldsymbol{C}}) = \acute{\boldsymbol{C}} = \boldsymbol{0}$ only if $\acute{\boldsymbol{C}} = \boldsymbol{0}$.
2) $\mathbb{P}_{\mathcal{BHSC}_2}(\grave{\boldsymbol{A}}\grave{\boldsymbol{C}}' + \grave{\boldsymbol{A}}\acute{\boldsymbol{C}}') = \grave{\boldsymbol{A}}\acute{\boldsymbol{C}}' = \boldsymbol{0}$ only if $\acute{\boldsymbol{C}}' = \boldsymbol{0}$.
3) $\mathbb{P}_{\mathcal{BSC}_2}(\grave{\boldsymbol{A}}\grave{\boldsymbol{C}}' + \grave{\boldsymbol{A}}\acute{\boldsymbol{C}}') = \grave{\boldsymbol{A}}\grave{\boldsymbol{C}}' = \boldsymbol{0}$ only if $\grave{\boldsymbol{C}}' = \boldsymbol{0}$. ∎

The following lemma can be generalized to arbitrary numbers of blocks [43], but the proof is particularly simple for $N = 2$ blocks.

Lemma 8: Let $\grave{\boldsymbol{A}} \in \mathcal{BSC}_2^{K\times K}$. Then, $\grave{\boldsymbol{A}}^{-1} \in \mathcal{BSC}_2^{K\times K}$ if it exists.

Proof: $\mathrm{null}[\grave{\boldsymbol{A}}] = \{\boldsymbol{0}\}$ if $\grave{\boldsymbol{A}}^{-1}$ exists. Thus, item 2) of Lemma 7 applies to $\grave{\boldsymbol{A}}\grave{\boldsymbol{A}}^{-1} = \mathbf{I}_{2K} \in \mathcal{BSC}_2^{K\times K}$. ∎

D. Symmetric Matrices

In our recent work [34], a decomposition of the space of real symmetric $2M \times 2M$ matrices into a so-called power shaping space $\mathcal{P}^M$ and an impropriety space $\mathcal{N}^M$ was proposed. We come back to the notions of power shaping and impropriety in Section VI, where we consider the application of $\mathcal{BSC}_2$ and $\mathcal{BHSC}_2$ matrices for describing statistical properties of complex signals. However, at this point, we want to provide a formal definition of these two subspaces by means of the notations defined in this section.

Definition 8: The sets of real symmetric $\mathcal{BSC}_2$ and $\mathcal{BHSC}_2$ matrices are denoted by

$$\mathcal{P}^M = \mathbb{S}^{2M} \cap \mathcal{BSC}_2^{M\times M} \tag{20}$$

$$\mathcal{N}^M = \mathbb{S}^{2M} \cap \mathcal{BHSC}_2^{M\times M}. \tag{21}$$

In the context of complex-valued signal processing, $\mathcal{P}^M$ may be called power shaping space, and $\mathcal{N}^M$ may be called noncircularity space or impropriety space.

Note that the following lemma, which is reproduced from [34], is not a trivial consequence of Lemma 5.

Lemma 9: $\mathcal{N}^M$ is the orthogonal complement of $\mathcal{P}^M$ in $\mathbb{S}^{2M}$.

Proof: Orthogonality follows from Lemma 5. To obtain symmetric $\grave{\boldsymbol{P}} \in \mathcal{P}^M$ and $\acute{\boldsymbol{N}} \in \mathcal{N}^M$ from (13), we need that $\boldsymbol{P}_1$, $\boldsymbol{N}_1$, and $\boldsymbol{N}_2$ are symmetric whereas $\boldsymbol{P}_2$ must be skew-symmetric. Thus, we can count that $\mathcal{P}^M$ is M^2-dimensional, and $\mathcal{N}^M$ is $(M^2 + M)$-dimensional. This adds up to $(2M^2 + M)$, which is the dimensionality of $\mathbb{S}^{2M}$. ∎

For some proofs later in this paper, the following lemma is helpful.

Lemma 10: Let $\grave{\boldsymbol{P}} \in \mathcal{P}^M$, $\acute{\boldsymbol{N}} \in \mathcal{N}^M$, and

$$\grave{\boldsymbol{J}}_M = \begin{bmatrix} \mathbf{0} & -\mathbf{I}_M \\ \mathbf{I}_M & \mathbf{0} \end{bmatrix}. \tag{22}$$

Then, we have

1) $\grave{\boldsymbol{J}}_M^{-1} = \grave{\boldsymbol{J}}_M^{\mathrm{T}} = -\grave{\boldsymbol{J}}_M$,
2) $\grave{\boldsymbol{J}}_M^{\mathrm{T}} \grave{\boldsymbol{P}} \grave{\boldsymbol{J}}_M = \grave{\boldsymbol{P}}$,
3) $\grave{\boldsymbol{J}}_M^{\mathrm{T}} \acute{\boldsymbol{N}} \grave{\boldsymbol{J}}_M = -\acute{\boldsymbol{N}}$,
4) $\boldsymbol{x}^{\mathrm{T}} \grave{\boldsymbol{J}}_M \boldsymbol{x} = 0,\ \forall \boldsymbol{x} \in \mathbb{R}^{2M}$,
5) $\boldsymbol{x} \in \mathrm{null}[\grave{\boldsymbol{P}}] \Leftrightarrow \grave{\boldsymbol{J}}_M \boldsymbol{x} \in \mathrm{null}[\grave{\boldsymbol{P}}]$,
6) $\boldsymbol{x} \in \mathrm{null}[\acute{\boldsymbol{N}}] \Leftrightarrow \grave{\boldsymbol{J}}_M \boldsymbol{x} \in \mathrm{null}[\acute{\boldsymbol{N}}]$.

Proof: Immediate after inserting the block structures of the matrices and the partitioning $\boldsymbol{x}^{\mathrm{T}} = [\boldsymbol{x}_1^{\mathrm{T}}, \boldsymbol{x}_2^{\mathrm{T}}]$. ■

E. Eigenvalue Decompositions

We conclude this section by some statements concerning eigenvalue decompositions (EVDs) and singular value decompositions (SVDs) of block-skew-circulant matrices.[6]

Lemma 11: For a symmetric $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{A}} \in \mathcal{P}^M$, the eigenvalue decomposition can be written as $\grave{\boldsymbol{A}} = \grave{\boldsymbol{Q}} \grave{\boldsymbol{\Lambda}} \grave{\boldsymbol{Q}}^{\mathrm{T}}$ with $\grave{\boldsymbol{Q}} \in \mathcal{BSC}_2^{M \times M}$ and a diagonal $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{\Lambda}} \in \mathcal{P}^M$.

Proof: The proof relies on $\grave{\boldsymbol{J}}_M$ from Lemma 10 and can be found in Appendix A. ■

Eigenvalue decompositions are not only ambiguous with respect to permutations of the eigenvalues and eigenvectors and to scaling of the eigenvectors, but also with respect to the choice of the bases of eigenspaces corresponding to eigenvalues with multiplicity larger than one. The latter is of particular importance for symmetric $\mathcal{BSC}_2$ matrices whose eigenvalues always have even multiplicity ($\lambda_{i+M} = \lambda_i$, $i = \{1, \ldots, M\}$ since $\grave{\boldsymbol{\Lambda}}$ is $\mathcal{BSC}_2$). This ambiguity is exploited in the proof of Lemma 11 to find a modal matrix that is $\mathcal{BSC}_2$.

Definition 9: We refer to an EVD of the form given in Lemma 11 as a *standard eigenvalue decomposition* of a symmetric $\mathcal{BSC}_2$ matrix.

[6]The EVD of a $\mathcal{BSC}$ matrix was previously considered in [44], but only for the special case that the blocks are themselves skew-circulant matrices.

With the standard EVD, we can prove the following lemma.

Lemma 12: A symmetric positive-semidefinite $\mathcal{BSC}_2$ matrix $\mathbf{0} \preceq \grave{\boldsymbol{A}} \in \mathcal{P}^M$ has a symmetric $\mathcal{BSC}_2$ square root $\grave{\boldsymbol{S}} \in \mathcal{P}^M$ with $\grave{\boldsymbol{A}} = \grave{\boldsymbol{S}}\grave{\boldsymbol{S}}$.

Proof: Let $\grave{\boldsymbol{S}} = \grave{\boldsymbol{Q}}\grave{\boldsymbol{\Lambda}}^{\frac{1}{2}}\grave{\boldsymbol{Q}}^{\mathrm{T}} = \grave{\boldsymbol{S}}^{\mathrm{T}}$, where $\grave{\boldsymbol{Q}}\grave{\boldsymbol{\Lambda}}\grave{\boldsymbol{Q}}^{\mathrm{T}} = \grave{\boldsymbol{A}}$ is a standard EVD of $\grave{\boldsymbol{A}}$. Since the square root of the diagonal matrix $\grave{\boldsymbol{\Lambda}}^{\frac{1}{2}}$ is an element-wise operation, it does not destroy the $\mathcal{BSC}_2$ structure, so that $\grave{\boldsymbol{S}}$ is $\mathcal{BSC}_2$ due to Lemma 3. ■

Similarly, we can obtain an EVD of a symmetric $\mathcal{BHSC}_2$ matrix.

Lemma 13: For a symmetric $\mathcal{BHSC}_2$ matrix $\acute{\boldsymbol{B}} \in \mathcal{N}^M$, the eigenvalue decomposition can be written as $\acute{\boldsymbol{B}} = \grave{\boldsymbol{Q}}\acute{\boldsymbol{\Lambda}}\grave{\boldsymbol{Q}}^{\mathrm{T}}$ with $\grave{\boldsymbol{Q}} \in \mathcal{BSC}_2^{M\times M}$ and a diagonal $\mathcal{BHSC}_2$ matrix $\acute{\boldsymbol{\Lambda}} \in \mathcal{N}^M$.

Proof: Apply the same reasoning as in the proof of Lemma 11, but use item 3 of Lemma 10 instead of item 2. ■

Note that it would also be possible to decompose $\acute{\boldsymbol{B}} \in \mathcal{N}^M$ as $\acute{\boldsymbol{B}} = \acute{\boldsymbol{Q}}\acute{\boldsymbol{\Lambda}}\acute{\boldsymbol{Q}}^{\mathrm{T}}$ with $\acute{\boldsymbol{Q}} \in \mathcal{BHSC}_2^{M\times M}$. Nevertheless, we choose to state the following definition.

Definition 10: We refer to an EVD of the form given in Lemma 13 as a *standard eigenvalue decomposition* of a symmetric $\mathcal{BHSC}_2$ matrix.

Lemma 14: Symmetric $\mathcal{BHSC}_2$ matrices $\acute{\boldsymbol{B}} \in \mathcal{N}^M$, $\acute{\boldsymbol{B}} \neq \mathbf{0}$ are indefinite.

Proof: Since $\acute{\boldsymbol{\Lambda}} \in \mathcal{N}^M$ in the standard EVD, $\lambda_{i+M} = -\lambda_i$, $i = \{1, \ldots, M\}$ (and $\lambda_i \neq 0$ for some i). ■

F. Singular Value Decompositions

When ignoring the convention that singular values are usually sorted in descending order, we obtain the following lemma for general $\mathcal{BSC}_2$ matrices.

Lemma 15: A reduced singular value decomposition of $\grave{\boldsymbol{A}} \in \mathcal{BSC}_2^{K\times L}$ is given by $\grave{\boldsymbol{A}} = \grave{\boldsymbol{U}}\grave{\boldsymbol{\Sigma}}\grave{\boldsymbol{V}}^{\mathrm{T}}$ with $\grave{\boldsymbol{U}} \in \mathcal{BSC}_2^{K\times M}$, $\grave{\boldsymbol{V}} \in \mathcal{BSC}_2^{L\times M}$, and a diagonal $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{\Sigma}} \in \mathcal{P}^M$, where $\grave{\boldsymbol{U}}^{\mathrm{T}}\grave{\boldsymbol{U}} = \grave{\boldsymbol{V}}^{\mathrm{T}}\grave{\boldsymbol{V}} = \mathbf{I}_{2M}$, and $M = \min\{K, L\}$.

Proof: See Appendix A. ■

For the proof of Lemma 15, it is again necessary to exploit the ambiguities of eigenvalue decompositions discussed below Lemma 11, which translate to ambiguities of the singular value decomposition.

Definition 11: We refer to an SVD of the form given in Lemma 15 as a *standard singular value decomposition* of a $\mathcal{BSC}_2$ matrix.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TSP.2015.2395992

In a similar manner, we can define a standard singular value decomposition of $\mathcal{BHSC}_2$ matrices.

Definition 12: We define the *standard singular value decomposition* of a $\mathcal{BHSC}_2$ matrix $\acute{\boldsymbol{B}} \in \mathcal{BHSC}_2^{K\times L}$ as follows. For $K \leq L$: $\acute{\boldsymbol{B}} = \grave{\boldsymbol{U}}\grave{\boldsymbol{\Sigma}}\acute{\boldsymbol{V}}^{\mathrm{T}}$ with $\grave{\boldsymbol{U}} \in \mathcal{BSC}_2^{K\times K}$, $\acute{\boldsymbol{V}} \in \mathcal{BHSC}_2^{L\times K}$, and a diagonal $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{\Sigma}} \in \mathcal{P}^K$, where $\grave{\boldsymbol{U}}^{\mathrm{T}}\grave{\boldsymbol{U}} = \acute{\boldsymbol{V}}^{\mathrm{T}}\acute{\boldsymbol{V}} = \mathbf{I}_{2K}$. For $K \geq L$: $\acute{\boldsymbol{B}} = \acute{\boldsymbol{U}}\grave{\boldsymbol{\Sigma}}\grave{\boldsymbol{V}}^{\mathrm{T}}$ with $\acute{\boldsymbol{U}} \in \mathcal{BHSC}_2^{K\times L}$, $\grave{\boldsymbol{V}} \in \mathcal{BSC}_2^{L\times L}$, and a diagonal matrix $\grave{\boldsymbol{\Sigma}} \in \mathcal{P}^L$, where $\acute{\boldsymbol{U}}^{\mathrm{T}}\acute{\boldsymbol{U}} = \grave{\boldsymbol{V}}^{\mathrm{T}}\grave{\boldsymbol{V}} = \mathbf{I}_{2L}$.

Lemma 16: A standard singular value decomposition as given in Definition 12 can be found for any $\grave{\boldsymbol{B}} \in \mathcal{BHSC}_2^{K\times L}$.

Proof: See Appendix A. ■

Note that we could also decompose a $\mathcal{BHSC}_2$ matrix into a product $\acute{\boldsymbol{B}} = \acute{\boldsymbol{U}}\acute{\boldsymbol{\Sigma}}\acute{\boldsymbol{V}}^{\mathrm{T}}$, where all matrices are $\mathcal{BHSC}_2$, $\acute{\boldsymbol{\Sigma}} \in \mathcal{N}^M$ is diagonal, and $\acute{\boldsymbol{U}}^{\mathrm{T}}\acute{\boldsymbol{U}} = \acute{\boldsymbol{V}}^{\mathrm{T}}\acute{\boldsymbol{V}} = \mathbf{I}_{2M}$, where $M = \min\{K, L\}$. However, this would not be an SVD since half of the nonzero diagonal elements of $\acute{\boldsymbol{\Sigma}}$ would be negative.

Above results on the SVD enable us to show the following properties.

Lemma 17: The rank of a $\mathcal{BSC}_2$ or $\mathcal{BHSC}_2$ matrix is even.

Proof: Follows from Lemma 15 and Lemma 16. ■

Lemma 18: The pseudoinverse of a $\mathcal{BSC}_2$ ($\mathcal{BHSC}_2$) matrix is $\mathcal{BSC}_2$ ($\mathcal{BHSC}_2$).

Proof: Using the standard SVD of $\grave{\boldsymbol{A}} = \grave{\boldsymbol{U}}\grave{\boldsymbol{\Sigma}}\grave{\boldsymbol{V}}^{\mathrm{T}} \in \mathcal{BSC}_2^{K\times L}$, we can write the pseudoinverse as $\grave{\boldsymbol{A}}^{+} = \grave{\boldsymbol{V}}\grave{\boldsymbol{\Sigma}}^{+}\grave{\boldsymbol{U}}^{\mathrm{T}}$. The matrix $\grave{\boldsymbol{\Sigma}}^{+}$ is obtained by inverting the nonzero diagonal elements of $\grave{\boldsymbol{\Sigma}}$ and leaving the zero elements unchanged, which does not destroy the $\mathcal{BSC}_2$ structure. Consequently, $\grave{\boldsymbol{A}}^{+} \in \mathcal{BSC}_2^{L\times K}$ due to Lemma 3. The proof for $\mathcal{BHSC}_2$ matrices follows the same lines. ■

IV. Relations to Complex Matrices

Using the properties derived above, we can discuss relations

$$\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{K\times L} \quad \leftrightarrow \quad \grave{\boldsymbol{A}} = \begin{bmatrix} \Re(\underset{\sim}{\boldsymbol{A}}) & -\Im(\underset{\sim}{\boldsymbol{A}}) \\ \Im(\underset{\sim}{\boldsymbol{A}}) & \Re(\underset{\sim}{\boldsymbol{A}}) \end{bmatrix} \in \mathcal{BSC}_2^{K\times L} \tag{23}$$

between complex matrices and real block-skew-circulant matrices.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TSP.2015.2395992

A. Fundamentals

We first translate some of the properties found in [10] to the framework of block-skew-circulant matrices.

Lemma 19: For real $\mathcal{BSC}_2$ matrices $\grave{\boldsymbol{A}}, \grave{\boldsymbol{C}}, \grave{\boldsymbol{D}}$ with the structure from (13) and their complex equivalents $\underset{\sim}{\boldsymbol{A}} = \boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2$, $\underset{\sim}{\boldsymbol{C}} = \boldsymbol{C}_1 + \mathrm{j}\boldsymbol{C}_2$, $\underset{\sim}{\boldsymbol{D}} = \boldsymbol{D}_1 + \mathrm{j}\boldsymbol{D}_2$, we have that

1) $\grave{\boldsymbol{A}} = \grave{\boldsymbol{C}}\grave{\boldsymbol{D}} \Leftrightarrow \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{D}}$,
2) $\grave{\boldsymbol{A}} = \grave{\boldsymbol{C}} + \grave{\boldsymbol{D}} \Leftrightarrow \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{C}} + \underset{\sim}{\boldsymbol{D}}$,
3) $\grave{\boldsymbol{A}} = \grave{\boldsymbol{C}}^{\mathrm{T}} \Leftrightarrow \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{C}}^{\mathrm{H}}$,
4) $\grave{\boldsymbol{A}} = \grave{\boldsymbol{C}}^{-1} \Leftrightarrow \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{C}}^{-1}$,
5) $\grave{\boldsymbol{A}}$ is orthonormal $\Leftrightarrow \underset{\sim}{\boldsymbol{A}}$ is unitary.

Proof: See [10, Lemma 1] and [10, Corollary 1]. ■

B. Eigenvalue Decompositions

In [49], a method to derive the eigenvalues of certain kinds of block matrices was presented. Using this approach, we can find the eigenvalues and eigenvectors of $\mathcal{BSC}_2$ matrices.

Lemma 20: $\grave{\boldsymbol{A}}\underset{\sim}{\boldsymbol{q}} = \underset{\sim}{\boldsymbol{q}}\underset{\sim}{\lambda}$ for the $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{A}}$ from (13) if and only if $(\boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2)\underset{\sim}{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{x}}\underset{\sim}{\varphi}$ with

$$\underset{\sim}{\boldsymbol{q}} = \begin{bmatrix} \underset{\sim}{\boldsymbol{x}} \\ -\mathrm{j}\underset{\sim}{\boldsymbol{x}} \end{bmatrix}, \ \underset{\sim}{\lambda} = \underset{\sim}{\varphi} \quad \text{or} \quad \underset{\sim}{\boldsymbol{q}} = \begin{bmatrix} \underset{\sim}{\boldsymbol{x}}^* \\ \mathrm{j}\underset{\sim}{\boldsymbol{x}}^* \end{bmatrix}, \ \underset{\sim}{\lambda} = \underset{\sim}{\varphi}^*. \tag{24}$$

Proof: See Appendix A. ■

The lemma implies that eigenvalues of $\mathcal{BSC}_2$ matrices are either real with even multiplicity (this is the case for all eigenvalues in Lemma 11) or they are complex conjugate pairs.

For symmetric $\grave{\boldsymbol{A}} \in \mathcal{P}^M$, the complex eigenvectors obtained in Lemma 20 do not contradict the real-valued ones obtained in Lemma 11: as a subspace of $\mathbb{C}^{2M}$, the two-dimensional eigenspace corresponding to the double eigenvalue $\underset{\sim}{\varphi} = \underset{\sim}{\varphi}^*$ is

$$\operatorname{span}\left\{ \begin{bmatrix} \Re\underset{\sim}{\boldsymbol{x}} \\ \Im\underset{\sim}{\boldsymbol{x}} \end{bmatrix}, \begin{bmatrix} -\Im\underset{\sim}{\boldsymbol{x}} \\ \Re\underset{\sim}{\boldsymbol{x}} \end{bmatrix} \right\} = \operatorname{span}\left\{ \begin{bmatrix} \underset{\sim}{\boldsymbol{x}} \\ -\mathrm{j}\underset{\sim}{\boldsymbol{x}} \end{bmatrix}, \begin{bmatrix} \underset{\sim}{\boldsymbol{x}}^* \\ \mathrm{j}\underset{\sim}{\boldsymbol{x}}^* \end{bmatrix} \right\}. \tag{25}$$

This means that Lemma 20 does not directly deliver a standard EVD of symmetric $\mathcal{BSC}_2$ matrices, but that the result can be easily transformed to a standard EVD using (25).

For [10, Lemma 1, (4e)], which is reproduced in the following, Lemma 20 delivers an alternative proof.

Lemma 21: For a $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{A}}$ as in (13),

$$\begin{aligned}\det(\grave{\boldsymbol{A}}) &= \det(\boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2)\det(\boldsymbol{A}_1 - \mathrm{j}\boldsymbol{A}_2) \\ &= \det\left((\boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2)(\boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2)^{\mathrm{H}}\right). \end{aligned} \tag{26}$$

Proof: Follows from Lemma 20 since the determinant of a square matrix is the product of its eigenvalues. ■

In addition, we can provide an alternative proof for the following Lemma, which is known from [10, Corollary 2] and [50, Problem 4.2.1].

Lemma 22: $\grave{\boldsymbol{A}} \in \mathcal{P}^M$ with the block structure from (13) is positive semidefinite if and only if $\boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2$ is positive semidefinite.

Proof: Follows from Lemma 20 and from the fact that a Hermitian matrix is positive semidefinite if and only if all its eigenvalues are nonnegative. ■

To conclude this subsection, we turn our attention to $\mathcal{BHSC}_2$ matrices, whose eigenvalues can be obtained from [49] as well.

Lemma 23: $\underset{\sim}{\varphi}$ is an eigenvalue of the $\mathcal{BHSC}_2$ matrix $\acute{\boldsymbol{B}}$ from (13) if and only if $\underset{\sim}{\varphi}^2$ is an eigenvalue of $\boldsymbol{B}_1^2 + \boldsymbol{B}_2^2 + \mathrm{j}(\boldsymbol{B}_1\boldsymbol{B}_2 - \boldsymbol{B}_2\boldsymbol{B}_1)$.

Proof: See [49, Example below Th. 5]. ■

Consequently, the eigenvalues of real $\mathcal{BHSC}_2$ matrices must form real pairs $\underset{\sim}{\varphi}, -\underset{\sim}{\varphi}$ or complex quads $\underset{\sim}{\varphi}, \underset{\sim}{\varphi}^*, -\underset{\sim}{\varphi}, -\underset{\sim}{\varphi}^*$. Note that this confirms Lemma 4.

C. Singular Value Decompositions

We can also relate the singular values and singular vectors of $\grave{\boldsymbol{A}}$ to the singular values and singular vectors of $\boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2$.

Lemma 24: The set of singular values of $\grave{\boldsymbol{A}} \in \mathcal{BSC}_2^{K\times L}$ from (13) is the same as the set of singular values of $\boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2$, but the multiplicity of each singular value is doubled.

Proof: Apply Lemma 20 and (25) to $\grave{\boldsymbol{A}}\grave{\boldsymbol{A}}^{\mathrm{T}}$ or $\grave{\boldsymbol{A}}^{\mathrm{T}}\grave{\boldsymbol{A}}$. ■

This delivers a simple proof for the following lemma, which was stated in our previous work [32] without a detailed proof (due to lack of space).

Lemma 25: The rank of the $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{A}}$ from (13) fulfills $\mathrm{Rank}[\grave{\boldsymbol{A}}] = 2\,\mathrm{Rank}[\boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2]$.

Proof: Follows from Lemma 24. ∎

Note that this property was also shown in [51], but only for Hermitian matrices.

V. Linear and Widely Linear Operations

In Section II, we have seen that a complex linear mapping, i.e., the multiplication of a matrix $\underset{\sim}{\boldsymbol{A}}$ and a vector $\underset{\sim}{\boldsymbol{x}}$, can be written as the multiplication of the corresponding $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{A}}$ and the composite real vector $\check{\boldsymbol{x}}$ [cf. (7)].

In a similar manner, we can study a conjugate linear mapping in the composite real representation.

Theorem 1: A conjugate linear mapping $\underset{\sim}{\boldsymbol{x}} \mapsto \underset{\sim}{\boldsymbol{f}}_{\mathrm{CL}}(\underset{\sim}{\boldsymbol{x}}) = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{x}}^*$ can be represented by the multiplication of a $\mathcal{BHSC}_2$ matrix and the composite real vector $\check{\boldsymbol{x}}$.

Proof: Let

$$\acute{\boldsymbol{J}}_N = \begin{bmatrix} \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_N \end{bmatrix} \in \mathcal{BHSC}_2^{N\times N}. \tag{27}$$

The composite real representation of $\underset{\sim}{\boldsymbol{x}}^* \in \mathbb{C}^N$ is given by $\acute{\boldsymbol{J}}_N \check{\boldsymbol{x}} \in \mathbb{R}^{2N}$. Consequently, $\check{\boldsymbol{f}}_{\mathrm{CL}}(\check{\boldsymbol{x}}) = \grave{\boldsymbol{B}}\acute{\boldsymbol{J}}_N\check{\boldsymbol{x}} = \acute{\boldsymbol{B}}\check{\boldsymbol{x}}$, where $\grave{\boldsymbol{B}} \in \mathcal{BSC}_2^{M\times N}$ is the real-valued equivalent (23) of $\underset{\sim}{\boldsymbol{B}} \in \mathbb{C}^{M\times N}$, and $\acute{\boldsymbol{B}} \in \mathcal{BHSC}_2^{M\times N}$ due to Lemma 3. ∎

This enables us to give a second variant of the composite real representation of a widely linear mapping.

Theorem 2: The composite real representation (6) of a widely linear mapping of the form (2) is given by

$$\check{\boldsymbol{f}}_{\mathrm{WL}}(\check{\boldsymbol{x}}) = \boldsymbol{A}_{\mathrm{WL}}\check{\boldsymbol{x}} = (\grave{\boldsymbol{A}}_{\mathrm{L}} + \acute{\boldsymbol{A}}_{\mathrm{CL}})\check{\boldsymbol{x}} = \grave{\boldsymbol{A}}_{\mathrm{L}}\check{\boldsymbol{x}} + \grave{\boldsymbol{A}}_{\mathrm{CL}}\acute{\boldsymbol{J}}_N\check{\boldsymbol{x}} \tag{28}$$

with $\acute{\boldsymbol{J}}_N$ from (27), where the $\mathcal{BSC}_2$ matrices $\grave{\boldsymbol{A}}_{\mathrm{L}}$ and $\grave{\boldsymbol{A}}_{\mathrm{CL}}$ are the real-valued equivalents (23) of $\underset{\sim}{\boldsymbol{A}}_{\mathrm{L}}$ and $\underset{\sim}{\boldsymbol{A}}_{\mathrm{CL}}$, respectively. The complex form (2) of a widely linear mapping is obtained from the composite real representation (6) by setting $\underset{\sim}{\boldsymbol{A}}_{\mathrm{L}}$ and $\underset{\sim}{\boldsymbol{A}}_{\mathrm{CL}}$ to the complex equivalents (23) of the $\mathcal{BSC}_2$ matrices

$$\grave{\boldsymbol{A}}_{\mathrm{L}} = \mathbb{P}_{\mathcal{BSC}_2}(\boldsymbol{A}_{\mathrm{WL}}) \quad \text{and} \quad \grave{\boldsymbol{A}}_{\mathrm{CL}} = \mathbb{P}_{\mathcal{BHSC}_2}(\boldsymbol{A}_{\mathrm{WL}})\acute{\boldsymbol{J}}_N \tag{29}$$

respectively.

Proof: Follows from Theorem 1 and (7). ∎

Corollary 1: A widely linear mapping is linear if the composite real representation (6) fulfills $\mathbb{P}_{\mathcal{BHSC}_2}(\boldsymbol{A}_{\mathrm{WL}}) = \mathbf{0}$.

VI. Second-Order Properties of Complex Random Vectors

In this section, we apply the proposed framework in combination with the definitions given in Section II to characterize complex random vectors.

A. Fundamentals

Let $\underset{\sim}{\boldsymbol{x}} \in \mathbb{C}^M$ be a complex random vector, $\breve{\boldsymbol{x}} \in \mathbb{R}^{2M}$ its composite real representation, and $\boldsymbol{C}_{\breve{\boldsymbol{x}}} \in \mathbb{R}^{2M\times 2M}$ the (real-valued) covariance matrix of the composite real vector $\breve{\boldsymbol{x}}$ with the block structure given in (8).

In our recent work [34], we decomposed $\boldsymbol{C}_{\breve{\boldsymbol{x}}}$ into a term depending on the complex covariance matrix and a term depending on the pseudocovariance matrix. In the following, we show how this decomposition fits into the framework of block-skew-circulant matrices.

Theorem 3: The $\mathcal{BSC}_2$ component $\mathbb{P}_{\mathcal{BSC}_2}(\boldsymbol{C}_{\breve{\boldsymbol{x}}})$ of the composite real covariance matrix depends only on the complex covariance matrix $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$. The $\mathcal{BHSC}_2$ component $\mathbb{P}_{\mathcal{BHSC}_2}(\boldsymbol{C}_{\breve{\boldsymbol{x}}})$ of the composite real covariance matrix depends only on the pseudocovariance matrix $\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}$.

Proof: The covariance matrix and pseudocovariance matrix of the complex random vector $\underset{\sim}{\boldsymbol{x}}$ can be expressed as [5]

$$\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}} = \boldsymbol{C}_{\Re\underset{\sim}{\boldsymbol{x}}} + \boldsymbol{C}_{\Im\underset{\sim}{\boldsymbol{x}}} + \mathrm{j}(\boldsymbol{C}^{\mathrm{T}}_{\Re\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{x}}} - \boldsymbol{C}_{\Re\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{x}}}) \tag{30}$$

$$\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}} = \boldsymbol{C}_{\Re\underset{\sim}{\boldsymbol{x}}} - \boldsymbol{C}_{\Im\underset{\sim}{\boldsymbol{x}}} + \mathrm{j}(\boldsymbol{C}^{\mathrm{T}}_{\Re\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{x}}} + \boldsymbol{C}_{\Re\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{x}}}). \tag{31}$$

Solving for $\boldsymbol{C}_{\Re\underset{\sim}{\boldsymbol{x}}}$, $\boldsymbol{C}_{\Im\underset{\sim}{\boldsymbol{x}}}$, and $\boldsymbol{C}_{\Re\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{x}}}$ yields

$$\boldsymbol{C}_{\breve{\boldsymbol{x}}} = \underbrace{\frac{1}{2}\begin{bmatrix}\Re(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}) & -\Im(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}) \\ \Im(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}) & \Re(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}})\end{bmatrix}}_{\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}} + \underbrace{\frac{1}{2}\begin{bmatrix}\Re(\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}) & \Im(\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}) \\ \Im(\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}) & -\Re(\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}})\end{bmatrix}}_{\acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}}}. \tag{32}$$

Obviously, $\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}$ has the structure of a $\mathcal{BSC}_2$ matrix while $\acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}}$ has the structure of a $\mathcal{BHSC}_2$ matrix. ■

Corollary 2: A complex random vector $\underset{\sim}{\boldsymbol{x}}$ is proper if the impropriety matrix $\acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}}$ vanishes, i.e., if $\mathbb{P}_{\mathcal{BHSC}_2}(\boldsymbol{C}_{\breve{\boldsymbol{x}}}) = \boldsymbol{0}$. Otherwise, $\underset{\sim}{\boldsymbol{x}}$ is improper.

Note that both $\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}$ and $\acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}}$ must be symmetric to obtain a symmetric matrix $\boldsymbol{C}_{\breve{\boldsymbol{x}}}$. Thus, $\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}$ lies in the power shaping space $\mathcal{P}^M$ defined in (20) and $\acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}}$ lies in the impropriety space $\mathcal{N}^M$ defined in (21). This makes clear why the names for $\mathcal{P}^M$ and $\mathcal{N}^M$ are a sensible choice. While

$\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}} \in \mathcal{P}^M$ depends only on the complex covariance matrix, which describes the power shaping of the complex random vector $\underset{\sim}{\boldsymbol{x}}$, $\acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}} \in \mathcal{N}^N$ is determined by the pseudocovariance matrix, which describes the impropriety (or noncircularity).

By summing up arbitrary elements of $\mathcal{P}^M$ and $\mathcal{N}^M$, we obtain a symmetric $\boldsymbol{C}_{\check{\boldsymbol{x}}}$, but this is not sufficient to obtain a valid covariance matrix. In addition, we need that $\boldsymbol{C}_{\check{\boldsymbol{x}}} \succeq \mathbf{0}$, i.e., $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ must be positive semidefinite. For this, we find necessary and sufficient conditions in the following.

In the complex formulation, it is a common practice to check whether the generalized Schur complement of the augmented covariance matrix, $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}} - \tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}(\underset{\sim}{\boldsymbol{C}}^*_{\underset{\sim}{\boldsymbol{x}}})^+\tilde{\underset{\sim}{\boldsymbol{C}}}^*_{\underset{\sim}{\boldsymbol{x}}}$, is positive semidefinite. Together with the nullspace condition $\mathrm{null}[\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}] \subseteq \mathrm{null}[\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}]$, positive-semidefiniteness of $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$, and symmetry of $\tilde{\underset{\sim}{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}$, this is a necessary and sufficient condition for having a valid pair of covariance matrix and pseudocovariance matrix [4], [52]. The composite real counterpart of this test reads as follows.

Theorem 4: For $\grave{\boldsymbol{P}} \in \mathcal{P}^M$ and $\acute{\boldsymbol{N}} \in \mathcal{N}^M$, we have $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} \succeq \mathbf{0}$ if and only if

1) $\grave{\boldsymbol{P}} \succeq \mathbf{0}$,
2) $\grave{\boldsymbol{P}} - \acute{\boldsymbol{N}}\grave{\boldsymbol{P}}^+\acute{\boldsymbol{N}} \succeq \mathbf{0}$, and
3) $\mathrm{null}[\grave{\boldsymbol{P}}] \subseteq \mathrm{null}[\acute{\boldsymbol{N}}]$.

Proof: $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} \succeq \mathbf{0}$ means that $\boldsymbol{x}^{\mathrm{T}}(\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}})\boldsymbol{x} \geq 0,\ \forall \boldsymbol{x} \in \mathbb{R}^{2M}$. In particular, the inequality also holds for $\boldsymbol{x}' = \grave{\boldsymbol{J}}_M\boldsymbol{x}$ with $\grave{\boldsymbol{J}}_M$ from (22). Using Lemma 10, we have

$$0 \leq \boldsymbol{x}^{\mathrm{T}}(\grave{\boldsymbol{J}}_M^{\mathrm{T}}\grave{\boldsymbol{P}}\grave{\boldsymbol{J}}_M + \grave{\boldsymbol{J}}_M^{\mathrm{T}}\acute{\boldsymbol{N}}\grave{\boldsymbol{J}}_M)\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}(\grave{\boldsymbol{P}} - \acute{\boldsymbol{N}})\boldsymbol{x} \tag{33}$$

i.e., we have $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} \succeq \mathbf{0} \Leftrightarrow \grave{\boldsymbol{P}} - \acute{\boldsymbol{N}} \succeq \mathbf{0}$. Thus, $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} \succeq \mathbf{0}$ is equivalent to

$$\begin{bmatrix} \grave{\boldsymbol{P}} & \acute{\boldsymbol{N}} \\ \acute{\boldsymbol{N}} & \grave{\boldsymbol{P}} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \mathbf{I}_M & \mathbf{I}_M \\ \mathbf{I}_M & -\mathbf{I}_M \end{bmatrix} \underbrace{\begin{bmatrix} \grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} & \\ & \grave{\boldsymbol{P}} - \acute{\boldsymbol{N}} \end{bmatrix}}_{\succeq \mathbf{0}} \begin{bmatrix} \mathbf{I}_M & \mathbf{I}_M \\ \mathbf{I}_M & -\mathbf{I}_M \end{bmatrix} \succeq \mathbf{0}. \tag{34}$$

Since $\grave{\boldsymbol{P}} - \acute{\boldsymbol{N}}\grave{\boldsymbol{P}}^+\acute{\boldsymbol{N}}$ is the generalized Schur complement of $\grave{\boldsymbol{P}}$ in the matrix in (34), we can conclude from [53, Appendix A.5] that (34) is equivalent to $\grave{\boldsymbol{P}} \succeq \mathbf{0}$, $\grave{\boldsymbol{P}} - \acute{\boldsymbol{N}}\grave{\boldsymbol{P}}^+\acute{\boldsymbol{N}} \succeq \mathbf{0}$, and $\mathrm{range}[\acute{\boldsymbol{N}}] \subseteq \mathrm{range}[\grave{\boldsymbol{P}}]$. Due to $\acute{\boldsymbol{P}}^{\mathrm{T}} = \acute{\boldsymbol{P}}$ and $\acute{\boldsymbol{N}}^{\mathrm{T}} = \acute{\boldsymbol{N}}$, this is equivalent to the conditions in the theorem. ■

Unlike in the complex formulation, it is questionable whether a test via the Schur complement makes sense in the composite real representation. Instead we can directly verify whether the sum $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}}$ is positive-semidefinite, which is a much simpler test. Nevertheless, the positive-semidefiniteness condition and the nullspace condition turn out to be very useful for characterizing pairs of power shaping matrices and impropriety matrices (see, e.g., next subsection). Therefore, we summarize them in the following corollary.

Corollary 3: For any valid pair of power shaping matrix $\grave{\boldsymbol{P}} \in \mathcal{P}^M$ and impropriety matrix $\acute{\boldsymbol{N}} \in \mathcal{N}^M$, we have

1) $\grave{\boldsymbol{P}} \succeq \boldsymbol{0}$,
2) $\mathrm{null}[\grave{\boldsymbol{P}}] \subseteq \mathrm{null}[\acute{\boldsymbol{N}}]$,
3) $\mathrm{null}[\grave{\boldsymbol{P}}] \subseteq \mathrm{null}[\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}}]$,
4) $\mathrm{Rank}[\acute{\boldsymbol{N}}] \leq \mathrm{Rank}[\grave{\boldsymbol{P}}]$.

Proof: All items directly follow from Theorem 4, but unlike Theorem 4, this weaker statement can be proven without making use of the properties of the generalized Schur complement. See Appendix A for this alternative proof. ■

B. Cross-Covariances and Joint Propriety

Let $\underset{\sim}{\boldsymbol{x}} \in \mathbb{C}^M$ and $\underset{\sim}{\boldsymbol{y}} \in \mathbb{C}^N$ be complex random vectors. For simplicity, we assume $\mathrm{E}\big[\underset{\sim}{\boldsymbol{x}}\big] = \boldsymbol{0}$ and $\mathrm{E}\big[\underset{\sim}{\boldsymbol{y}}\big] = \boldsymbol{0}$, but the considerations can be easily extended to vectors with nonzero mean.

The vectors $\underset{\sim}{\boldsymbol{x}}$ and $\underset{\sim}{\boldsymbol{y}}$ are called *jointly proper* if the composite vector $\underset{\sim}{\boldsymbol{z}}^{\mathrm{T}} = [\underset{\sim}{\boldsymbol{x}}^{\mathrm{T}}, \underset{\sim}{\boldsymbol{y}}^{\mathrm{T}}] \in \mathbb{C}^{M+N}$ is proper (e.g., [4, Section 2.3]).

Theorem 5: $\underset{\sim}{\boldsymbol{x}} \in \mathbb{C}^M$ and $\underset{\sim}{\boldsymbol{y}} \in \mathbb{C}^N$ are jointly proper if and only if $\boldsymbol{C}_{\check{\boldsymbol{x}}} \in \mathcal{BSC}_2^{M\times M}$, $\boldsymbol{C}_{\check{\boldsymbol{y}}} \in \mathcal{BSC}_2^{N\times N}$ and

$$\boldsymbol{C}_{\check{\boldsymbol{x}},\check{\boldsymbol{y}}} = \begin{bmatrix} \mathrm{E}\big[\Re\underset{\sim}{\boldsymbol{x}}\Re\underset{\sim}{\boldsymbol{y}}^{\mathrm{T}}\big] & \mathrm{E}\big[\Re\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{y}}^{\mathrm{T}}\big] \\ \mathrm{E}\big[\Im\underset{\sim}{\boldsymbol{x}}\Re\underset{\sim}{\boldsymbol{y}}^{\mathrm{T}}\big] & \mathrm{E}\big[\Im\underset{\sim}{\boldsymbol{x}}\Im\underset{\sim}{\boldsymbol{y}}^{\mathrm{T}}\big] \end{bmatrix} \in \mathcal{BSC}_2^{M\times N}. \tag{35}$$

Proof: We have to verify whether

$$\mathrm{E}\left[\check{\boldsymbol{z}}\check{\boldsymbol{z}}^{\mathrm{T}}\right] = \left[\begin{array}{cc:cc} \mathrm{E}\left[\Re\boldsymbol{x}\Re\boldsymbol{x}^{\mathrm{T}}\right] & \mathrm{E}\left[\Re\boldsymbol{x}\Re\boldsymbol{y}^{\mathrm{T}}\right] & \mathrm{E}\left[\Re\boldsymbol{x}\Im\boldsymbol{x}^{\mathrm{T}}\right] & \mathrm{E}\left[\Re\boldsymbol{x}\Im\boldsymbol{y}^{\mathrm{T}}\right] \\ \mathrm{E}\left[\Re\boldsymbol{y}\Re\boldsymbol{x}^{\mathrm{T}}\right] & \mathrm{E}\left[\Re\boldsymbol{y}\Re\boldsymbol{y}^{\mathrm{T}}\right] & \mathrm{E}\left[\Re\boldsymbol{y}\Im\boldsymbol{x}^{\mathrm{T}}\right] & \mathrm{E}\left[\Re\boldsymbol{y}\Im\boldsymbol{y}^{\mathrm{T}}\right] \\ \hdashline \mathrm{E}\left[\Im\boldsymbol{x}\Re\boldsymbol{x}^{\mathrm{T}}\right] & \mathrm{E}\left[\Im\boldsymbol{x}\Re\boldsymbol{y}^{\mathrm{T}}\right] & \mathrm{E}\left[\Im\boldsymbol{x}\Im\boldsymbol{x}^{\mathrm{T}}\right] & \mathrm{E}\left[\Im\boldsymbol{x}\Im\boldsymbol{y}^{\mathrm{T}}\right] \\ \mathrm{E}\left[\Im\boldsymbol{y}\Re\boldsymbol{x}^{\mathrm{T}}\right] & \mathrm{E}\left[\Im\boldsymbol{y}\Re\boldsymbol{y}^{\mathrm{T}}\right] & \mathrm{E}\left[\Im\boldsymbol{y}\Im\boldsymbol{x}^{\mathrm{T}}\right] & \mathrm{E}\left[\Im\boldsymbol{y}\Im\boldsymbol{y}^{\mathrm{T}}\right] \end{array}\right] \tag{36}$$

$\in \mathbb{R}^{2(M+N)\times 2(M+N)}$ is a $\mathcal{BSC}_2$ matrix. To this end, we have to partition the matrix into four blocks consisting of four subblocks each, as indicated by the dashed lines. Clearly, we have equality of a pair of blocks if and only if the equality holds for each of the four pairs of subblocks. ∎

The intuition behind Theorem 5 is that $\boldsymbol{x}$ and $\boldsymbol{y}$ must be individually proper, the correlation between the two real parts must be the same as the correlation between the two imaginary parts, and the correlation between $\Re\boldsymbol{x}$ and $\Im\boldsymbol{y}$ must be minus the correlation between $\Im\boldsymbol{x}$ and $\Re\boldsymbol{y}$.

In [4, Section 2.3], the last two conditions were subsumed under the notion of cross-propriety between $\boldsymbol{x}$ and $\boldsymbol{y}$: $\boldsymbol{x}$ and $\boldsymbol{y}$ are called *cross-proper*, if the pseudo-cross-covariance matrix $\tilde{\boldsymbol{C}}_{\boldsymbol{x},\boldsymbol{y}}$ vanishes. To transfer this to the framework of block-skew-circulant matrices, let us apply the partitioning from (32) to the composite real cross-covariance matrix $\boldsymbol{C}_{\check{\boldsymbol{x}},\check{\boldsymbol{y}}}$:

$$\boldsymbol{C}_{\check{\boldsymbol{x}},\check{\boldsymbol{y}}} = \grave{\boldsymbol{P}}_{\boldsymbol{x},\boldsymbol{y}} + \acute{\boldsymbol{N}}_{\boldsymbol{x},\boldsymbol{y}} = \frac{1}{2}\begin{bmatrix} \Re(\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}) & -\Im(\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}) \\ \Im(\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}) & \Re(\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}) \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \Re(\tilde{\boldsymbol{C}}_{\boldsymbol{x},\boldsymbol{y}}) & \Im(\tilde{\boldsymbol{C}}_{\boldsymbol{x},\boldsymbol{y}}) \\ \Im(\tilde{\boldsymbol{C}}_{\boldsymbol{x},\boldsymbol{y}}) & -\Re(\tilde{\boldsymbol{C}}_{\boldsymbol{x},\boldsymbol{y}}) \end{bmatrix} \tag{37}$$

where $\grave{\boldsymbol{P}}_{\boldsymbol{x},\boldsymbol{y}} \in \mathcal{BSC}_2^{M\times N}$ and $\acute{\boldsymbol{N}}_{\boldsymbol{x},\boldsymbol{y}} \in \mathcal{BHSC}_2^{M\times N}$. This shows that the composite real cross-covariance matrix can be partitioned into a cross-power-shaping matrix $\grave{\boldsymbol{P}}_{\boldsymbol{x},\boldsymbol{y}}$ with $\mathcal{BSC}_2$ structure and a cross-impropriety matrix $\acute{\boldsymbol{N}}_{\boldsymbol{x},\boldsymbol{y}}$ with $\mathcal{BHSC}_2$ structure. We have $\acute{\boldsymbol{N}}_{\boldsymbol{x},\boldsymbol{y}} = \boldsymbol{0}$ if and only if $\tilde{\boldsymbol{C}}_{\boldsymbol{x},\boldsymbol{y}} = \boldsymbol{0}$, i.e., if $\boldsymbol{x}$ and $\boldsymbol{y}$ are cross-proper.

Since cross-covariance matrices are not necessarily symmetric and in general not even square matrices, $\grave{\boldsymbol{P}}_{\boldsymbol{x},\boldsymbol{y}}$ and $\acute{\boldsymbol{N}}_{\boldsymbol{x},\boldsymbol{y}}$ are in general not elements of the power shaping space $\mathcal{P}^M$ or the impropriety space $\mathcal{N}^M$ defined in Definition 8.

C. Entropy

By definition, the differential entropy of a complex random vector is the same as the differential entropy of the composite real vector (e.g., [4, Section 2.2.3]), i.e., $\mathrm{h}(\underset{\sim}{\boldsymbol{x}}) = \mathrm{h}(\check{\boldsymbol{x}})$. For a general (proper or improper) complex Gaussian random vector $\underset{\sim}{\boldsymbol{x}} \in \mathbb{C}^M$, we therefore have

$$\begin{aligned}
\mathrm{h}(\underset{\sim}{\boldsymbol{x}}) = \mathrm{h}(\check{\boldsymbol{x}}) &= \frac{1}{2} \log\det(2\pi\mathrm{e}\boldsymbol{C}_{\check{\boldsymbol{x}}}) \\
&= \frac{1}{2} \log\det\left(2\pi\mathrm{e}(\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}} + \acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}})\right) \\
&= \frac{1}{2} \log\det(2\pi\mathrm{e}\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}) + \frac{1}{2} \log\det\left(\mathbf{I}_{2M} + \grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}^{-1} \acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}}\right)
\end{aligned} \tag{38}$$

if $\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}^{-1}$ exists, i.e., if $\det(\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}) \neq 0$. Due to item 4) of Corollary 3 and due to Lemma 25, this is always the case if $\det(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}) \neq 0$, i.e., if the distribution is not degenerate. This leads to the following theorem.

Theorem 6: The differential entropy of a complex Gaussian random vector $\underset{\sim}{\boldsymbol{x}}$ with $\det(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}) \neq 0$ can be decomposed into

$$\mathrm{h}(\underset{\sim}{\boldsymbol{x}}) = \mathrm{h}_{\text{proper}}(\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}) + \underbrace{\Delta\,\mathrm{h}(\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}, \acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}})}_{\leq 0} \tag{39}$$

where $\mathrm{h}_{\text{proper}}(\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}})$ is the differential entropy of a proper complex Gaussian random vector with the same power shaping matrix $\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}$, and $\Delta\,\mathrm{h}(\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}, \acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}})$ is a nonpositive correction term accounting for the possible impropriety of $\underset{\sim}{\boldsymbol{x}}$, which vanishes if and only if $\acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}} = \mathbf{0}$.

Proof: See Appendix A. ∎

Corollary 4: The differential entropy of a complex Gaussian random vector $\underset{\sim}{\boldsymbol{x}}$ is maximized if $\acute{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}} = \mathbf{0}$, i.e., if $\mathbb{P}_{\mathcal{BHSC}_2}(\boldsymbol{C}_{\check{\boldsymbol{x}}}) = \mathbf{0}$.

So far, such a decomposition had only been reported based on the formulation with covariance matrix and pseudocovariance matrix, where $\mathrm{h}(\underset{\sim}{\boldsymbol{x}})$ can be written as (e.g. [29])

$$\mathrm{h}(\underset{\sim}{\boldsymbol{x}}) = \underbrace{\log\det(\pi\mathrm{e}\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}})}_{\mathrm{h}_{\text{proper}}\left(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}\right)} + \underbrace{\frac{1}{2}\log\det(\mathbf{I}_M - \underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}^{-*} \underset{\sim}{\tilde{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}^{\mathrm{H}} \underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}^{-1} \underset{\sim}{\tilde{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}})}_{\Delta\,\mathrm{h}\left(\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}, \underset{\sim}{\tilde{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}\right) \leq 0}. \tag{40}$$

Recalling that there is a bijective mapping between $\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}$ and $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$ as well as between $\grave{\boldsymbol{N}}_{\underset{\sim}{\boldsymbol{x}}}$ and $\underset{\sim}{\tilde{\boldsymbol{C}}}_{\underset{\sim}{\boldsymbol{x}}}$, we can see the connection between (39) and (40). However, it is important to note that the

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TSP.2015.2395992

framework of block-skew-circulant matrices has delivered a derivation for (39) that is independent from the existing literature on complex-valued signal processing.

VII. Application Examples

Having established the framework of block-skew-circulant matrices for complex-valued signal processing, we now present application examples to demonstrate the usefulness of this framework. The examples come from the area of signal processing for communications since our focus in studying improper signals has been in this area. However, we are confident that there is a large potential for the application of the framework also in other areas of complex-valued signal processing.

A. Widely Linear MMSE Filters

If we want to estimate a zero-mean random vector $\underset{\sim}{\boldsymbol{x}} \in \mathbb{C}^M$ from a zero-mean observation $\underset{\sim}{\boldsymbol{y}}$ by means of a widely linear minimum mean square error (MMSE) estimator, we can do so by applying a linear MMSE estimator to the composite real representation $\check{\boldsymbol{y}}$, i.e., the complex estimate $\hat{\underset{\sim}{\boldsymbol{x}}}$ is given by

$$\hat{\underset{\sim}{\boldsymbol{x}}} = \begin{bmatrix} \mathbf{I}_M & \mathrm{j}\mathbf{I}_M \end{bmatrix} \boldsymbol{G}_{\mathrm{LMMSE}}\, \check{\boldsymbol{y}} = \begin{bmatrix} \mathbf{I}_M & \mathrm{j}\mathbf{I}_M \end{bmatrix} \boldsymbol{C}_{\check{\boldsymbol{x}},\check{\boldsymbol{y}}} \boldsymbol{C}_{\check{\boldsymbol{y}}}^{+} \check{\boldsymbol{y}}. \tag{41}$$

If $\underset{\sim}{\boldsymbol{x}}$ and $\underset{\sim}{\boldsymbol{y}}$ are jointly proper, $\boldsymbol{C}_{\check{\boldsymbol{y}}}$ and $\boldsymbol{C}_{\check{\boldsymbol{y}},\check{\boldsymbol{x}}}$ are $\mathcal{BSC}_2$ matrices (see Corollary 2 and Theorem 5). Then, due to Lemmas 18 and 3, we have that $\boldsymbol{G}_{\mathrm{LMMSE}}$ is $\mathcal{BSC}_2$, which corresponds to a (strictly) linear complex filter (see Corollary 1).

From this, we see that an MMSE-optimal widely linear filter becomes linear if the involved signals are jointly proper. This fact is well known (e.g., [6], [11]), but the framework of block-skew-circulant matrices delivers an alternative proof.

B. MIMO Capacity

In [10], it was shown that proper Gaussian signals are the capacity-achieving input distribution of a point-to-point multiple-input multiple-output (MIMO) communication system with proper Gaussian noise. Using the framework of block-skew-circulant matrices, we can easily extend this result to general complex Gaussian noise.

To this end, we apply the derivation of the MIMO capacity given in [10] to the composite real representation with the channel matrix $\grave{\boldsymbol{H}} \in \mathcal{BSC}_2^{N\times M}$ and noise covariance matrix $\boldsymbol{C}_{\check{\boldsymbol{n}}} \in \mathbb{S}^{2N}$. We obtain that real-valued Gaussian inputs with covariance matrix $\boldsymbol{C}_{\check{\boldsymbol{x}}} = \boldsymbol{U}\boldsymbol{Q}\boldsymbol{U}^{\mathrm{T}}$ achieve the capacity, where $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}} = \grave{\boldsymbol{H}}^{\mathrm{T}}\boldsymbol{C}_{\check{\boldsymbol{n}}}^{-1}\grave{\boldsymbol{H}}$ is an eigenvalue decomposition, and $[\boldsymbol{Q}]_{i,i} = \max\left\{\mu - \frac{1}{\lambda_i},\ 0\right\}$ is a waterfilling power allocation (see, e.g., [54]).

In the special case that the noise is proper, we have $\boldsymbol{C}_{\check{\boldsymbol{n}}} = \grave{\boldsymbol{C}}_{\check{\boldsymbol{n}}} \in \mathcal{P}^N \subset \mathcal{BSC}_2^{N\times N}$. In this case, $\grave{\boldsymbol{H}}^{\mathrm{T}}\grave{\boldsymbol{C}}_{\boldsymbol{n}}^{-1}\grave{\boldsymbol{H}} \in \mathcal{P}^M \subset \mathcal{BSC}_2^{M\times M}$ due to Lemmas 1, 3, and 8. Using the standard EVD (see Lemma 11), we then obtain a block-skew-circulant $\boldsymbol{U}$ and a block-skew-circulant $\boldsymbol{\Lambda}$. The latter means that $\lambda_i = \lambda_{i+M}$ for $i = 1, \ldots, M$, which implies $[\boldsymbol{Q}]_{i,i} = [\boldsymbol{Q}]_{i+M,i+M}$. Finally, as a product of $\mathcal{BSC}_2$ matrices, $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ is a $\mathcal{BSC}_2$ matrix as well, meaning that the optimal input $\boldsymbol{x}$ for the complex system is proper.

On the other hand, if the noise $\boldsymbol{n}$ is improper, $\boldsymbol{C}_{\check{\boldsymbol{n}}}$ is not a $\mathcal{BSC}_2$ matrix, and the optimal $\boldsymbol{C}_{\check{\boldsymbol{x}}}$ is (in general) not a $\mathcal{BSC}_2$ matrix, meaning that the optimal $\boldsymbol{x}$ is (in general) improper.

C. Optimality of Proper Signaling in Gaussian MIMO Broadcast Channels with Shaping Constraints

In multiple-input multiple-output (MIMO) broadcast channels with proper Gaussian noise, the sum rate capacity under a sum power constraint is achievable using so-called dirty paper coding (DPC) with proper Gaussian signals [55]. However, until recently, it had not been shown that the optimality of proper signals also holds under a shaping constraint, i.e., a constraint on the sum transmit covariance matrix instead of on the sum power.[7] In our recent works [34], [35], we used the minimax duality with linear conic constraints from [56], [57] in combination with a power shaping matrix and an impropriety matrix to show this more general result. As discussed in Section III-D, these matrices fit into the framework proposed in this paper. Thus, this recently applied proof technique can be considered as an application of the framework of block-skew-circulant matrices and is worth being mentioned here. In order to briefly sketch the key idea of this approach, we reproduce the proof for the special case of a sum rate maximization in a system with two users. The more general weighted sum rate maximization for an arbitrary number of users is studied in [35].

[7]From [33], it can be concluded that Gaussian signals achieve the complete capacity region, but not whether proper or improper signals are needed.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TSP.2015.2395992

We use $\boldsymbol{x}_k$ and $\boldsymbol{\xi}_k$, $k \in \{1, 2\}$ for the input signals in the downlink and in the dual uplink, respectively. Moreover, we use $\boldsymbol{\eta}_k$ for the downlink noise, and $\boldsymbol{\eta}$ for the uplink noise. The number of downlink transmit antennas is denoted by M, and the number of antennas at the kth downlink receiver is N_k.

We write the sum rate maximization in the downlink as a minimax problem in the composite real representation.

$$\min_{\substack{(\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_1}\succeq\mathbf{0},\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_2}\succeq\mathbf{0}):\ (\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_1},\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_2})\in\mathcal{Y}^\perp \\ \mathrm{tr}[\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_1}+\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_2}]=N_1+N_2}} \ \max_{\substack{(\boldsymbol{C}_{\check{\boldsymbol{x}}_1}\succeq\mathbf{0},\boldsymbol{C}_{\check{\boldsymbol{x}}_2}\succeq\mathbf{0}),\acute{\boldsymbol{Z}}\in\mathcal{Z} \\ \boldsymbol{C}_{\check{\boldsymbol{x}}_1}+\boldsymbol{C}_{\check{\boldsymbol{x}}_2}\preceq\grave{\boldsymbol{P}}+\acute{\boldsymbol{Z}}}} R_{\mathrm{DL}} \tag{42}$$

where $\mathcal{Z} = \mathcal{N}^M$ and

$$\mathcal{Y}^\perp = \{(\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_1}, \grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_2}) \in \mathbb{S}^{2N_1} \times \mathbb{S}^{2N_2} \mid \grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_k} = \alpha \mathbf{I}_{2N_k}\ \forall k,\ \alpha \in \mathbb{R}\}. \tag{43}$$

A shaping constraint $\sum_{k=1}^K \underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}_k} \preceq \underset{\sim}{\boldsymbol{C}}$ on the complex sum transmit covariance matrix corresponds to a constraint $\sum_{k=1}^K \grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}_k} \preceq \grave{\boldsymbol{P}} = \frac{1}{2}\grave{\boldsymbol{C}}$ on the power shaping matrix, where $\grave{\boldsymbol{C}}$ is the composite real equivalent (23) of $\underset{\sim}{\boldsymbol{C}}$. Consequently, the impropriety component can be chosen arbitrarily. This is modeled by the possibility of adding an arbitrary $\acute{\boldsymbol{Z}} \in \mathcal{N}^M$ from the impropriety space to the right hand side of the composite real shaping constraint $\boldsymbol{C}_{\check{\boldsymbol{x}}_1} + \boldsymbol{C}_{\check{\boldsymbol{x}}_2} \preceq \grave{\boldsymbol{P}} + \acute{\boldsymbol{Z}}$.

The feasible set of the worst-case noise minimization contains $(\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_1}, \grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}_2}) = (\frac{1}{2}\mathbf{I}_{2N_1}, \frac{1}{2}\mathbf{I}_{2N_2})$ as the only element, i.e., the noise statistics are fixed. However, the formulation as a minimax problem enables us to apply the minimax uplink-downlink duality from [56], [57]. This duality was shown in [56] for proper complex signals, but by repeating the derivation from [56] for real-valued systems, we obtain that the above optimization has the same optimal value as the following composite real minimax uplink problem.

$$\min_{\substack{\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}}\succeq\mathbf{0},\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}}\in\mathcal{Z}^\perp \\ \mathrm{tr}[\grave{\boldsymbol{P}}\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}}]=N_1+N_2}} \ \max_{\substack{(\boldsymbol{C}_{\check{\boldsymbol{\xi}}_1}\succeq\mathbf{0},\boldsymbol{C}_{\check{\boldsymbol{\xi}}_2}\succeq\mathbf{0}),(\boldsymbol{Y}_1,\boldsymbol{Y}_2)\in\mathcal{Y} \\ \boldsymbol{C}_{\check{\boldsymbol{\xi}}_k}\preceq\mathbf{I}_{2N_k}+\boldsymbol{Y}_k,\quad k=1,2}} R_{\mathrm{UL}}. \tag{44}$$

The constraints of the maximization look complicated at the first glance, but the shaping constraint in combination with the slack variables $(\boldsymbol{Y}_1, \boldsymbol{Y}_2)$ from the subspace $\mathcal{Y}$ (which is the orthogonal complement of $\mathcal{Y}^\perp$) is in fact equivalent to a sum power constraint [34], [35], [57]. The worst-case noise minimization has the constraint that the composite real noise covariance matrix has to lie in $\mathcal{Z}^\perp$. The key point is now Lemma 9, which tells us that $\mathcal{Z}^\perp = \mathcal{P}^M$, i.e., the uplink noise with $\grave{\boldsymbol{C}}_{\check{\boldsymbol{\eta}}} \in \mathcal{P}^M$ has to be proper.

For this minimax uplink problem, the optimal transmit strategy can easily be shown to consist of proper signals [34, Lemma 4]. Transforming this strategy back to the downlink as described in [56] leads to a proper strategy in the downlink as well.

We can sum up that a key point of the proof is the fact that the power shaping space and the impropriety space are orthogonal complements of each other. The minimax uplink-downlink duality thus tells us that allowing arbitrary impropriety components for the downlink transmit signals corresponds to constraining the uplink noise to have a vanishing impropriety component. Therefore, the proof sketched above can be considered as a prime example of a useful application of the framework proposed in this paper.

As stated above, a more general proof for weighted sum rate maximization in a MIMO broadcast channel with an arbitrary number of users can be found in [35]. Moreover, the same framework makes it possible to also prove that the worst-case downlink noise in a MIMO broadcast channel with shaping constraints is proper [35]. Finally, the combination of the minimax uplink-downlink duality and the framework of block-skew-circulant matrices is also helpful to prove results for other system models. In [34], we applied the same technique as a sub-step to show that proper signals are the optimal Gaussian signals for partial decode-and-forward in Gaussian MIMO relay channels. Without exploiting the properties of power shaping matrices and impropriety matrices, this line of proof would not have been possible.

D. *Quality of Service Feasibility in MIMO Broadcast Channels with Widely Linear Transceivers*

As an example in which improper signaling leads to performance gains, we consider the problem of achieving required minimal instantaneous rates for all users in a multiple-input multiple-output (MIMO) broadcast channel without interference cancellation (such as dirty paper coding, DPC). In such a setting, the question of feasibility arises, i.e., it might happen that certain rate requirements cannot be fulfilled even if arbitrarily high transmit power is spent [32], [58]–[60].

For the case where the transmit and receive filters are restricted to be complex linear filters and the per-user signals are proper Gaussian, it was shown in [59], [60] that rates that fulfill

$$\sum_{k\in\mathcal{K}}(1-2^{-r_k}) < \operatorname{Rank}[\underset{\sim}{\boldsymbol{H}}_{\mathcal{K}}] \quad \forall\mathcal{K}\subseteq\{1,\ldots,K\} \tag{45}$$

can be achieved, where $\underset{\sim}{\boldsymbol{H}}_{\mathcal{K}} \in \mathbb{C}^{M\times\sum_{k\in\mathcal{K}} N_k}$ is a matrix that comprises the transposed channel matrices $\underset{\sim}{\boldsymbol{H}}_k^{\mathrm{T}} \in \mathbb{C}^{M\times N_k}$ of all users $k \in \mathcal{K}$ as a block row [60].

In our recent work [32], we derived the feasibility region for the case where the transmit and receive filters are allowed to be widely linear filters and the per-user transmit signals are allowed to be improper. By applying the results from [59], [60] to the composite real representation of a complex MIMO broadcast channel with widely linear filters, we could derive that the feasibility region is given by

$$\sum_{k\in\mathcal{K}}(1-2^{-2r_k}) < \operatorname{Rank}[\boldsymbol{H}'_{\mathcal{K}}] \quad \forall\mathcal{K}\subseteq\{1,\ldots,K\}. \tag{46}$$

where $\boldsymbol{H}'_{\mathcal{K}} \in \mathbb{R}^{2M\times\sum_{k\in\mathcal{K}} 2N_k}$ comprises the composite real equivalents $\grave{\boldsymbol{H}}_k^{\mathrm{T}} \in \mathbb{R}^{2M\times 2N_k}$ of the transposed channel matrices for $k \in \mathcal{K}$. Applying Lemma 25, which was already stated in our previous work [32] (but without detailed proof), we obtain that $\operatorname{Rank}[\boldsymbol{H}'_{\mathcal{K}}] = 2\operatorname{Rank}[\underset{\sim}{\boldsymbol{H}}_{\mathcal{K}}]$. Making use of this, the QoS feasibility region (46) with improper signaling can be shown to be larger than the QoS feasibility region (45) with proper signaling (see [32]).

E. Algorithm Analysis

Above, we have seen in various examples that an advantage of the composite real representation is that methods which have originally been developed for real-valued signals and systems can be applied to treat improper complex signals. This not only applies to theoretical tools as in the cases discussed above, but also to numerical algorithms. Such algorithms could be applied to design widely linear transceivers for the transmission of improper signaling to outperform conventional proper signaling. We demonstrate this application of the framework by means of the analysis of a simple gradient-projection algorithm and leave the analysis of more complicated algorithms open for future research.

A gradient-projection algorithm to optimize the transmit filters $\boldsymbol{T}_k$ in the dual uplink of a MIMO broadcast channel with linear transceivers was proposed, e.g., in [61], [62]. Transferred to the composite real representation, the gradient of the weighted sum rate is given by

$$\frac{\partial \sum_{k'=1}^{K} w_{k'} R_{k'}}{\partial \boldsymbol{T}_k^*} = \boldsymbol{A}_k \boldsymbol{T}_k \tag{47}$$

where the scalars $w_{k'}$ are constant weighting factors,

$$\boldsymbol{A}_k = \frac{1}{\ln 2}\grave{\boldsymbol{H}}_k \left(\sum_{k'=1}^{K} w_{k'}\boldsymbol{X}^{-1} - \sum_{k'\neq k} w_{k'}\boldsymbol{X}_{k'}^{-1}\right)\grave{\boldsymbol{H}}_k^{\mathrm{T}} \tag{48}$$

$$\boldsymbol{X} = \mathbf{I}_{2M} + \sum_{k=1}^{K} \grave{\boldsymbol{H}}_k^{\mathrm{T}} \boldsymbol{T}_k \boldsymbol{T}_k^{\mathrm{T}} \grave{\boldsymbol{H}}_k \tag{49}$$

and $\boldsymbol{X}_{k'} = \boldsymbol{X} - \grave{\boldsymbol{H}}_{k'}^{\mathrm{T}} \boldsymbol{T}_{k'} \boldsymbol{T}_{k'}^{\mathrm{T}} \grave{\boldsymbol{H}}_{k'}$. The gradient-projection update is performed by setting each beamforming matrix $\boldsymbol{T}_k$ to a linear combination of the respective old beamforming matrix and the gradient, i.e.,

$$\boldsymbol{T}_k \leftarrow (a_k \mathbf{I}_{2N_k} + b_k \boldsymbol{A}_k)\boldsymbol{T}_k \tag{50}$$

where the scaling factors a_k and b_k are chosen according to a step size rule and subject to a sum power constraint [61], [62]. The matrices $\grave{\boldsymbol{H}}_k \in \mathcal{BSC}_2^{N_k \times M}$ are the composite real channel matrices. Now, assume that all filter matrices $\boldsymbol{T}_k$ are $\mathcal{BSC}_2$ matrices. The operations contained in the update rule are transposition, addition, multiplication, and matrix inversion. According to Lemmas 1, 3, and 8, all these operations preserve the $\mathcal{BSC}_2$ structure, so that the new filter matrices after the update are again block-skew-circulant.

Consequently, when initialized with $\mathcal{BSC}_2$ matrices, the gradient-projection algorithm converges to a solution with $\mathcal{BSC}_2$ structure. On the other hand, when initialized with matrices that do not have the $\mathcal{BSC}_2$ structure, a solution that does not have this structure either can be obtained.

VIII. Summary and Outlook

Based on block-skew-circulant matrices and block-Hankel-skew-circulant matrices, we have proposed a new framework to characterize widely linear filters and improper signals. After providing a wide-ranging collection of formulae for these kinds of matrices, we have presented a variety of application examples from the field of signal processing for communications. We hope that other researchers will decide to adopt the proposed framework and will come up with various useful applications also from other fields of complex-valued signal processing.

Moreover, there is still potential to extend the framework by proving further properties and deriving additional calculation rules. For instance, future research should study the composite real counterparts of further concepts that were originally introduced in the literature on augmented complex representations. Possible examples are the circularity coefficients and the canonical coordinates studied in [4], but also the widely linear principal component analysis (see [4]).

Another line of possible research is to focus on problems that involve scalar complex random variables (such as in the sum rate maximization for single-antenna interference channels in [29]), which become matrix-valued in the composite real representation. Since scalar expressions are generally more tractable than matrix-valued expressions, the question arises whether the fact that the underlying complex description is scalar leads to desirable additional structural properties of the composite real representation that can be exploited.

APPENDIX A
VARIOUS PROOFS

Proof of Lemma 11: As $\grave{\boldsymbol{A}}$ is symmetric, each eigenvalue $\phi \in \mathbb{R}$ is real-valued, and the corresponding eigenvector $\boldsymbol{q} = \begin{bmatrix}\boldsymbol{u}^{\mathrm{T}} & \boldsymbol{v}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{2M}$ can be chosen to be real-valued. Using the orthogonal $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{J}}_M$ from Lemma 10, we have

$$\begin{aligned}\grave{\boldsymbol{A}}\boldsymbol{q} = \boldsymbol{q}\phi \;\Leftrightarrow\; & \grave{\boldsymbol{J}}_M\grave{\boldsymbol{A}}\grave{\boldsymbol{J}}_M^{\mathrm{T}}\grave{\boldsymbol{J}}_M\boldsymbol{q} = \grave{\boldsymbol{J}}_M\boldsymbol{q}\phi \\ \Leftrightarrow\; & \grave{\boldsymbol{A}}\grave{\boldsymbol{J}}_M\boldsymbol{q} = \grave{\boldsymbol{J}}_M\boldsymbol{q}\phi \end{aligned} \tag{51}$$

and $\boldsymbol{q}^{\mathrm{T}}\grave{\boldsymbol{J}}_M\boldsymbol{q} = 0$. This means that for each eigenvalue ϕ, we have a pair of orthogonal eigenvectors $\boldsymbol{q}$ and $\grave{\boldsymbol{J}}_M\boldsymbol{q}$. Properly arranging these eigenvalues and eigenvectors in $\mathcal{BSC}_2$ matrices, we obtain

$$\grave{\boldsymbol{A}}\underbrace{\begin{bmatrix}\boldsymbol{U} & -\boldsymbol{V}\\ \boldsymbol{V} & \boldsymbol{U}\end{bmatrix}}_{\grave{\boldsymbol{Q}}\in\mathcal{BSC}_2^{M\times M}} = \underbrace{\begin{bmatrix}\boldsymbol{U} & -\boldsymbol{V}\\ \boldsymbol{V} & \boldsymbol{U}\end{bmatrix}}_{\grave{\boldsymbol{Q}}\in\mathcal{BSC}_2^{M\times M}}\underbrace{\begin{bmatrix}\boldsymbol{\Phi} & \\ & \boldsymbol{\Phi}\end{bmatrix}}_{\grave{\boldsymbol{\Lambda}}\in\mathcal{P}^M} \tag{52}$$

where $\grave{\boldsymbol{Q}}^{-1} = \grave{\boldsymbol{Q}}^{\mathrm{T}}$ since the eigenvectors of a symmetric matrix form an orthogonal matrix. ∎

Proof of Lemma 15: For $K \leq L$, we apply Lemma 11 to $\grave{\boldsymbol{U}}\grave{\boldsymbol{\Sigma}}^2\grave{\boldsymbol{U}}^{\mathrm{T}} = \grave{\boldsymbol{A}}\grave{\boldsymbol{A}}^{\mathrm{T}} \in \mathcal{P}^K$ in order to obtain an orthogonal matrix $\grave{\boldsymbol{U}} \in \mathcal{BSC}_2^{K\times K}$ and a diagonal matrix $\grave{\boldsymbol{\Sigma}} \in \mathcal{P}^K$. Then, a reduced SVD of $\grave{\boldsymbol{A}}$ is obtained by solving $\grave{\boldsymbol{A}} = \grave{\boldsymbol{U}}\grave{\boldsymbol{\Sigma}}\grave{\boldsymbol{V}}^{\mathrm{T}}$ for $\grave{\boldsymbol{V}}^{\mathrm{T}}$ [63, Ch. 14]. Let us first assume that $\grave{\boldsymbol{A}}$ has full rank, i.e., all diagonal entries of $\grave{\boldsymbol{\Sigma}}$ are nonzero. Then, since $\grave{\boldsymbol{A}}$, $\grave{\boldsymbol{U}}$, and $\grave{\boldsymbol{\Sigma}}$ are $\mathcal{BSC}_2$ Lemma 7 implies that $\grave{\boldsymbol{V}} \in \mathcal{BSC}_2^{L\times K}$. If we now consider the rank-deficient case by setting some diagonal entries of $\grave{\boldsymbol{\Sigma}}$ to zero, we can keep the corresponding singular vectors unchanged, i.e., we can keep $\grave{\boldsymbol{V}} \in \mathcal{BSC}_2^{L\times K}$. For $K \geq L$, we start with $\grave{\boldsymbol{V}}\grave{\boldsymbol{\Sigma}}^2\grave{\boldsymbol{V}}^{\mathrm{T}} = \grave{\boldsymbol{A}}^{\mathrm{T}}\grave{\boldsymbol{A}} \in \mathcal{P}^L$ and solve for $\grave{\boldsymbol{U}}$ afterwards. ∎

Proof of Lemma 16: For $K \leq L$: since $\acute{\boldsymbol{B}}\acute{\boldsymbol{B}}^{\mathrm{T}} \in \mathcal{P}^K$ due to Lemma 3, the same steps as in the proof of Lemma 15 can be applied. The only difference is that for $\grave{\boldsymbol{U}}\grave{\boldsymbol{\Sigma}}\acute{\boldsymbol{V}}^{\mathrm{T}} = \acute{\boldsymbol{B}} \in \mathcal{BHSC}_2^{K\times L}$

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TSP.2015.2395992

with $\mathcal{BSC}_2$ matrices $\grave{\boldsymbol{U}}$ and $\grave{\boldsymbol{\Sigma}}$, Lemma 7 implies that $\acute{\boldsymbol{V}}$ is $\mathcal{BHSC}_2$. The proof for $K \geq L$ is analogous. ■

Proof of Lemma 20: A $\mathcal{BSC}_2$ matrix $\grave{\boldsymbol{A}}$ can be written as[8]

$$\grave{\boldsymbol{A}} = \mathbf{I}_2 \otimes \boldsymbol{A}_1 + \grave{\boldsymbol{J}}_1 \otimes \boldsymbol{A}_2 \quad \text{with} \quad \grave{\boldsymbol{J}}_1 = \begin{bmatrix} & -1 \\ 1 & \end{bmatrix}. \tag{53}$$

The matrices $\mathbf{I}_2$ and $\grave{\boldsymbol{J}}_1$ are jointly diagonalized by

$$\underset{\sim}{\boldsymbol{U}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -\mathrm{j} & \mathrm{j} \end{bmatrix} \tag{54}$$

and we have $\mathbf{I}_2 = \underset{\sim}{\boldsymbol{U}} \operatorname{diag}(1,1) \underset{\sim}{\boldsymbol{U}}^{\mathrm{H}}$ and $\grave{\boldsymbol{J}}_1 = \boldsymbol{U} \operatorname{diag}(\mathrm{j}, -\mathrm{j}) \boldsymbol{U}^{\mathrm{H}}$. Therefore, it follows from [49, Th. 1] that every eigenvalue of

$$\underset{\sim}{\boldsymbol{M}}_1 = 1 \otimes \boldsymbol{A}_1 + \quad \mathrm{j} \otimes \boldsymbol{A}_2 = \boldsymbol{A}_1 + \mathrm{j}\boldsymbol{A}_2 \tag{55}$$

$$\text{or} \quad \underset{\sim}{\boldsymbol{M}}_2 = 1 \otimes \boldsymbol{A}_1 + (-\mathrm{j}) \otimes \boldsymbol{A}_2 = \boldsymbol{A}_1 - \mathrm{j}\boldsymbol{A}_2 \tag{56}$$

is an eigenvalue of $\grave{\boldsymbol{A}}$, and, conversely, every eigenvalue of $\grave{\boldsymbol{A}}$ is an eigenvalue of $\underset{\sim}{\boldsymbol{M}}_1$ or $\underset{\sim}{\boldsymbol{M}}_2$. The corresponding eigenvectors are given by

$$\underset{\sim}{\boldsymbol{q}}_1 = \begin{bmatrix} 1 \\ -\mathrm{j} \end{bmatrix} \otimes \underset{\sim}{\boldsymbol{x}}_1 \qquad \text{and} \qquad \underset{\sim}{\boldsymbol{q}}_2 = \begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix} \otimes \underset{\sim}{\boldsymbol{x}}_2 \tag{57}$$

due to [49, Corollary 2], where $\underset{\sim}{\boldsymbol{x}}_1$ and $\underset{\sim}{\boldsymbol{x}}_2 = \underset{\sim}{\boldsymbol{x}}_1^*$ are eigenvectors of $\underset{\sim}{\boldsymbol{M}}_1$ and $\underset{\sim}{\boldsymbol{M}}_2 = \underset{\sim}{\boldsymbol{M}}_1^*$ respectively. ■

Alternative Proof of Corollary 3: Since $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} \succeq \mathbf{0} \Leftrightarrow \grave{\boldsymbol{P}} - \acute{\boldsymbol{N}} \succeq \mathbf{0}$ (see Proof of Theorem 4), we have that

$$0 \leq \boldsymbol{x}^{\mathrm{T}}(\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}})\boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}}(\grave{\boldsymbol{P}} - \acute{\boldsymbol{N}})\boldsymbol{x} = 2\boldsymbol{x}^{\mathrm{T}}\grave{\boldsymbol{P}}\boldsymbol{x}. \tag{58}$$

This shows that $\boldsymbol{x}^{\mathrm{T}}\grave{\boldsymbol{P}}\boldsymbol{x} \geq 0$, $\forall \boldsymbol{x} \in \mathbb{R}^{2M}$. Now let $\boldsymbol{x} \in \operatorname{null}[\grave{\boldsymbol{P}}]$. Then, $0 \leq \boldsymbol{x}^{\mathrm{T}}\acute{\boldsymbol{N}}\boldsymbol{x}$. Moreover, $\grave{\boldsymbol{J}}_M\boldsymbol{x} \in \operatorname{null}[\grave{\boldsymbol{P}}]$ and $0 \leq \boldsymbol{x}^{\mathrm{T}}\grave{\boldsymbol{J}}_M^{\mathrm{T}}\acute{\boldsymbol{N}}\grave{\boldsymbol{J}}_M\boldsymbol{x} = -\boldsymbol{x}^{\mathrm{T}}\acute{\boldsymbol{N}}\boldsymbol{x}$ due to Lemma 10. This implies that $\boldsymbol{x}^{\mathrm{T}}\acute{\boldsymbol{N}}\boldsymbol{x} = 0$. Now suppose that there exists a $\boldsymbol{y}$ such that $\boldsymbol{y}^{\mathrm{T}}\acute{\boldsymbol{N}}\boldsymbol{x} \neq 0$. Then,

$$(\alpha\boldsymbol{x} + \boldsymbol{y})^{\mathrm{T}}(\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}})(\alpha\boldsymbol{x} + \boldsymbol{y}) = \underbrace{\boldsymbol{y}^{\mathrm{T}}(\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}})\boldsymbol{y}}_{=const. \geq 0} + 2\alpha\boldsymbol{y}^{\mathrm{T}}\acute{\boldsymbol{N}}\boldsymbol{x} \tag{59}$$

[8] Note that [49] uses a different definition of the $\otimes$ operator where the order of the operands is reversed.

and there exists an $\alpha \in \mathbb{R}$ such that the sum on the right hand side becomes negative, which contradicts $\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} \succeq \mathbf{0}$. Thus, $\boldsymbol{y}^{\mathrm{T}}\acute{\boldsymbol{N}}\boldsymbol{x} = 0$ for all $\boldsymbol{y}$, which implies $\boldsymbol{x} \in \mathrm{null}[\acute{\boldsymbol{N}}]$. Items 3) and 4) are necessary since they are implications of 2). ■

Proof of Theorem 6: Using Lemma 21, and the relation between $\grave{\boldsymbol{P}}_{\underset{\sim}{\boldsymbol{x}}}$ and $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$ in (32), we have that the first summand in (38) is equal to $\log\det(\pi \mathrm{e}\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}})$, which is the differential entropy of a proper complex Gaussian random vector with covariance matrix $\underset{\sim}{\boldsymbol{C}}_{\underset{\sim}{\boldsymbol{x}}}$. We have to show that the determinant $d = \det\left(\mathbf{I}_{2M} + \grave{\boldsymbol{P}}^{-1}\acute{\boldsymbol{N}}\right)$ takes a value between zero and one. Since $\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}^{\mathrm{H}} \succeq \boldsymbol{A}\boldsymbol{X}'\boldsymbol{A}^{\mathrm{H}}$ for $\boldsymbol{X} \succeq \boldsymbol{X}'$ [48, Section 7.7],

$$\grave{\boldsymbol{P}} + \acute{\boldsymbol{N}} \succeq \mathbf{0} \quad \Leftrightarrow \quad \mathbf{I}_{2M} + \grave{\boldsymbol{P}}^{-\frac{1}{2}}\acute{\boldsymbol{N}}\grave{\boldsymbol{P}}^{-\frac{1}{2}} \succeq \mathbf{0} \tag{60}$$

and, thus,

$$d = \det\left(\mathbf{I}_{2M} + \grave{\boldsymbol{P}}^{-\frac{1}{2}}\acute{\boldsymbol{N}}\grave{\boldsymbol{P}}^{-\frac{1}{2}}\right) \geq 0. \tag{61}$$

Since $\grave{\boldsymbol{P}}^{-\frac{1}{2}}$ is $\mathcal{BSC}_2$ due to Lemmas 12 and 8, $\grave{\boldsymbol{P}}^{-\frac{1}{2}}\acute{\boldsymbol{N}}\grave{\boldsymbol{P}}^{-\frac{1}{2}}$ is $\mathcal{BHSC}_2$ due to Lemma 3. Using the standard EVD of this symmetric $\mathcal{BHSC}_2$ matrix (see Lemma 13), we obtain

$$\begin{aligned} d &= \det\left(\mathbf{I}_{2M} + \grave{\boldsymbol{Q}}\acute{\boldsymbol{\Lambda}}\grave{\boldsymbol{Q}}^{\mathrm{T}}\right) = \det\left(\mathbf{I}_{2M} + \acute{\boldsymbol{\Lambda}}\right) = \prod_{i=1}^{2M}(1+\lambda_i) \\ &= \prod_{i=1}^{M}(1+\lambda_i)(1-\lambda_i) = \prod_{i=1}^{M}\underbrace{(1-\lambda_i^2)}_{\leq 1} \leq 1 \end{aligned} \tag{62}$$

where we have used $\lambda_{i+M} = -\lambda_i$ for $i = 1, \ldots, M$. This shows that $\frac{1}{2}\log d \leq 0$ with equality if and only if $\lambda_i = 0\ \forall i$, i.e., if and only if $\acute{\boldsymbol{N}} = \mathbf{0}$. ■

References

[1] R. Meyer, R. Schober, and W. Gerstacker, "Method for cancelling interference during TDMA transmission and/or FDMA transmission," International Patent WO 02/054 660 A1, 2002.

[2] J. Hofmann, E. Riddington, V. Rexhepi-van der Pol, S. Parolari, G. Sébire, and M. Säily, "3GPP Release 9 and beyond," in *GSM/EDGE: Evolution and Performance*, M. Säily, G. Sébire, and E. Riddington, Eds. John Wiley & Sons, Ltd., 2011, ch. 4.

[3] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling—settling the Høst-Madsen–Nosratinia conjecture," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4552–4565, Sep. 2010.

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

[5] 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.

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

[7] 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.

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

[9] D. P. Mandic and S. L. Goh, *Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models*. West Sussex, UK: John Wiley & Sons, 2009.

[10] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[11] C. Hellings and W. Utschick, "Performance gains due to improper signals in MIMO broadcast channels with widely linear transceivers," in *Proc. Int. Conf. Acoust., Speech, and Signal Process. (ICASSP) 2013*, May 2013, pp. 4379–4383.

[12] A. Hakkarainen, J. Werner, M. Renfors, K. Dandekar, and M. Valkama, "RF-aware widely linear MMSE beamforming," in *Proc. 10th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2013, pp. 582–586.

[13] Y. C. Yoon and H. Leib, "Maximizing SNR in improper complex noise and applications to CDMA," *IEEE Commun. Lett.*, vol. 1, no. 1, pp. 5–8, Jan. 1997.

[14] H. Gerstacker, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1512–1523, Sep. 2003.

[15] D. Darsena, G. Gelli, L. Paura, and F. Verde, "Widely linear equalization and blind channel identification for interference-contaminated multicarrier systems," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 1163–1177, Mar. 2005.

[16] D. Mattera, L. Paura, and F. Sterle, "Widely linear MMSE transceiver for real-valued sequences over MIMO channel," presented at the 14th European Signal Process. Conf. (EUSIPCO), Florence, Italy, Sep. 4–8, 2006.

[17] 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, Jun. 2006.

[18] A. Mirbagheri, K. N. Plataniotis, and S. Pasupathy, "An enhanced widely linear CDMA receiver with OQPSK modulation," *IEEE Trans. Commun.*, vol. 54, no. 2, pp. 261–272, Feb. 2006.

[19] A. S. Cacciapuoti, G. Gelli, and F. Verde, "FIR zero-forcing multiuser detection and code designs for downlink MC-CDMA," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4737–4751, Oct. 2007.

[20] D. Darsena, G. Gelli, and F. Verde, "Universal linear precoding for NBI-proof widely linear equalization in MC systems," *EURASIP J. Wireless Commun. and Networking*, vol. 2008, pp. 8:1–8:11, Jan. 2008.

[21] P. Xiao and M. Sellathurai, "Iterative receiver design for MIMO systems with improper signal constellations," presented at the Int. Conf. Commun. (ICC) 2009, Dresden, Germany, Jun. 14–18, 2009.

[22] ——, "Improved linear transmit processing for single-user and multi-user MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1768–1779, Mar. 2010.

[23] J. Steinwandt, R. C. de Lamare, L. Wang, N. Song, and M. Haardt, "Widely linear adaptive beamforming algorithm based on the conjugate gradient method," presented at the Int. ITG Workshop on Smart Antennas (WSA) 2011, Aachen, Germany, Feb. 24–25, 2011.

[24] Z. Ho, A. Zappone, and E. Jorswieck, "Optimal non-regenerative relay processing with improper signals," in *Proc. 10th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2013, pp. 567–571.

[25] J. Zhang and M. Haardt, "Widely linear signal processing for two-way relaying with MIMO amplify and forward relays," in *Proc. 10th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2013, pp. 264–268.

[26] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[27] C. Lameiro and I. Santamaría, "Degrees-of-freedom for the 4-user SISO interference channel with improper signaling," in *Proc. Int. Conf. Commun. (ICC) 2013*, Jun. 2013, pp. 1646–1650.

[28] 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, Sep. 2012.

[29] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2899–2913, Jun. 2013.

[30] Y. Zeng, R. Zhang, E. Gunawan, and Y. L. Guan, "MISO interference channel with improper Gaussian signaling," in *Proc. Int. Conf. Acoust., Speech, and Signal Process. (ICASSP) 2013*, May 2013, pp. 4374–4378.

[31] ——, "Optimized transmission with improper Gaussian signaling in the K-user MISO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6303–6313, Dec. 2013.

[32] C. Hellings, M. Joham, and W. Utschick, "QoS feasibility in MIMO broadcast channels with widely linear transceivers," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1134–1137, Nov. 2013.

[33] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[34] C. Hellings, L. Gerdes, L. Weiland, and W. Utschick, "On optimal Gaussian signaling in MIMO relay channels with partial decode-and-forward," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3153–3164, Jun. 2014.

[35] C. Hellings, L. Weiland, and W. Utschick, "Optimality of proper signaling in Gaussian MIMO broadcast channels with shaping constraints," in *Proc. Int. Conf. Acoust., Speech, and Signal Process. (ICASSP) 2014*, May 2014, pp. 3498–3502.

[36] P. J. Schreier, L. L. Scharf, and A. Hanssen, "A generalized likelihood ratio test for impropriety of complex signals," *IEEE Signal Process. Lett.*, vol. 13, no. 7, pp. 433–436, Jul. 2006.

[37] D. Darsena, G. Gelli, and F. Verde, "Widely-linear precoders and decoders for MIMO channels," in *Proc. 10th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2013, pp. 587–591.

[38] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[39] H. Karner, J. Schneid, and C. W. Ueberhuber, "Spectral decomposition of real circulant matrices," *Linear Algebra and its Appl.*, vol. 367, pp. 301–311, Jul. 2003.

[40] N. Akhondi and F. Toutounian, "Accelerated circulant and skew circulant splitting methods for Hermitian positive definite Toeplitz systems," *Adv. Numerical Anal.*, vol. 2012, Article ID 973407, pp. 1–17, 2012.

[41] Y. Gao, Z. Jiang, and Y. Gong, "On the determinants and inverses of skew circulant and skew left circulant matrices with Fibonacci and Lucas numbers," *WSEAS Trans. Math.*, vol. 12, no. 4, pp. 472–481, Apr. 2013.

[42] G. J. Tee, "Eigenvectors of block circulant and alternating circulant matrices," *New Zealand J. Math.*, vol. 36, pp. 195–211, 2007.

[43] L. Ducas and P. Q. Nguyen, "Faster Gaussian lattice sampling using lazy floating-point arithmetic," in *Advances in Cryptology – ASIACRYPT 2012*, ser. Lecture Notes in Computer Science. Springer, 2012, vol. 7658, pp. 415–432.

[44] T. M. Foltz, "Trigonometric transforms for image reconstruction," Ph.D. Dissertation, Air Force Institute of Technology, Wright-Patterson AFB, OH, USA, 1998.

[45] X.-G. Lv and T.-Z. Huang, "The inverses of block Toeplitz matrices," *J. Math.*, vol. 2013, Article ID 207176, pp. 1–8, 2013.

[46] S. T. Flammia, N. C. Menicucci, and O. Pfister, "The optical frequency comb as a one-way quantum computer," *J. Physics B: Atomic, Molecular and Optical Physics*, vol. 42, no. 114009, pp. 1–12, May 2009.

[47] K. D. Ikramov and V. N. Chugunov, "A characterization of the Toeplitz and Hankel circulants," *J. Math. Sciences*, vol. 176, no. 1, pp. 38–43, Jul. 2011.

[48] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.

[49] B. Friedman, "Eigenvalues of composite matrices," *Math. Proc. Cambridge Philosophical Soc.*, vol. 57, no. 1, pp. 37–49, Jan. 1961.

[50] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore and London: The Johns Hopkins University Press, 1996.

[51] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: identifiability, uniqueness, and separability," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.

[52] P. Wahlberg and P. J. Schreier, "Spectral relations for multidimensional complex improper stationary and (almost) cyclostationary processes," *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1670–1682, Apr. 2008.

[53] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2009, 7th printing with corrections.

[54] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.

[55] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[56] A. Dotzler, M. Riemensberger, and W. Utschick, "Minimax duality for MIMO interference networks," 2014, to be published. Preprint available at http://www.msv.ei.tum.de/ando/MiniMaxDuality.pdf.

[57] A. Dotzler, M. Riemensberger, W. Utschick, and G. Dietl, "Interference robustness for cellular MIMO networks," in *Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2012, pp. 229–233.

[58] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[59] R. Hunger and M. Joham, "A complete description of the QoS feasibility region in the vector broadcast channel," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3870–3878, Jul. 2010.

[60] M. Joham and R. Hunger, "Feasible rate region of the MIMO broadcast channel with linear transceivers," in *Proc. Int. ITG Workshop on Smart Antennas (WSA) 2010*, Feb. 2010, pp. 342–349.

[61] R. Böhnke and K.-D. Kammeyer, "Weighted sum rate maximization for MIMO-OFDM systems with linear and dirty paper precoding," presented at the 7th Int. ITG Conf. Source and Channel Coding (SCC), Ulm, Germany, Jan. 14–16, 2008.

[62] R. Hunger, D. Schmidt, and M. Joham, "A combinatorial approach to maximizing the sum rate in the MIMO BC with linear precoding," in *Conf. Record 42nd Asilomar Conf. Signals, Syst., and Comput. (ACSSC)*, Oct. 2008, pp. 316–320.

[63] B. Dasgupta, *Applied Mathematical Methods*. Dorling Kindersley, 2006.

Christoph Hellings (S'10) received the B.Sc. degree and the Dipl.-Ing. degree in electrical engineering (both with high distinction) from the Technische Universität München (TUM) in 2008 and 2010, respectively.

He is currently working towards the doctoral degree at the Associate Institute for Signal Processing, TUM. His research interests include energy efficient communications, improper signaling, and the optimization of linear precoding strategies in multi-carrier systems.

Mr. Hellings was awarded a scholarship of the Max Weber Program of the Bavarian state, and for his Diploma thesis, he received an award of the German Association for Electrical, Electronic & Information Technologies (VDE). In 2011, he received an outstanding teaching assistant award from the student representatives of the Department of Electrical Engineering and Information Technology, TUM.

Wolfgang Utschick (SM'06) completed several accredited industrial training programs before he received the diploma (93) and doctoral degrees (98) in electrical engineering, both with honors, from Technische Universitt Mnchen (TUM). Since 2002 Dr. Utschick has been appointed Professor at TUM where he is director of the Fachgebiet Methoden der Signalverarbeitung (Signal Processing Group). Dr. Utschick teaches courses on Signal Processing, Stochastic Processes, and Optimization Theory in the field of Wireless Communications and Signal Processing applications. Since 2011 he is serving as a regular guest professor at Singapore's new autonomous university, Singapore Institute of Technology (SIT). He holds some 20 patents in the field of multiantenna signal processing and has authored and co-authored more than 300 technical articles in international journals and conference proceedings. He edited several books and is founder and editor of the Springer book series Foundations in Signal Processing, Communications and Networking. Dr. Utschick has been Principal Investigator in multiple research projects funded by the German Research Fund (DFG). He is currently the coordinator and spokesman of the German DFG focus program Communications over Interference limited Networks (COIN) which is devoted to topics as cooperative communications, crosslayer design, ad-hoc wireless networks, etc. He is a member of the VDE and senior member of the IEEE, where he serves as an elected member for the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications and Networking since 2010 during his second term. Since 2011 he is member of the steering committee of the Department for Electrical and Computer Engineering at TUM.