

# Maximally Improper Signaling in Underlay MIMO Cognitive Radio Networks

Christian Lameiro\*, *Member, IEEE*, Ignacio Santamaría†, *Senior Member, IEEE*, Peter J. Schreier\*, *Senior Member, IEEE*, and Wolfgang Utschick‡, *Senior Member, IEEE*

**Abstract**—Improper Gaussian signaling is a well-known technique that has been shown to improve performance in different multi-user scenarios. In this paper, we analyze the benefit of improper signaling in underlay cognitive radio when users are equipped with multiple antennas. Specifically, we assume that the primary user is protected by the so-called interference temperature constraint, which guarantees a prescribed rate requirement. In this setting, we study how the maximum tolerable interference power changes when the interference is additionally constrained to be maximally improper (strictly noncircular, or rectilinear). We observe that the correlation structure of a maximally improper interference is an additional degree of freedom that can be exploited to improve the SU performance. Because of that, we propose two different protection strategies for the PU where this structure is either constrained or unconstrained, and derive the interference temperature threshold for both cases. We then focus on the secondary user and provide designs of the transmission parameters under the proposed protection strategies.

**Index Terms**—Improper Gaussian signaling, interference temperature, underlay cognitive radio, MIMO systems, majorization theory, transceiver optimization.

## I. INTRODUCTION

An improper complex random vector is correlated with its complex conjugate [1], as opposed to a proper one. Improper complex random vectors can be used to model real-world communication signals in different contexts. For example, some digital constellations, such as binary shift keying (BPSK) or Gaussian minimum shift keying (GMSK), yield an improper transmit signal. Also, hardware imperfections can be modeled as additive improper noise [2]. The correlation structure of an improper signal can be exploited at the receiver by widely-linear processing [1], [3], [4], which is linear in both the signal and its complex conjugate. The design of widely-linear receivers has been widely studied in the literature [5]–[8].

From an information-theoretic point of view, a proper Gaussian random vector is known to maximize the entropy for a given covariance matrix, which is known as the maximum entropy theorem [9]. Because of that, the proper Gaussian distribution achieves capacity in point-to-point, multiple-access

and broadcast channels [10]. This is also the reason why this distribution is a common assumption in the theoretical analysis of wireless communication scenarios, even though the capacity-achieving distribution is, in general, unknown except for some special cases.

The use of improper signals, rather than proper ones, has recently emerged as an efficient means of improving the performance in multiuser wireless communication networks. The potential superiority of improper Gaussian signaling (IGS) has been shown in different interference-limited scenarios, such as the interference channel [11]–[16], relay channels [17], [18], and cognitive radio networks [19]–[24], although recent results indicate that proper Gaussian signaling (PGS) can be sufficient if appropriate time sharing protocols are considered [25], [26]. The latter will generally not be applicable for cognitive radio networks, since the required coordinated scheduling for a time-sharing protocol is probably not feasible in such scenarios.

On another front, cognitive radio (CR) [27], where licensed or primary users (PU) share the spectrum with unlicensed or secondary users (SU), has been proposed as a promising technology to improve the spectrum utilization in 5G wireless communication systems [28]. Among the different CR paradigms (see [27] for an overview), underlay cognitive radio (UCR) exploits the fact that the PUs are typically not fully loaded, i.e., they operate below capacity, and thus tolerate a certain amount of interference. Therefore, SUs can access the channel provided that the interference level at the PU is tolerable. This is typically accomplished by constraining the interference power below a given threshold [29]–[31], the so-called interference temperature (IT) limit. Because a UCR scenario is interference-limited (as the performance of the SU is limited by the interference they cause to the primary receiver), IGS has also been shown to pay off in this context [19]–[22]. However, all existing works consider single-antenna transceivers.

When users have multiple antennas and the PU is protected by an interference power constraint, the interference power threshold can be set such that the PU is ensured a prescribed data rate. As the total interference power does not fully capture the spatial structure of the interference in the multiple-antenna case, this threshold has to be determined assuming the worst-case interference covariance matrix in order to guarantee the instantaneous rate of the PU [32]. How this threshold changes when the interference is improper is not straightforward, as it depends on its degree of impropriety or non-circularity as well as on its spatial correlation structure. In this paper we aim at filling this gap by studying how conventional total interference

\*C. Lameiro and P. J. Schreier are with the Signal & System Theory Group, Universität Paderborn, Germany (email: {christian.lameiro, peter.schreier}@sst.upb.de).

†I. Santamaría is with the Department of Communications Engineering, University of Cantabria, Spain (e-mail: i.santamaria@unican.es).

‡W. Utschick is with the Associate Institute for Signal Processing, Technische Universität München, Germany (e-mail: utschick@tum.de).

The work of C. Lameiro and P. J. Schreier was supported by the German Research Foundation (DFG) under grants SCHR 1384/6-1 and LA 4107/1-1. The work of I. Santamaría was supported by the Ministerio de Economía y Competitividad (MINECO) and AEI/FEDER funds of the UE, Spain, under project CARMEN (TEC2016-75067-C4-4-R).

power constraints can be combined with additional constraints that force the interference to be improper. The motivation is that these additional constraints further limit the structure of the interference, thus alleviating the impact of the worst-case interference and in turn permitting a higher interference power threshold. Similar ideas have been proposed for PGS by additionally constraining the spatial structure of the interference, but at the cost of increased cooperation between primary and secondary networks [33]. This paper is an extension of our previous work [34]. With respect to that work, we provide here one additional protection strategy, include detailed proofs for all results, and analyze the performance of the SU under the proposed protection strategies.

The rest of the paper is organized as follows. Section II introduces the system model. Two different strategies based on IGS are presented in Section III. In Section IV, we derive the IT thresholds for the two proposed strategies, while Section V addresses the optimization of the SU. Numerical examples are provided in Section VI. Section VII presents the concluding remarks.

### A. Preliminaries

We start with some definitions and properties of improper complex random vectors that will be used throughout the paper. We refer the reader to [1] for a comprehensive treatment of the subject.

The complementary covariance matrix of a complex random vector  $\mathbf{x}$  is defined as  $\tilde{\mathbf{R}}_{xx} = E\{\mathbf{x}\mathbf{x}^T\}$ , where  $E\{\cdot\}$  denotes expectation. If  $\tilde{\mathbf{R}}_{xx} = \mathbf{0}$ , we call  $\mathbf{x}$  proper, otherwise improper. Without loss of generality, the complementary covariance matrix can be expressed as [1, Section 3.2.3]<sup>1</sup>  $\tilde{\mathbf{R}}_{xx} = \mathbf{R}_{xx}^{1/2} \mathbf{F} \mathbf{C} \mathbf{F}^T \mathbf{R}_{xx}^{*/2}$ , where  $\mathbf{R}_{xx} = E\{\mathbf{x}\mathbf{x}^H\}$  is the covariance matrix,  $\mathbf{F}$  is a unitary matrix, which we call improper signature matrix, and  $\mathbf{C}$  is a diagonal matrix containing the circularity coefficients, which measure the degree of impropriety and belong to the range  $[0, 1]$ . If  $\mathbf{C} = \mathbf{I}$ , we call  $\mathbf{x}$  maximally improper (or rectilinear). Finally, it is usually useful to express the second-order statistics of  $\mathbf{x}$  through the augmented covariance matrix, which is defined as

$$\mathbf{R}_{xx} = E\{\underline{\mathbf{x}}\underline{\mathbf{x}}^H\} = \begin{bmatrix} \mathbf{R}_{xx} & \tilde{\mathbf{R}}_{xx} \\ \tilde{\mathbf{R}}_{xx}^* & \mathbf{R}_{xx}^* \end{bmatrix}, \quad (1)$$

where  $\underline{\mathbf{x}} = [\mathbf{x}^T \ \mathbf{x}^H]^T$ .

## II. SYSTEM MODEL

We consider a UCR scenario, where a multiple-input multiple-output (MIMO) PU shares the spectrum with a MIMO SU. For the sake of exposition, we consider an even number of antennas that is the same at both sides of the link,<sup>2</sup> with  $N$  antennas in the PU link and  $M$  in the SU link. The direct links for the PU and SU are  $\mathbf{H}_{pp} \in \mathbb{C}^{N \times N}$  and  $\mathbf{H}_{ss} \in \mathbb{C}^{M \times M}$ , respectively, while  $\mathbf{H}_{ps} \in \mathbb{C}^{N \times M}$  is the cross-channel between the secondary transmitter and the primary

receiver. The received signals at the primary and secondary receiver are respectively given by

$$\mathbf{y}_p = \mathbf{H}_{pp}\mathbf{s}_p + \mathbf{H}_{ps}\mathbf{s}_s + \mathbf{n}_p, \quad (2)$$

$$\mathbf{y}'_s = \mathbf{H}'_{ss}\mathbf{s}_s + \mathbf{H}_{sp}\mathbf{s}_p + \mathbf{n}'_s, \quad (3)$$

where  $\mathbf{n}_p$  ( $\mathbf{n}'_s$ ) is the additive noise at the primary (secondary) receiver, which is assumed to be proper Gaussian with covariance matrix  $\sigma^2\mathbf{I}$ ; and  $\mathbf{s}_p$  ( $\mathbf{s}_s$ ) is the signal transmitted by the primary (secondary) transmitter. To avoid cumbersome expressions, we consider an equivalent model for the received signal at the SU, which eliminates the interfering term by an appropriate scaling of  $\mathbf{H}'_{ss}$ . In particular, by multiplying  $\mathbf{y}'_s$  with the inverse square root of the interference-plus-noise covariance matrix times the noise standard deviation, we obtain the equivalent model for the SU<sup>3</sup>

$$\mathbf{y}_s = \sigma (\mathbf{H}_{sp}\mathbf{Q}_p\mathbf{H}_{sp}^H + \sigma^2\mathbf{I})^{-1/2} \mathbf{y}'_s = \mathbf{H}_{ss}\mathbf{s}_s + \mathbf{n}_s, \quad (4)$$

where  $\mathbf{n}_s$  and  $\mathbf{n}'_s$  follow the same distribution. The primary transmitter is unaware of the secondary system and thus uses PGS with an arbitrary and fixed covariance matrix  $\mathbf{Q}_p \in \mathbb{S}_+^N$ , where  $\mathbb{S}_+^N$  denotes the set of  $N \times N$  positive semidefinite Hermitian matrices. This also means that the PU is unaware of the SU channel coefficients, number of SUs, etc. Additionally, the SU does not have access to  $\mathbf{Q}_p$  and  $\mathbf{H}_{pp}$ , but can acquire the cross-channel  $\mathbf{H}_{ps}$ .

In order to allow the SU to access the channel and at the same time protect the PU from the secondary transmissions, an IT constraint (i.e., interference power constraint) is typically imposed to the secondary network. This way, the performance of the PU can be guaranteed in spite of the secondary transmission without requiring explicit cooperation between secondary and primary systems. The IT constraint imposed on the secondary transmitter is  $\text{Tr}(\mathbf{H}_{ps}\mathbf{Q}_s\mathbf{H}_{ps}^H) \leq t$ , where  $\mathbf{Q}_s \in \mathbb{S}_+^M$  is the transmit covariance matrix of the SU and  $t$  is the IT threshold. The value of  $t$  must be determined such that the instantaneous rate of the PU is ensured [32]. Since the PU is unaware of the secondary network, the IT threshold has to be obtained for the most harmful interference covariance matrix  $\mathbf{K} \in \mathbb{S}_+^N$  at the PU that permits achieving the required rate [32]. That is, let  $\mathbf{K}^*$  denote the most harmful interference covariance matrix and  $t^*$  its associated IT threshold, so that the PU rate is guaranteed. Then

$$\begin{aligned} \bar{R} &= R_p(\mathbf{K}^*) \leq R_p(\mathbf{H}_{ps}\mathbf{Q}_s\mathbf{H}_{ps}^H), \\ \forall \mathbf{Q}_s &\in \{\mathbf{Q} \succeq \mathbf{0} : \text{Tr}(\mathbf{H}_{ps}\mathbf{Q}_s\mathbf{H}_{ps}^H) \leq t^*\}, \end{aligned} \quad (5)$$

where  $\bar{R}$  is the rate constraint, and  $R_p(\mathbf{K})$  is the PU rate when the interference is proper with covariance matrix  $\mathbf{K}$ , which is given by

$$R_p(\mathbf{K}) = \log_2 |\mathbf{I} + (\sigma^2\mathbf{I} + \mathbf{K})^{-1}\mathbf{H}_{pp}\mathbf{Q}_p\mathbf{H}_{pp}^H|. \quad (6)$$

This approach, however, is very conservative and may limit the performance of the secondary system. This is because it is worst-case oriented (the instantaneous PU rate has to be

<sup>1</sup>We use in this paper the unique positive-semidefinite square root for all matrix square roots.

<sup>2</sup>Our results can easily be extended to an arbitrary number of antennas.

<sup>3</sup>Note that this is possible because the transmit covariance matrix of the PU is fixed. Additionally, this matrix is only required at the SU receiver to carry out the optimizations described in Section V, and it can be estimated by listening to the PU signal prior to transmission.

guaranteed with probability one), and because there are many unexploited degrees of freedom in  $\mathbf{K}$  (as only the trace of  $\mathbf{K}$  is constrained). While the former cannot be changed as it ensures the PU performance, accounting for the latter may help obtain more suitable protection strategies. That is, by cleverly imposing additional constraints on the interference covariance matrix, the tolerable interference power  $t$  can increase without compromising the PU performance. This, in turn, may result, despite the additional constraints, in a better performance for the SU. Specifically, we study how an additional constraint on the degree of impropriety of the interference affects the IT limit  $t$  and the SU performance.

### III. PROTECTION STRATEGIES BASED ON IMPROPER SIGNALING

This section presents the proposed protection strategies. Even though the foregoing section has introduced a system model with a single SU, the proposed strategies are valid for arbitrary secondary networks, as no information about the secondary system is used, and only a worst-case interference covariance matrix (which can eventually be caused by one or multiple SUs) is considered.

In our previous work for single-input single-output (SISO) UCR networks [19] we have shown that, whenever IGS permits the SU to achieve higher rate than PGS, maximally IGS is optimal provided that the power budget is large enough. When the power budget is not sufficiently high, the optimal circularity coefficient depends on the particular channel realization of each link. This means that, in order to find suitable constraint values for the circularity coefficients of the interference, i.e., for  $\mathbf{C}$ , the channel matrices involving the secondary transmitter need to be considered by the PU. Since this information is not available at the primary node, as the PU has no knowledge about the secondary network, a fixed choice or constraint for  $\mathbf{C}$  has to be considered instead, which must be suitable for most channel realizations and secondary network configurations. Based on the optimality of maximally IGS in the high-power regime, we will constrain the interference to be maximally improper for the MIMO setup, which will also simplify the analysis. Hence, the interference complementary covariance matrix is constrained to be of the form

$$\tilde{\mathbf{K}} = \mathbf{K}^{1/2} \mathbf{F}_p \mathbf{F}_p^T \mathbf{K}^{*/2}, \quad (7)$$

where  $\mathbf{F}_p \in \mathbb{U}^N$  is the improper signature matrix of the interference at the primary receiver, with  $\mathbb{U}^N$  being the set of  $N \times N$  unitary matrices. This matrix is the key to our analysis as it represents, along with the covariance matrix  $\mathbf{K}$ , the available degrees of freedom for the optimal design of a maximally improper interference. When the interference is maximally improper, the PU rate is given by [1]

$$R_p(\mathbf{K}, \mathbf{F}_p) = \frac{1}{2} \log_2 \left| \mathbf{I} + \begin{bmatrix} \sigma^2 \mathbf{I} + \mathbf{K} & \mathbf{K}^{1/2} \mathbf{F}_p \mathbf{F}_p^T \mathbf{K}^{*/2} \\ \mathbf{K}^{*/2} \mathbf{F}_p^* \mathbf{F}_p^H \mathbf{K}^{1/2} & \sigma^2 \mathbf{I} + \mathbf{K}^* \end{bmatrix}^{-1} \right. \\ \left. \times \begin{bmatrix} \mathbf{H}_{pp} \mathbf{Q}_p \mathbf{H}_{pp}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{pp}^* \mathbf{Q}_p^* \mathbf{H}_{pp}^T \end{bmatrix} \right|. \quad (8)$$

We will then consider the following constraints on the interference:

$$\text{Tr}(\mathbf{K}) \leq t, \text{ and } \tilde{\mathbf{K}} = \mathbf{K}^{1/2} \mathbf{F}_p \mathbf{F}_p^T \mathbf{K}^{*/2}, \mathbf{F}_p \in \mathbb{F}, \quad (9)$$

where  $\mathbb{F} \subseteq \mathbb{U}^N$ . The above set of constraints includes, along with the conventional IT constraint, a maximally improper constraint through a signature matrix  $\mathbf{F}_p$  that must be contained within a predefined set,  $\mathbb{F}$ . Our proposed scheme consists in the following two steps.

- 1) The primary receiver obtains the maximum value of  $t$  that ensures its rate by solving

$$\underset{t}{\text{maximize}} \quad t, \quad (10a)$$

$$\text{subject to} \quad R_p(\mathbf{K}, \mathbf{F}_p) \geq \bar{R}, \forall \mathbf{K} \in \mathbb{K}(t), \forall \mathbf{F}_p \in \mathbb{F}, \quad (10b)$$

where  $\mathbb{K}(t) = \{\mathbf{K} \succeq \mathbf{0} : \text{Tr}(\mathbf{K}) \leq t\}$  is the set of admissible interference covariance matrices, and  $\bar{R}$  is the prescribed rate constraint. Notice that the rate constraint has to be fulfilled for all the matrices  $\mathbf{K}$  and  $\mathbf{F}_p$  that belong to their respective sets, and therefore the only optimization variable is the maximum tolerable interference power  $t$ . The primary receiver broadcasts the value of  $t$  through a control channel, or, alternatively, sends its value to a central unit, so that it can be acquired by the secondary network.

- 2) The SU optimizes its transmission scheme subject to the interference constraints by solving

$$\underset{\mathbf{Q}_s, \tilde{\mathbf{Q}}_s, \mathbf{F}_p}{\text{maximize}} \quad R_s(\mathbf{Q}_s, \tilde{\mathbf{Q}}_s) = \\ \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \begin{bmatrix} \mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H & \mathbf{H}_{ss} \tilde{\mathbf{Q}}_s \mathbf{H}_{ss}^T \\ \mathbf{H}_{ss}^* \tilde{\mathbf{Q}}_s^* \mathbf{H}_{ss}^H & \mathbf{H}_{ss}^* \mathbf{Q}_s^* \mathbf{H}_{ss}^T \end{bmatrix} \right| \quad (11a)$$

$$\text{subject to} \quad \text{Tr}(\mathbf{Q}_s) \leq P, \quad (11b)$$

$$\begin{bmatrix} \mathbf{Q}_s & \tilde{\mathbf{Q}}_s \\ \tilde{\mathbf{Q}}_s^* & \mathbf{Q}_s^* \end{bmatrix} \succeq \mathbf{0}, \quad (11c)$$

$$\text{Tr}(\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H) \leq t, \quad (11d)$$

$$\mathbf{H}_{ps} \tilde{\mathbf{Q}}_s \mathbf{H}_{ps}^T \\ = (\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H)^{1/2} \mathbf{F}_p \mathbf{F}_p^T (\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H)^{*/2}, \quad (11e)$$

$$\mathbf{F}_p \in \mathbb{F}. \quad (11f)$$

Constraints (11e) and (11f) force the SU to transmit a maximally improper signal, such that its improper signature at the PU receiver is contained in  $\mathbb{F}$ , while (11d) restricts the total interference power. Notice that, as long as  $t$  is obtained by solving (10), these constraints ensure that the PU rate is satisfied.

To perform the proposed scheme there are some issues, which we address in this paper, namely,

- 1) What is a good choice for  $\mathbb{F}$ ?
- 2) Given  $\mathbb{F}$ , how can we solve (10) to find the optimal  $t$ ?
- 3) Given  $\mathbb{F}$  and  $t$ , how can we solve (11) to find the optimal parameters of the SU?

### A. Constraints on the improper signature

We start with discussing the choice of  $\mathbb{F}$ , as finding the solution to (10) and (11) depends on it. In order to find a suitable choice for  $\mathbb{F}$ , we need to analyze the impact that this set has on the value of  $t$  that solves (10), as well as on the maximum rate of the SU that is the solution to (11). In this regard, it is important to note that each receiver sees a different improper signature matrix. That is, let the transmit complementary covariance matrix of the SU be

$$\tilde{\mathbf{Q}}_s = \mathbf{Q}_s^{1/2} \mathbf{F} \mathbf{F}^T \mathbf{Q}_s^{*/2}. \quad (12)$$

The complementary covariance matrix of the received signal at the PU is given by (11e), and that at the secondary receiver is

$$\begin{aligned} \mathbf{H}_{ss} \tilde{\mathbf{Q}}_s \mathbf{H}_{ss}^T &= \mathbf{H}_{ss} \mathbf{Q}_s^{1/2} \mathbf{F} \mathbf{F}^T \mathbf{Q}_s^{*/2} \mathbf{H}_{ss}^T \\ &= (\mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H)^{1/2} \mathbf{F}_s \mathbf{F}_s^T (\mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H)^{*/2}. \end{aligned} \quad (13)$$

Therefore, we have in general  $\mathbf{F} \neq \mathbf{F}_p \neq \mathbf{F}_s$ . Let us first study the effect of  $\mathbf{F}_s$ , i.e., the improper signature matrix as seen by the secondary receiver, on the performance achieved by the SU. In our previous work [35], we have shown that

$$R_s(\mathbf{Q}_s, \tilde{\mathbf{Q}}_s^{1/2} \mathbf{F}_w \mathbf{F}_w^T \tilde{\mathbf{Q}}_s^{*/2}) \leq R_s(\mathbf{Q}_s, \tilde{\mathbf{Q}}_s^{1/2} \mathbf{F} \mathbf{F}^T \tilde{\mathbf{Q}}_s^{*/2}) \quad (14)$$

$$R_s(\mathbf{Q}_s, \tilde{\mathbf{Q}}_s^{1/2} \mathbf{F}_b \mathbf{F}_b^T \tilde{\mathbf{Q}}_s^{*/2}) \geq R_s(\mathbf{Q}_s, \tilde{\mathbf{Q}}_s^{1/2} \mathbf{F} \mathbf{F}^T \tilde{\mathbf{Q}}_s^{*/2}), \quad (15)$$

where  $\mathbf{F}_w$  and  $\mathbf{F}_b$ , whose subscripts stand respectively for *worst* and *best*, are such that

$$\begin{aligned} \mathbf{H}_{ss} \mathbf{Q}_s^{1/2} \mathbf{F}_w \mathbf{F}_w^T \mathbf{Q}_s^{*/2} \mathbf{H}_{ss}^T &= (\mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H)^{1/2} \mathbf{V}_s \mathbf{V}_s^T \\ &\quad \times (\mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H)^{*/2}, \end{aligned} \quad (16)$$

$$\begin{aligned} \mathbf{H}_{ss} \mathbf{Q}_s^{1/2} \mathbf{F}_b \mathbf{F}_b^T \mathbf{Q}_s^{*/2} \mathbf{H}_{ss}^T &= (\mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H)^{1/2} \mathbf{V}_s \mathbf{J} \mathbf{V}_s^T \\ &\quad \times (\mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H)^{*/2}, \end{aligned} \quad (17)$$

where  $\mathbf{J}$  is the exchange matrix,<sup>4</sup> and  $\mathbf{V}_s$  is the matrix of eigenvectors of  $\mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H$ , arranged such that the respective eigenvalues are sorted in decreasing order. Expression (14) states that, when the SU transmits a maximally improper signal with a given covariance matrix, its rate is minimum if the improper signature matrix makes each mode fully correlated with its complex conjugate. By contrast, (15) states that the SU rate is maximum if the strongest signal modes are correlated with the complex conjugate of the weakest modes. Equations (16) and (17) indicate that the improper signature matrices at the secondary receiver that lead to the worst and best SU performance, respectively, are  $\mathbf{F}_{s,w} = \mathbf{V}_s$  and  $\mathbf{F}_{s,b} = \mathbf{V}_s \mathbf{J}^{1/2}$ .

It is important to note that the bounds in (14) and (15) are for a fixed covariance matrix  $\mathbf{Q}_s$ . For example suppose that  $\mathbb{F} = \mathbb{U}^N$ , i.e., no specific constraint is imposed on  $\mathbf{F}_p$  other than being unitary. By (14) and (15), this means that  $\mathbf{F} = \mathbf{F}_b$  after solving problem (11). However, the value of  $t$  obtained after solving (10) has to be such that the PU rate is satisfied for all unitary matrices  $\mathbf{F}_p$ . Alternatively, if the set  $\mathbb{F}$  is chosen to be a specific unitary matrix, the value of  $t$  that solves (10) can be made higher by choosing this unitary matrix wisely, as the PU rate has to be ensured only for this unitary matrix.

<sup>4</sup>The exchange matrix is a square matrix with ones on the counterdiagonal and zeros elsewhere.

However, in this case the SU has no freedom in designing matrix  $\mathbf{F}$  when solving (11). We therefore observe a trade-off: reducing the size of  $\mathbb{F}$  makes constraints (11e) and (11f) more stringent but permits increasing the value of  $t$ , thus relaxing constraint (11d).

The relationship between  $t$ ,  $\mathbb{F}$ , and the SU rate is difficult to obtain and depends on all the channel matrices. Additionally, when the PU computes the value of  $t$  by solving (10), this relationship cannot be exploited due to the unawareness of the SU. Therefore, we propose two strategies that either impose no constraint on  $\mathbf{F}_p$  or reduce its feasibility set to a single point:

- 1) **Strategy 1.** The interference must obey a power constraint and be maximally improper:

$$\mathbb{F} = \mathbb{U}^N, \text{ and } \text{Tr}(\mathbf{K}) \leq t_{\min}. \quad (18)$$

- 2) **Strategy 2.** The interference must obey a power constraint and be maximally improper with a specific improper signature matrix at the PU receiver:

$$\mathbb{F} = \mathbf{S} \in \mathbb{U}^N, \text{ and } \text{Tr}(\mathbf{K}) \leq t_{\max}. \quad (19)$$

Strategy 1 is worst-case oriented since no constraints on  $\mathbf{F}_p$  are imposed, which means that the PU rate must be ensured for all  $\mathbf{F}_p \in \mathbb{U}^N$ . Such an uncertainty translates into an IT threshold that is more conservative and, thus, lower than that obtained by Strategy 2. Alternatively, Strategy 2 eliminates the uncertainty on  $\mathbf{F}_p$  by imposing a specific improper signature matrix and is thus best-case oriented. Specifically, if  $\mathbf{F}_p$  is constrained to match the least harmful improper signature matrix for the PU,  $\mathbf{S}$ , the IT threshold can be higher. Therefore, we denote the IT thresholds of Strategy 1 and Strategy 2 as  $t_{\min}$  and  $t_{\max}$ , respectively, and we have  $t_{\min} \leq t \leq t_{\max}$  for any other set  $\mathbb{F}$ . Next section addresses the solution to (10) for these two strategies and the computation of  $\mathbf{S}$ .

## IV. DERIVATION OF THE THRESHOLDS

### A. IT threshold under Strategy 1

The PU obtains the interference power limit for Strategy 1 as the solution of (10) for  $\mathbb{F} = \mathbb{U}^N$ :

$$\text{maximize}_t \quad t, \quad (20a)$$

$$\text{subject to} \quad R_p(\mathbf{K}, \mathbf{F}_p) \geq \bar{R}, \quad \forall \mathbf{K} \in \mathbb{K}(t), \quad \forall \mathbf{F}_p \in \mathbb{U}^N. \quad (20b)$$

Let  $t_{\min}$  be its optimal value. It is then clear that  $\min_{\mathbf{K} \in \mathbb{K}(t_{\min}), \mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) = \bar{R}$ , otherwise  $t_{\min}$  could increase without violating the constraints and would therefore not be the optimal solution of (20). Hence, the foregoing optimization problem can alternatively be solved by finding the most harmful interference covariance matrix  $\mathbf{K}$  and improper signature matrix  $\mathbf{F}_p$ . These matrices can be obtained as the solution of the following optimization problem.

$$\text{minimize}_{t, \mathbf{K}, \mathbf{F}_p} \quad t, \quad (21a)$$

$$\text{subject to} \quad R_p(\mathbf{K}, \mathbf{F}_p) \leq \bar{R}, \quad (21b)$$

$$\mathbf{K} \in \mathbb{K}(t), \quad (21c)$$

$$\mathbf{F}_p \in \mathbb{U}^N. \quad (21d)$$

Notice that  $\mathbf{K}$  and  $\mathbf{F}_p$  are now optimization variables. In order to obtain the solution of the above problem, we first present the following lemma.

*Lemma 1:* Let  $\phi_1 \geq \phi_2 \geq \dots \geq \phi_N$  and  $\mathbf{V}_p$  be the eigenvalues and the matrix of eigenvectors of  $\mathbf{H}_{pp} \mathbf{Q}_p \mathbf{H}_{pp}^H$ , respectively. Then,

$$\begin{aligned} & \min_{\substack{\mathbf{K} \in \mathbb{K}(t) \\ \mathbf{F}_p \in \mathbb{U}^N}} R_p(\mathbf{K}, \mathbf{F}_p) = \\ & \min_{\substack{\sum_{\theta_i \geq 0} \theta_i \leq t \\ \theta_i \geq 0}} \sum_{i=1}^{\frac{N}{2}} \left[ \log_2 \left( 1 + \frac{\phi_i}{\sigma^2 + \theta_i} \right) + \log_2 \left( 1 + \frac{\phi_{N-i+1}}{\sigma^2} \right) \right], \end{aligned} \quad (22)$$

where the value of  $\mathbf{F}_p$  yielding (22) is  $\mathbf{F}_{p,1} = \mathbf{V}_p \mathbf{J}^{1/2}$ , with  $\mathbf{J}$  being the exchange matrix.

*Proof:* Please refer to Appendix A. ■

Lemma 1 characterizes the most harmful maximally improper interference under an interference power constraint. The most harmful improper signature implies full correlation between pairs of modes exhibiting the highest and lowest interference levels. This structure is indeed the same as the one that maximizes the SU rate with a fixed covariance matrix (see the left-hand side of (15)). Notice, however, that the improper signature matrix at the primary receiver,  $\mathbf{F}_p$ , is in general different than that at the secondary receiver,  $\mathbf{F}_s$ . Additionally, the most harmful interference covariance matrix is rank-deficient, consisting of non-null interference in the first  $N/2$  modes, and zero interference in the remaining modes. Consequently, the most harmful maximally improper interference degenerates to a half-rank proper interference. This may seem contradictory at first glance, but it is indeed consistent with our parameterization of complementary covariance matrices in (7). As a matter of fact, there is a thin line between full-rank maximally improper and half-rank proper. To illustrate this, assume that the smallest  $N/2$  eigenvalues of the covariance matrix of a maximally improper signal are equal to  $\epsilon$ , and that the improper signature matrix is as in Lemma 1. For an arbitrarily small  $\epsilon > 0$ , this signal is full-rank and maximally improper, but it is half-rank proper when  $\epsilon = 0$ .

Since  $\mathbf{F}_p$  and the eigenvectors of  $\mathbf{K}$  only affect the rate constraint in (21), we can make use of Lemma 1 to rewrite this problem as

$$\begin{aligned} & \underset{\{\theta_i \geq 0\}_{i=1}^N}{\text{minimize}} \quad \sum_{i=1}^N \theta_i, \quad (23a) \\ & \text{subject to} \quad \sum_{i=1}^{\frac{N}{2}} \left[ \log_2 \left( 1 + \frac{\phi_i}{\sigma^2 + \theta_i} \right) \right. \\ & \quad \left. + \log_2 \left( 1 + \frac{\phi_{N-i+1}}{\sigma^2} \right) \right] \leq \bar{R}. \quad (23b) \end{aligned}$$

Since  $\frac{\partial^2 \log(1+x^{-1})}{\partial x^2} = \frac{2x+1}{x^2(x+1)} > 0$  for  $x > 0$ , the above problem is a convex optimization problem. Furthermore, it satisfies Slater's condition, i.e., the feasible set has a non-empty interior [36]. Hence, the Karush-Kuhn-Tucker (KKT)

conditions are necessary and sufficient for optimality, and they yield the multilevel water-filling solution

$$\theta_{\min,i} = \begin{cases} \left[ \sqrt{\phi_i \left( \frac{1}{4} \phi_i + \mu \right)} - \left( \frac{1}{2} \phi_i + \sigma^2 \right) \right]^+ & i = 1, \dots, \frac{N}{2}, \\ 0, & i = \frac{N}{2} + 1, \dots, N, \end{cases} \quad (24)$$

where  $\phi_i$  is the  $i$ th eigenvalue of  $\mathbf{H}_{pp} \mathbf{Q}_p \mathbf{H}_{pp}^H$  and  $\mu$  is chosen such that the rate constraint holds with equality. The IT threshold under Strategy 1 is then

$$t_{\min} = \sum_{i=1}^N \theta_{\min,i}. \quad (25)$$

### B. IT threshold under Strategy 2

In Strategy 2, we constrain the improper signature matrix  $\mathbf{F}_p$  to be equal to the least harmful such matrix for the PU, denoted here as  $\mathbf{S}$ . We consider this matrix to be a function of the interference covariance matrix, which is possible because the SUs know the actual interference covariance matrix provoked by their transmitted signals. Following these lines, the improper signature matrix is selected such that it maximizes the interference power limit, which the PU carries out solving the following optimization problem

$$\underset{t}{\text{maximize}} \quad t, \quad (26a)$$

$$\text{subject to} \quad \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) \geq \bar{R}, \quad \forall \mathbf{K} \in \mathbb{K}(t). \quad (26b)$$

Let  $t_{\max}$  be the optimal solution of this problem. It is then clear that  $\min_{\mathbf{K} \in \mathbb{K}(t_{\max})} \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) = \bar{R}$ , otherwise  $t_{\max}$  would not be the optimal solution. Therefore, problem (26) can be rewritten as

$$\underset{t, \mathbf{K} \in \mathbb{K}(t)}{\text{minimize}} \quad t, \quad (27a)$$

$$\text{subject to} \quad \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) \leq \bar{R}. \quad (27b)$$

In the following, we will derive a lower bound on the optimal solution of the above problem, which coincides with its optimal solution under some conditions that will also be determined. To this end, we first present the following lemma.

*Lemma 2:* Let  $\phi_1 \geq \phi_2 \geq \dots \geq \phi_N$  and  $\mathbf{V}_p$  be the eigenvalues and the matrix of eigenvectors of  $\mathbf{H}_{pp} \mathbf{Q}_p \mathbf{H}_{pp}^H$ , respectively. Then,

$$\begin{aligned} & \min_{\mathbf{K} \in \mathbb{K}(t)} \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) = \\ & \min_{\substack{\sum_{\theta_i \geq 0} \theta_i \leq t \\ \theta_i \geq 0}} \sum_{i=1}^N \frac{1}{2} \log_2 \left[ 1 + \frac{\phi_i}{\sigma^2} \left( 1 + \frac{\sigma^2 + \phi_i}{\sigma^2 + 2\theta_i} \right) \right], \end{aligned} \quad (28)$$

for  $t \geq \bar{t}$ , where  $\bar{t} = \sum_{i=1}^N \bar{\theta}_i$ , with

$$\bar{\theta}_i = \frac{1}{2} \left[ \sqrt{\phi_i \left( \frac{1}{4} \phi_i + \bar{\mu} \right)} - \left( \frac{1}{2} \phi_i + \sigma^2 \right) \right]^+, \quad (29)$$

$$\bar{\mu} = \frac{\phi_1 \phi_2 \left( 1 + \frac{\phi_1}{\sigma^2} \right) \left( 1 + \frac{\phi_2}{\sigma^2} \right)}{\sigma^2 \left( \frac{\phi_1 \phi_2}{\sigma^4} - 1 \right)^2}. \quad (30)$$

If  $t < \bar{t}$ , then

$$\begin{aligned} & \min_{\substack{\sum_{\theta_i \leq t} \\ \theta_i \geq 0}} \sum_{i=1}^N \frac{1}{2} \log_2 \left[ 1 + \frac{\phi_i}{\sigma^2} \left( 1 + \frac{\sigma^2 + \phi_i}{\sigma^2 + 2\theta_i} \right) \right] \\ & \leq \min_{\mathbf{K} \in \mathbb{K}(t)} \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) < \\ & \min_{\substack{\sum_{\theta_i \leq \bar{t}} \\ \theta_i \geq 0}} \sum_{i=1}^N \frac{1}{2} \log_2 \left[ 1 + \frac{\phi_i}{\sigma^2} \left( 1 + \frac{\sigma^2 + \phi_i}{\sigma^2 + 2\theta_i} \right) \right]. \end{aligned} \quad (31)$$

Furthermore, the improper signature matrix leading to (28) is  $\mathbf{F}_{p,2} = \mathbf{U}$ , with  $\mathbf{U}$  being the matrix of eigenvectors of  $\mathbf{K}$ .

*Proof:* Please refer to Appendix B. ■

Lemma 2 states that, at least for  $t \geq \bar{t}$ , the least harmful improper signature implies that the interference at each stream is maximally improper, while it is uncorrelated with the complex conjugate of the interference affecting the other streams. This means that the worst-case interference under Strategy 2 is, at least for  $t \geq \bar{t}$ , spatially-unconstrained maximally improper, as opposed to Strategy 1, where the worst-case interference is spatially-constrained proper. Please note that this result amends our respective statement in [34], which is herewith diminished to the conjecture that (28) holds for all  $t \geq 0$ , as observed through extensive numerical simulations.

As with Strategy 1, we also observe in this case a correspondence with the analysis of the SU performance. Specifically, the value of  $\mathbf{F}_s$  that minimizes the rate of the SU (left-hand side of (14)) resembles the value of  $\mathbf{F}_{p,2}$  in Lemma 2, yielding the same improper correlation structure but at different receiver spaces. Again, this does not imply that setting  $\mathbb{F} = \mathbf{S} = \mathbf{F}_{p,2}$  makes the SU achieve its lower bound in (14), as  $\mathbf{F}_s$  will in general have a different structure due to different channel matrices.

Making use of Lemma 2, we can obtain a lower bound on the optimal solution of (26) by rewriting (27) as

$$\begin{aligned} & \underset{\{\theta_i \geq 0\}_{i=1}^N}{\text{minimize}} \quad \sum_{i=1}^N \theta_i, \quad (32a) \\ & \text{subject to} \quad \sum_{i=1}^N \frac{1}{2} \log_2 \left[ 1 + \frac{\phi_i}{\sigma^2} \left( 1 + \frac{\sigma^2 + \phi_i}{\sigma^2 + 2\theta_i} \right) \right] \leq \bar{R}. \end{aligned} \quad (32b)$$

According to Lemma 2, if the optimal solution of this problem fulfills  $t^* = \sum_{i=1}^N \theta_i^* \geq \bar{t}$ , it will be as well the optimal solution of the original problem, and it will be a lower bound otherwise. Problem (32) is convex and satisfies Slater's condition [36]. Hence, its optimal solution can be obtained using the KKT conditions, which yields the multilevel water-filling solution

$$\theta_{\max,i} = \frac{1}{2} \left[ \sqrt{\phi_i \left( \frac{1}{4} \phi_i + \mu' \right)} - \left( \frac{1}{2} \phi_i + \sigma^2 \right) \right]^+, \quad i = 1, \dots, N, \quad (33)$$

with  $\mu'$  such that the rate constraint holds with equality. We then obtain the IT threshold under Strategy 2 as

$$t_{\max} = \sum_{i=1}^N \theta_{\max,i}. \quad (34)$$

## V. OPTIMIZATION OF THE SECONDARY USER

In this section, we design the transmission strategy of the SU for the two proposed strategies. That is, we address problem (11) for  $\mathbb{F} = \mathbb{U}^N$  and for  $\mathbb{F} = \mathbf{S} = \mathbf{F}_{p,2}$ .

For Strategy 1, we will reveal the structure of the optimal solution. In particular, we will show that it admits a closed-form expression when only one of the power constraints (interference power or power budget) is active, which allows a more efficient computation of the transmission scheme as explained in the next subsection.

For Strategy 2, we show that an adequate approximation of the problem allows the application of an efficient modified steepest decent method [37].

### A. Optimization under Strategy 1

Under this strategy, the SU is constrained with a total interference power  $t_{\min}$  but can use an arbitrary unitary improper signature matrix,  $\mathbf{F}$ . Therefore, by (15), the improper signature matrix that maximizes the SU rate is  $\mathbf{F}^* = \mathbf{U}_s \mathbf{J}^{1/2}$  [35], where  $\mathbf{U}_s$  is the matrix of eigenvectors of  $\mathbf{Q}_s^{1/2} \mathbf{H}_{ss}^H \mathbf{H}_{ss} \mathbf{Q}_s^{1/2}$ , arranged such that the corresponding eigenvalues are sorted in decreasing order. The resulting achievable rate is [35]

$$\begin{aligned} R_s(\mathbf{Q}_s) = & \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{Q}_s^{1/2} \mathbf{G}_{ss} \mathbf{Q}_s^{1/2} \right. \\ & \left. + \frac{1}{\sigma^2} \mathbf{U}_s \mathbf{J} \mathbf{U}_s^T (\mathbf{Q}_s^{1/2} \mathbf{G}_{ss} \mathbf{Q}_s^{1/2})^* \mathbf{U}_s^* \mathbf{J} \mathbf{U}_s^H \right|, \end{aligned} \quad (35)$$

where  $\mathbf{G}_{ss} = \mathbf{H}_{ss}^H \mathbf{H}_{ss}$ , and we have dropped the dependence of  $R_s$  on  $\mathbf{Q}_s$ , as  $\mathbf{F}$  has already been fixed. Clearly,  $\mathbf{F}^*$  depends on the eigenvalues of  $\mathbf{Q}_s$ , which makes the problem difficult to solve. In [35], we proposed a simple but effective alternating optimization algorithm to find a suboptimal solution. However, we show in the following that the optimal solution can be found in closed form when only one of the power constraints (either power budget or interference power) is active, which allow us to provide insights into the optimal solution for the general case. To this end, we express  $\mathbf{F}^* = \mathbf{U}_s \mathbf{\Pi}^{1/2}$ , where now  $\mathbf{U}_s$  contains the eigenvectors of  $\mathbf{Q}_s^{1/2} \mathbf{G}_{ss} \mathbf{Q}_s^{1/2}$  in an arbitrary order, and  $\mathbf{\Pi}$  is a symmetric permutation matrix. We analyze in the following the optimal structure of  $\mathbf{Q}_s$  for a fixed permutation  $\mathbf{\Pi}$ , which is the matrix that solves the problem

$$\begin{aligned} & \underset{\mathbf{Q}_s \succeq 0}{\text{maximize}} \quad \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{Q}_s^{1/2} \mathbf{G}_{ss} \mathbf{Q}_s^{1/2} \right. \\ & \quad \left. + \frac{1}{\sigma^2} \mathbf{U}_s \mathbf{\Pi} \mathbf{U}_s^T (\mathbf{Q}_s^{1/2} \mathbf{G}_{ss} \mathbf{Q}_s^{1/2})^* \mathbf{U}_s^* \mathbf{\Pi} \mathbf{U}_s^H \right| \end{aligned} \quad (36a)$$

$$\text{subject to} \quad \text{Tr}(\mathbf{Q}_s) \leq P, \quad (36b)$$

$$\text{Tr}(\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H) \leq t_{\min}. \quad (36c)$$

The structure of the optimal solution of this problem is given in the following lemma.

*Lemma 3:* The optimal solution of (36) can be written as

$$\mathbf{Q}_s^* = \mathbf{X}^* \mathbf{X}^{*H}, \quad \mathbf{X}^* = (\mathbf{H}_{ss}^H \mathbf{H}_{ss})^{-1/2} \mathbf{V} \mathbf{\Sigma}, \quad (37)$$

where  $\mathbf{V}$  is a unitary matrix and  $\mathbf{\Sigma}$  is a diagonal matrix with non-negative entries.

*Proof:* This result follows from Theorem 1 in [37]. Even though that result is not directly applicable since the cost function does not have the same form, we can use the same arguments to prove the lemma. Specifically, we just need to show that the optimal solution diagonalizes  $\mathbf{G}_{ss}$ . To show this, and following the lines of [37], suppose that there is a solution  $\mathbf{Q}_{s,1} = \mathbf{X}_1 \mathbf{X}_1^H$  such that  $\mathbf{X}_1^{1/2} \mathbf{G}_{ss} \mathbf{X}_1^{1/2}$  is not diagonal. Now consider the solution  $\mathbf{X}_2 = \mathbf{X}_1 \mathbf{R}$ , where  $\mathbf{R}$  is a unitary matrix. Notice that, since  $\mathbf{R}$  is unitary,  $\mathbf{Q}_{s,2} = \mathbf{X}_2 \mathbf{X}_2^H = \mathbf{Q}_{s,1}$ . Thus,  $\text{Tr}(\mathbf{Q}_{s,1}) = \text{Tr}(\mathbf{Q}_{s,2})$  and  $\text{Tr}(\mathbf{H}_{ps} \mathbf{Q}_{s,1} \mathbf{H}_{ps}^H) = \text{Tr}(\mathbf{H}_{ps} \mathbf{Q}_{s,2} \mathbf{H}_{ps}^H)$ . Furthermore, if  $\mathbf{U}_1$  contains the eigenvectors of  $\mathbf{X}_1^{1/2} \mathbf{G}_{ss} \mathbf{X}_1^{1/2}$  we have  $\mathbf{U}_2 = \mathbf{R}^H \mathbf{U}_1$  and the cost function becomes

$$\begin{aligned} & \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{X}_2^H \mathbf{G}_{ss} \mathbf{X}_2 + \frac{1}{\sigma^2} \mathbf{U}_2 \mathbf{\Pi} \mathbf{U}_2^T (\mathbf{X}_2^H \mathbf{G}_{ss} \mathbf{X}_2)^* \right. \\ & \times \left. \mathbf{U}_2^* \mathbf{\Pi} \mathbf{U}_2^H \right| = \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{R}^H \mathbf{X}_1^H \mathbf{G}_{ss} \mathbf{X}_1 \mathbf{R} + \frac{1}{\sigma^2} \mathbf{R}^H \mathbf{U}_1 \right. \\ & \times \left. \mathbf{\Pi} \mathbf{U}_1^T \mathbf{R}^* (\mathbf{R}^H \mathbf{X}_1^H \mathbf{G}_{ss} \mathbf{X}_1 \mathbf{R})^* \mathbf{R}^T \mathbf{U}_1^* \mathbf{\Pi} \mathbf{U}_1^H \mathbf{R} \right| = \frac{1}{2} \log_2 \left| \mathbf{I} \right. \\ & \left. + \frac{1}{\sigma^2} \mathbf{X}_1^H \mathbf{G}_{ss} \mathbf{X}_1 + \frac{1}{\sigma^2} \mathbf{U}_1 \mathbf{\Pi} \mathbf{U}_1^T (\mathbf{X}_1^H \mathbf{G}_{ss} \mathbf{X}_1)^* \mathbf{U}_1^* \mathbf{\Pi} \mathbf{U}_1^H \right|, \end{aligned} \quad (38)$$

where we have used  $\log |\mathbf{I} + \mathbf{A}\mathbf{B}| = \log |\mathbf{I} + \mathbf{B}\mathbf{A}|$  and the fact that  $\mathbf{R}$  is unitary. Therefore, if  $\mathbf{X}_1$  is an optimal solution, so is  $\mathbf{X}_2 = \mathbf{X}_1 \mathbf{R}$ . Finally,  $\mathbf{R}$  can be chosen such that  $\mathbf{X}_2^H \mathbf{G}_{ss} \mathbf{X}_2$  is diagonal. The structure of such a solution is given by (37) [37], which concludes the proof. ■

Let us first assume that only one of the power constraints is active, i.e., either the power budget constraint or the interference power constraint is active, but not both at the same time. To this end, we define

$$\mathbf{A} = \begin{cases} \mathbf{G}_{ss}^{-1}, & \text{if only (36b) is active,} \\ \mathbf{G}_{ss}^{-1/2} \mathbf{G}_{ps} \mathbf{G}_{ss}^{-1/2}, & \text{if only (36c) is active,} \end{cases} \quad (39)$$

where  $\mathbf{G}_{ps} = \mathbf{H}_{ps}^H \mathbf{H}_{ps}$ . Let  $\mathbf{A} = \mathbf{V}_A \mathbf{\Lambda}_A \mathbf{V}_A^H$  be the singular value decomposition (SVD). By [37] we then have  $\mathbf{V} = \mathbf{V}_A$ , and (36) becomes

$$\begin{aligned} & \text{maximize}_{\{p_i \geq 0\}_{i=1}^M} \sum_{i=1}^M \frac{1}{2} \log_2 \left( 1 + \frac{a_i p_i + a_{\pi(i)} p_{\pi(i)}}{\sigma^2} \right) \quad (40a) \\ & \text{subject to} \quad \sum_{i=1}^M p_i \leq \gamma, \quad (40b) \end{aligned}$$

where  $\pi(\cdot)$  is the permutation performed by  $\mathbf{\Pi}$ ,  $a_i^{-1} = (\mathbf{\Lambda}_A)_{ii}$ ,  $\mathbf{\Sigma} = \mathbf{\Lambda}_A^{-1/2} \text{diag}(\sqrt{p_1}, \dots, \sqrt{p_M})$ , and  $\gamma = P$  or  $\gamma = t_{\min}$  depending on the active constraint. Let us analyze the optimal solution to this problem. Assuming, without loss of generality, that  $a_i \geq a_{\pi(i)}$ , we have, for  $\pi(i) \neq i$ ,

$$\log_2 \left( 1 + \frac{a_i p_i + a_{\pi(i)} p_{\pi(i)}}{\sigma^2} \right) \leq \log_2 \left( 1 + \frac{a_i (p_i + p_{\pi(i)})}{\sigma^2} \right). \quad (41)$$

Therefore, the optimal power allocation fulfills  $p_{\pi(i)} = 0$  for  $a_i \geq a_{\pi(i)}$ . Furthermore, if we assume that  $a_1 \geq a_2 \geq \dots \geq a_M$ , it is then clear that the permutation that maximizes the achievable rate is  $\pi(i) = M - i + 1$ . Hence,

when only the transmit power constraint or the interference power constraint is active, the optimal maximally improper transmission degenerates to a half-rank proper one. Notice that this coincides with the structure of the worst-case maximally improper interference at the primary receiver derived in the previous section. This result validates Strategy 1, where the interference power limit is computed under the worst-case assumption, or, in other words, assuming that the SU uses its optimal transmit covariance matrix under a maximally improper constraint. The optimal solution of (40) is then given by the waterfilling power allocation

$$p_i^* = \begin{cases} \left( \nu - \frac{\sigma^2}{a_i} \right)^+ & i = 1, \dots, \frac{M}{2}, \\ 0 & i = \frac{M}{2} + 1, \dots, M, \end{cases} \quad (42)$$

with  $\nu$  such that  $\sum_{i=1}^M p_i^* = \gamma$ .

Now assume that both constraints, i.e., transmit power and interference power, are simultaneously active. By Lemma 3, (36) can be equivalently written as

$$\begin{aligned} & \text{maximize}_{\mathbf{V} \in \mathcal{U}^M, \{p_i \geq 0\}_{i=1}^M} \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{\Sigma} \mathbf{\Sigma}^H + \frac{1}{\sigma^2} \mathbf{\Pi} \mathbf{\Sigma} \mathbf{\Sigma}^H \mathbf{\Pi} \right| \quad (43a) \end{aligned}$$

$$\text{subject to} \quad \text{Tr}(\mathbf{V} \mathbf{G}_{ss}^{-1} \mathbf{V}^H \mathbf{\Sigma} \mathbf{\Sigma}^H) \leq P, \quad (43b)$$

$$\text{Tr}(\mathbf{V} \mathbf{G}_{ss}^{-1/2} \mathbf{G}_{ps} \mathbf{G}_{ss}^{-1/2} \mathbf{V}^H \mathbf{\Sigma} \mathbf{\Sigma}^H) \leq t_{\min}, \quad (43c)$$

$$\mathbf{\Sigma} = \text{diag}(p_1, \dots, p_M). \quad (43d)$$

In order to determine whether or not the structure of the optimal solution in the single-constraint case also applies to the general case, we analyze the optimal value of  $\mathbf{\Sigma}$  in terms of  $\mathbf{V}$  and the permutation  $\mathbf{\Pi}$ . To this end, we study the solution of the above problem for a fixed  $\mathbf{V}$ , which yields the problem

$$\text{maximize}_{\{p_i \geq 0\}_{i=1}^M} \frac{1}{2} \sum_{i=1}^M \log_2 \left( 1 + \frac{b_i p_i + b_{\pi(i)} p_{\pi(i)}}{\sigma^2} \right) \quad (44a)$$

$$\text{subject to} \quad \sum_{i=1}^M p_i \leq P, \quad (44b)$$

$$\sum_{i=1}^M c_i p_i \leq t_{\min}, \quad (44c)$$

where  $b_i^{-1} = [\mathbf{V} \mathbf{G}_{ss}^{-1} \mathbf{V}^H]_{ii}$ ,  $c_i = [\mathbf{V} \mathbf{G}_{ss}^{-1/2} \mathbf{G}_{ps} \mathbf{G}_{ss}^{-1/2} \mathbf{V}^H]_{ii}$ , and now  $\mathbf{\Sigma} = \text{diag}(b_1 \sqrt{p_1}, \dots, b_M \sqrt{p_M})$ . Since the foregoing problem is convex and satisfies Slater's condition [36], the KKT conditions are necessary and sufficient for optimality. The Lagrangian function yields

$$p_i = \frac{1}{c_i \beta + \mu - \nu_i} - \left( \frac{\sigma^2}{b_i} + p_{\pi(i)} \frac{b_{\pi(i)}}{b_i} \right), \quad (45)$$

$$p_{\pi(i)} = \frac{1}{c_{\pi(i)} \beta + \mu - \nu_{\pi(i)}} - \left( \frac{\sigma^2}{b_{\pi(i)}} + p_i \frac{b_i}{b_{\pi(i)}} \right), \quad (46)$$

where  $\beta$ ,  $\mu$  and  $\nu_i$  are the Lagrange multipliers associated with the interference power, transmit power and non-negativity constraints, respectively. Combining both expressions, we obtain

$$(b_i c_{\pi(i)} - \phi_{\pi(i)} c_i) \beta + (b_i - b_{\pi(i)}) \mu = b_i \nu_{\pi(i)} - b_{\pi(i)} \nu_i. \quad (47)$$

From the KKT conditions,  $p_i$  and  $p_{\pi(i)}$  are simultaneously non-zero if  $\nu_i = \nu_{\pi(i)} = 0$ . In such a case, (47) yields

$$\beta = \frac{b_i - b_{\pi(i)}}{b_{\pi(i)}c_i - b_i c_{\pi(i)}} \mu. \quad (48)$$

First we observe that, since  $\beta$  and  $\mu$  are both non-negative, this condition can only hold if  $b_i \geq b_{\pi(i)}$  and  $b_{\pi(i)}c_i \geq b_i c_{\pi(i)}$ . Second, the coefficients  $b_i$  and  $c_i$  depend on the channel matrices  $\mathbf{H}_{ss}$  and  $\mathbf{H}_{ps}$ . When these channel matrices are independently drawn from a continuous distribution, we have that

$$\frac{b_i - b_{\pi(i)}}{b_{\pi(i)}c_i - b_i c_{\pi(i)}} \neq \frac{b_j - b_{\pi(j)}}{b_{\pi(j)}c_j - b_j c_{\pi(j)}}, \quad (49)$$

almost surely for  $i \neq j$ . Because of that, (48) will be satisfied almost surely for no more than one index  $i$ . This means that the optimal solution will also follow the half-rank structure except for at most two data streams, which will be pairwise correlated, thus resembling again the structure of the worst-case interference covariance matrix under Strategy 1. The optimal permutation, however, cannot be analytically determined. Therefore, even though we have obtained interesting insights into the structure of the optimal solution for the general case, an efficient way to solve (43) is by the algorithm proposed in [35], although only a suboptimal solution can be guaranteed in general.

To sum up, we have obtained the following insights into the structure of the maximally improper scheme that maximizes the rate of the SU (problem (11)) under Strategy 1:

- When only one power constraint is active, the optimal scheme degenerates to a half-rank proper one. The optimal transmission parameters can then be obtained in closed form.
- When both power budget and interference power constraints are active, the optimal scheme is a half-rank proper one except for almost surely no more than two data streams, which are pairwise correlated (in the improper sense). No closed-form expression is available in this case, and a suboptimal solution can be obtained by the alternating optimization algorithm proposed in [35]. This algorithm has guaranteed convergence to a stationary point of the original problem [38].

Notice that the optimal structure for the case of one active constraint can be used as follows. First, the closed form solutions (42) are used for each power constraint. If each of them violates the other power constraint, then both constraints are active and the algorithm in [35] is resorted to. Otherwise, the solution for which the other power constraint is fulfilled is the optimal solution to the problem.

## B. Optimization under Strategy 2

Strategy 2 forces the improper signature matrix of the interference to be equal to the eigenvectors of the interference covariance matrix. Therefore, (11) can be rewritten in this case

as

$$\underset{\mathbf{Q}_s, \tilde{\mathbf{Q}}_s}{\text{maximize}} \quad \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \begin{bmatrix} \mathbf{H}_{ss} \mathbf{Q}_s \mathbf{H}_{ss}^H & \mathbf{H}_{ss} \tilde{\mathbf{Q}}_s \mathbf{H}_{ss}^H \\ \mathbf{H}_{ss}^* \tilde{\mathbf{Q}}_s^* \mathbf{H}_{ss}^H & \mathbf{H}_{ss}^* \mathbf{Q}_s^* \mathbf{H}_{ss}^H \end{bmatrix} \right| \quad (50a)$$

$$\text{subject to} \quad \text{Tr}(\mathbf{Q}_s) \leq P, \quad (50b)$$

$$\text{Tr}(\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H) \leq t_{\max}, \quad (50c)$$

$$\begin{bmatrix} \mathbf{Q}_s & \tilde{\mathbf{Q}}_s \\ \tilde{\mathbf{Q}}_s^* & \mathbf{Q}_s^* \end{bmatrix} \succeq \mathbf{0}, \quad (50d)$$

$$\mathbf{H}_{ps} \tilde{\mathbf{Q}}_s \mathbf{H}_{ps}^T = (\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H)^{1/2} \mathbf{F}_p \times \mathbf{F}_p^T (\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H)^{*/2}, \quad (50e)$$

$$\mathbf{F}_p = \text{eigvect}(\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H), \quad (50f)$$

where  $\text{eigvect}(\cdot)$  denotes the matrix of eigenvectors. The last two constraints make the above problem very difficult to solve optimally. In the following, we propose an efficient algorithm to find a suboptimal solution, which jointly designs  $\mathbf{Q}_s$  and  $\tilde{\mathbf{Q}}_s$ . To this end, let us first look at the required structure in the interference covariance matrix. Since the improper signature matrix of the interference has to be equal to the eigenvectors of the interference covariance matrix, the non-zero eigenvalues of the resulting interference augmented covariance matrix are  $2\lambda(\mathbf{H}_{ps} \mathbf{Q}_s \mathbf{H}_{ps}^H)$  [35, Lemma 1]. Furthermore, by expressing  $\tilde{\mathbf{Q}}_s = \mathbf{X}\mathbf{X}^T$  and  $\mathbf{Q}_s = \mathbf{X}\mathbf{X}^H$ , with  $\mathbf{X} = \mathbf{Q}_s^{1/2} \mathbf{F}$ ,  $\mathbf{F}$  being the improper signature matrix of the transmit signal, the non-zero eigenvalues are also equal to  $\lambda(\mathbf{X}^H \mathbf{H}_{ps}^H \mathbf{H}_{ps} \mathbf{X} + (\mathbf{X}^H \mathbf{H}_{ps}^H \mathbf{H}_{ps} \mathbf{X})^*)$ . Therefore,  $\mathbf{X}$  has to be such that  $\lambda(\mathbf{X}^H \mathbf{H}_{ps}^H \mathbf{H}_{ps} \mathbf{X} + (\mathbf{X}^H \mathbf{H}_{ps}^H \mathbf{H}_{ps} \mathbf{X})^*) = 2\lambda(\mathbf{H}_{ps} \mathbf{X} \mathbf{X}^H \mathbf{H}_{ps}^H)$ , which means  $\mathbf{X}^H \mathbf{H}_{ps}^H \mathbf{H}_{ps} \mathbf{X} = (\mathbf{X}^H \mathbf{H}_{ps}^H \mathbf{H}_{ps} \mathbf{X})^*$ . This can be satisfied by forcing  $\mathbf{X}^H \mathbf{H}_{ps}^H \mathbf{H}_{ps} \mathbf{X}$  to be diagonal, yielding [37]

$$\mathbf{X}^* = (\mathbf{H}_{ps}^H \mathbf{H}_{ps})^{-1/2} \mathbf{V} \mathbf{\Sigma} = \mathbf{G}_{ps}^{-1/2} \mathbf{V} \mathbf{\Sigma}, \quad (51)$$

where  $\mathbf{V}$  is a unitary matrix and  $\mathbf{\Sigma}$  is a diagonal matrix with non-negative entries. Plugging (51) into (50) we obtain the optimization problem

$$\underset{\substack{\mathbf{V} \in \mathbb{U}^M \\ \{p_i \geq 0\}_{i=1}^M}}{\text{maximize}} \quad \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{\Sigma} \mathbf{\Sigma}^H (\mathbf{V}^H \mathbf{A} \mathbf{V} + \mathbf{V}^T \mathbf{A}^* \mathbf{V}^*) \right| \quad (52a)$$

$$\text{subject to} \quad \text{Tr}(\mathbf{\Sigma} \mathbf{\Sigma}^H \mathbf{V}^H \mathbf{G}_{ps}^{-1} \mathbf{V}) \leq P, \quad (52b)$$

$$\text{Tr}(\mathbf{\Sigma} \mathbf{\Sigma}^H) \leq t_{\max}, \quad (52c)$$

$$\mathbf{\Sigma} = \text{diag}(p_1, \dots, p_M), \quad (52d)$$

where  $\mathbf{A} = \mathbf{G}_{ps}^{-1/2} \mathbf{G}_{ss} \mathbf{G}_{ps}^{-1/2}$ . The above problem is still non-convex, but we can find a suboptimal solution using the modified projected steepest descent algorithm proposed in [37], which converges to a stationary point.

## C. Extension to multiple SUs

Extending the above results to the multiuser case poses some challenges. First, a good design of the transmission scheme depends on the specific secondary network configuration. To achieve practical algorithms, distributed approaches should be aimed for. This poses the interesting question of how to efficiently and effectively decouple the constraints,

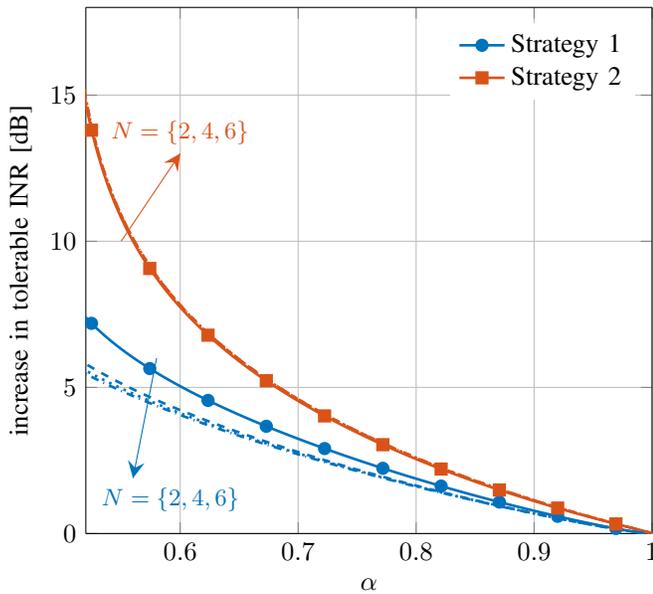


Fig. 1. Increase in tolerable INR with respect to PGS. We depict the results for different number  $N$  of antennas.

so that each SU has a local constraint on the interference it causes to the PU. While the total interference power constraint can be evenly divided among the SUs, how to transfer the constraint on the improper signature matrix, which has to be fulfilled by the aggregate interference covariance matrix, to local constraints is an interesting line of further work.

#### D. Discussion

In the foregoing analysis we have drawn the following insights:

- Strategy 1 constrains the interference to be maximally improper with arbitrary correlation structure. Because of that, the interference power limit is computed for the most harmful maximally improper structure, which is half-rank proper. Under this strategy, the SU designs its transmission scheme to maximize its rate under the interference power constraint and such that it is maximally improper. This optimization also yields a proper transmit signal that has half rank if only one of the power constraints is active. When both power constraints are active (power budget and interference power), the optimal design is half-rank except for no more than two data streams. This validates Strategy 1, as the worst-case assumption made when deriving the interference power threshold  $t$  closely follows the optimal transmission scheme of the SU.
- Strategy 2 increases the interference power threshold at the cost of an additional constraint on the improper structure of the interference. Specifically,  $t$  is maximized by forcing the interference to have the improper structure that is the least detrimental for the primary receiver. Under this strategy, the SU designs its transmission strategy under the additional constraint of matching this specific improper structure. By analyzing the improper structure that is the least favorable for the SU, we have observed that the improper structure imposed by Strategy

#### Step 1:

##### 1) Strategy 1:

- The PU obtains  $t$  using (24) and (25).

##### 2) Strategy 2:

- The PU obtains  $t$  using (33) and (34).

#### Step 2:

- 1) The PU broadcast the value of  $t$  and the strategy index, which are acquired by the SU.

#### Step 3:

##### 1) Strategy 1:

- The SU solves (36) to find its transmission scheme. Note that the SU transmits a maximally improper signal with an unconstrained improper signature.

##### 2) Strategy 2:

- The SU solves (50) to find its transmission scheme. Note that the SU transmits a maximally improper signal with a constrained improper signature.

Fig. 2. Workflow of the proposed strategies.

2 does not match the worst improper structure from the SU standpoint. This means that Strategy 2 allows maximizing the interference power threshold  $t$  by imposing an improper structure that is not the least favorable for the SU.

Finally, a summary of the workflow of the two strategies is shown in Fig. 2.

## VI. NUMERICAL RESULTS

We present in this section some numerical examples to illustrate our findings. For all the simulations, we set the noise variance  $\sigma^2 = 1$ , and thus the signal-to-noise ratio (SNR) is equal to the transmit power. The primary transmitter uses the optimal strategy in the absence of interference, namely, SVD of the direct link and waterfilling power allocation on the resulting eigenmodes. The rate constraint of the PU is set as a percentage of its capacity in the absence of interference, i.e.,  $\bar{R} = \alpha R_p(\mathbf{0}, \mathbf{0})$ , where  $\alpha \in [0, 1]$  is the loading factor. Unless stated otherwise, the elements of all channel matrices are independent and identically distributed (i.i.d.) as proper complex Gaussian with zero mean and unit variance. All results are averaged over 1000 independent channel realizations.

### A. Interference power limit

We first compare the two interference power limits,  $t_{\min}$  and  $t_{\max}$  corresponding to Strategy 1 and Strategy 2, respectively. Figure 1 depicts the increase in tolerable interference-to-noise ratio (INR) over PGS, as a function of  $\alpha$  and for different number  $N$  of antennas at the PU. We observe that the difference between  $t_{\max}$  and  $t_{\min}$  is larger as  $\alpha$  decreases and as the number of antennas increases. It is therefore expected

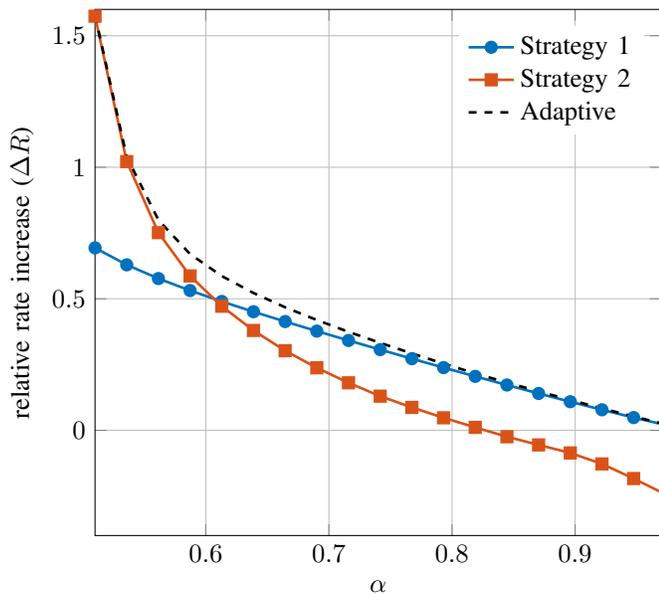


Fig. 3. Relative increase in achievable rate for the different IGS strategies as a function of  $\alpha$ . The power budget constraint is not active and the number of antennas is  $N = M = 2$ .

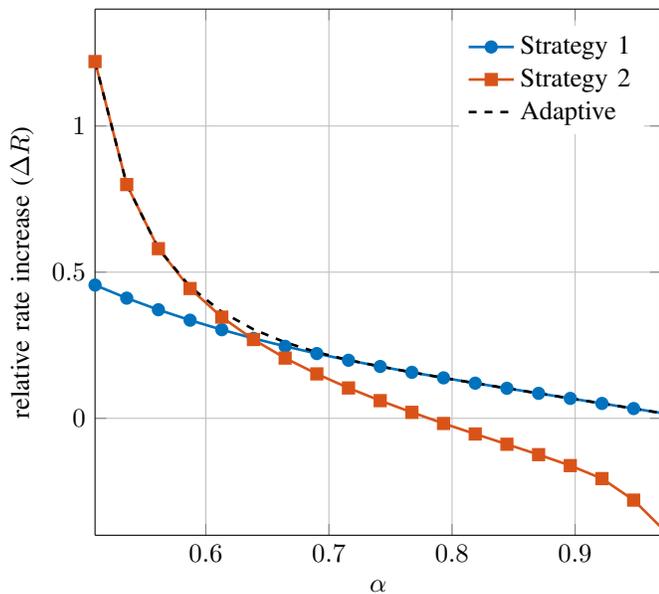


Fig. 4. Relative increase in achievable rate for the different IGS strategies as a function of  $\alpha$ . The power budget constraint is not active and the number of antennas for both users is  $N = M = 4$ .

that Strategy 2 will outperform Strategy 1 (in terms of SU performance) in these cases. Indeed, as  $\alpha$  approaches 0.5,  $t_{\max}$  goes towards infinity, while  $t_{\min}$  remains finite. This is because the PU can achieve half its maximum performance by neglecting, e.g., the imaginary part of each stream. Since Strategy 2 forces such an interference structure at the primary receiver, the result follows. As the number  $N$  of antennas increases, the threshold increase over PGS is approximately constant for Strategy 2, while, for Strategy 1, the gap between IGS and PGS decreases.

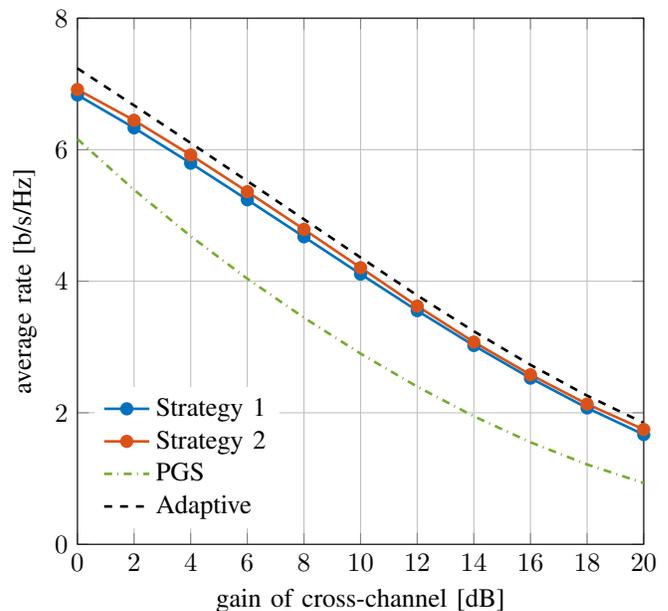


Fig. 5. Average achievable rate for the different techniques as a function of the gain of the cross-channel,  $\mathbf{H}_{ps}$ . The number of antennas is  $N = M = 2$ .

### B. SU performance

We now evaluate the performance achieved by the SU for PGS and the proposed IGS strategies. In order to better illustrate the behavior of the two strategies, we first drop the power budget constraint from the SU optimization, that is, we assume that the interference power constraint dominates the power budget constraint and thus it is not active. Figures 3 and 4 show the relative increase in average rate over PGS as a function of  $\alpha$  for  $N = M = 2$  and  $N = M = 4$ , respectively. The relative rate increase is defined as

$$\Delta R = \frac{R_{\text{imp}} - R_{\text{prop}}}{R_{\text{prop}}}, \quad (53)$$

where  $R_{\text{imp}}$  is the average rate achieved by IGS (either by Strategy 1 or Strategy 2) and  $R_{\text{prop}}$  is the average rate by PGS. The SNR of the PU is set to 20 dB. We also depict the performance achieved by an adaptive strategy that chooses the transmission scheme attaining the highest instantaneous rate. Such a scheme can be realized if the SU acquires the three interference power thresholds, e.g., if the PU feeds back these limits or they are provided by a central node [28].

In these examples, we observe that Strategy 2 outperforms Strategy 1 for small  $\alpha$ , but the former provides lower benefits as  $\alpha$  increases and is even outperformed by PGS when  $\alpha$  is high. While this observation applies to both Figs. 3 and 4, there are also some differences. When the number of antennas increases, the crossing point between both IGS strategies occurs at a slightly higher  $\alpha$ , while the crossing point between Strategy 2 and PGS shifts to the left. Furthermore, the relative rate increase is smaller for 4 than for 2 antennas. It is also worth pointing out that the adaptive scheme presents an interesting approach, as little additional feedback is required, but it permits dynamically choosing the optimal scheme depending on the current system parameters. Notice

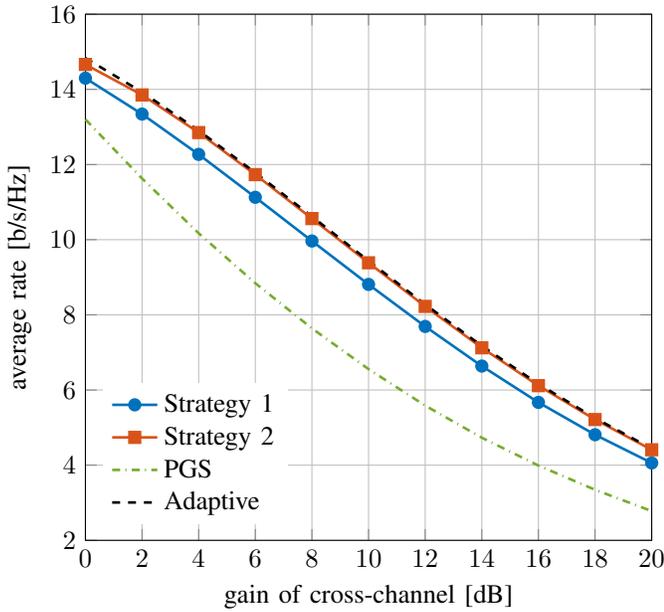


Fig. 6. Average achievable rate for the different techniques as a function of the gain of the cross-channel,  $\mathbf{H}_{ps}$ . The number of antennas for both users is  $N = M = 4$ .

that the adaptive scheme chooses the best strategy for each channel realization, which means that it dynamically selects the best strategy according to the instantaneous, rather than average, performance. This means that it may perform better than the maximum of the two strategies, as observed in Figs. 3 and 4.

For the next examples we fix  $\alpha = 0.6$  and the SNR of both users to 20 dB. We evaluate the proposed schemes for different gains of the cross-channel,  $\mathbf{H}_{ps}$ , measured as the variance of its coefficients. A higher gain represents a scenario where the impact of the interference is more significant. The average achievable rates are shown in Figs. 5 and 6 for  $N = M = 2$  and  $N = M = 4$ , respectively. We observed a significant performance gain of IGS over PGS in both cases, specially at intermediate cross-channel gains, which reflects again the effectiveness of IGS to reduce the impact of interference. In these examples, Strategy 2 outperforms Strategy 1, and the gap between both increases with the number of antennas.

## VII. CONCLUSION

In this paper we have considered a UCR scenario where users are equipped with multiple antennas. We have analyzed how to combine interference power constraints with additional constraints on the improper structure of the interference, such that an instantaneous rate of the PU is guaranteed. We have proposed two strategies that force the interference to be maximally improper, but which consider different restrictions on the improper correlation structure. Additionally, we have proposed algorithms to optimize the transmission scheme of the SU subject to these constraints. We have shown that the proposed strategies may significantly increase the performance of the SU compared to conventional PGS strategies. Furthermore, the advantages of the different transmission schemes can be

leveraged by an adaptive scheme with little feedback increase, which dynamically chooses the best scheme.

## APPENDIX

### A. Proof of Lemma 1

Let  $\mathbf{S}_p = \mathbf{H}_{pp}\mathbf{Q}_p\mathbf{H}_{pp}^H$  and  $\underline{\mathbf{S}}_p$  be the corresponding augmented covariance matrix. Since the interference is constrained to be maximally improper, the interference augmented covariance matrix  $\underline{\mathbf{K}}$  is half-rank, i.e.,  $\text{rank}(\underline{\mathbf{K}}) = N$ . Thus, the left-hand side of (22) is lower-bounded as

$$\begin{aligned} \min_{\substack{\mathbf{K} \in \mathbb{K}(t) \\ \mathbf{F}_p \in \mathbb{U}^N}} R_p(\mathbf{K}, \mathbf{F}_p) &\geq \min_{\substack{\underline{\mathbf{K}} \succeq \mathbf{0} \\ \text{rank}(\underline{\mathbf{K}}) = N \\ \text{Tr}(\underline{\mathbf{K}}) \leq 2t}} R_p(\underline{\mathbf{K}}) \\ &= \min_{\substack{\underline{\mathbf{K}} \succeq \mathbf{0} \\ \text{rank}(\underline{\mathbf{K}}) = N \\ \text{Tr}(\underline{\mathbf{K}}) \leq 2t}} \frac{1}{2} \log_2 |\sigma^2 \mathbf{I} + \underline{\mathbf{K}} + \mathbf{S}_p| - \frac{1}{2} \log_2 |\sigma^2 \mathbf{I} + \underline{\mathbf{K}}|, \end{aligned} \quad (54)$$

where, with a slight abuse of notation, we use  $R_p(\underline{\mathbf{K}})$  to indicate the PU rate when the interference augmented covariance matrix is  $\underline{\mathbf{K}}$ . Now consider the following property [39, 9.G.2.a]

$$|\mathbf{Y} + \mathbf{Z}| \geq \prod_{i=1}^N [\lambda_i(\mathbf{Y}) + \lambda_i(\mathbf{Z})], \quad (55)$$

for Hermitian matrices  $\mathbf{Y}$  and  $\mathbf{Z}$ , where  $\lambda_i(\cdot)$  denotes the  $i$ th largest eigenvalue. Combining (55) with (54) we obtain

$$\begin{aligned} \min_{\substack{\mathbf{K} \in \mathbb{K}(t) \\ \mathbf{F}_p \in \mathbb{U}^N}} R_p(\mathbf{K}, \mathbf{F}_p) &\geq \min_{\substack{\underline{\mathbf{K}} \succeq \mathbf{0} \\ \text{rank}(\underline{\mathbf{K}}) = N \\ \text{Tr}(\underline{\mathbf{K}}) \leq 2t}} \sum_{i=1}^{2N} \frac{1}{2} \log_2 (\sigma^2 + \lambda_i(\underline{\mathbf{K}}) + \lambda_i(\mathbf{S}_p)) \\ &\quad - \sum_{i=1}^{2N} \frac{1}{2} \log_2 (\sigma^2 + \lambda_i(\underline{\mathbf{K}})) \\ &= \min_{\substack{\underline{\mathbf{K}} \succeq \mathbf{0} \\ \text{rank}(\underline{\mathbf{K}}) = N \\ \text{Tr}(\underline{\mathbf{K}}) \leq 2t}} \sum_{i=1}^{2N} \frac{1}{2} \log_2 \left( 1 + \frac{\lambda_i(\mathbf{S}_p)}{\sigma^2 + \lambda_i(\underline{\mathbf{K}})} \right) \\ &= \min_{\substack{\theta_i \geq 0 \\ \sum_{i=1}^N \theta_i \leq 2t}} \sum_{i=1}^N \frac{1}{2} \log_2 \left( 1 + \frac{\lambda_i(\mathbf{S}_p)}{\sigma^2 + \lambda_i(\underline{\mathbf{K}})} \right) \\ &\quad + \sum_{j=N+1}^{2N} \frac{1}{2} \log_2 \left( 1 + \frac{\lambda_j(\mathbf{S}_p)}{\sigma^2} \right). \end{aligned} \quad (56)$$

Since the signal from the PU is proper, each eigenvalue of  $\underline{\mathbf{S}}_p$  has multiplicity two, i.e.,

$$\lambda(\underline{\mathbf{S}}_p) = [\phi_1, \phi_1, \phi_2, \phi_2, \dots, \phi_{N-1}, \phi_{N-1}, \phi_N, \phi_N]^T, \quad (57)$$

which implies that the eigenvalues of  $\underline{\mathbf{K}}$  that minimize the last lower bound in (56) are also repeated. Therefore, the last

lower bound in (56) can be simplified to

$$\begin{aligned} \min_{\substack{\mathbf{K} \in \mathbb{K}(t) \\ \mathbf{F}_p \in \mathbb{U}^N}} R_p(\mathbf{K}, \mathbf{F}) &\geq \min_{\substack{\theta_i \geq 0 \\ \sum_{i=1}^{\frac{N}{2}} \theta_i \leq t}} \sum_{i=1}^{\frac{N}{2}} \log_2 \left( 1 + \frac{\phi_i}{\sigma^2 + \theta_i} \right) \\ &+ \sum_{j=\frac{N}{2}+1}^N \log_2 \left( 1 + \frac{\phi_j}{\sigma^2} \right). \end{aligned} \quad (58)$$

Now we show that this lower bound is achievable as follows. Making the eigenvectors of  $\mathbf{K}$  equal to those of  $\mathbf{H}_{pp} \mathbf{Q}_p \mathbf{H}_{pp}^H$ ,  $\mathbf{U}$ , and taking  $\mathbf{F}_p = \mathbf{U} \mathbf{J}^{1/2}$ , the PU rate can be expressed as

$$\begin{aligned} R_p(\mathbf{U} \mathbf{\Theta} \mathbf{U}^H, \mathbf{U} \mathbf{J}^{1/2}) &= \frac{1}{2} \sum_{i=1}^N \log_2 \left[ 1 + \frac{\sigma^2 (\phi_i + \phi_{N-i+1}) + \phi_i \phi_{N-i+1} + \theta_{N-i+1} \phi_i + \theta_i \phi_{N-i+1}}{\sigma^2 (\sigma^2 + \theta_i + \theta_{N-i+1})} \right]. \end{aligned} \quad (59)$$

Taking  $\theta_i = 0$  for  $i = \frac{N}{2} + 1, \dots, N$ ,  $R_p(\mathbf{U} \mathbf{\Theta} \mathbf{U}^H, \mathbf{U} \mathbf{J}^{1/2})$  equals the lower bound in (58).

### B. Proof of Lemma 2

Let  $\mathbf{H}_{pp} \mathbf{Q}_p \mathbf{H}_{pp}^H = \mathbf{V}_p \mathbf{\Phi} \mathbf{V}_p^H$  be the SVD. The left-hand side of (28) can be lower bounded by

$$\min_{\mathbf{K} \in \mathbb{K}(t)} \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) \geq \min_{\mathbf{K} \in \mathbb{K}(t)} R_p(\mathbf{K}, \mathbf{U}), \quad (60)$$

where  $\mathbf{U}$  is the matrix of eigenvectors of  $\mathbf{K}$ . In this case, we can express the rate as

$$\begin{aligned} R_p(\mathbf{K}, \mathbf{U}) &= \log_2 \left| \sigma^2 \mathbf{I} + \mathbf{\Theta} + \mathbf{\Upsilon} \mathbf{\Phi} \mathbf{\Upsilon}^H \right| - \log_2 \left| \sigma^2 \mathbf{I} + \mathbf{\Theta} \right| \\ &+ \frac{1}{2} \log_2 \left| \mathbf{I} - \mathbf{\Theta}^{1/2} \left( \sigma^2 \mathbf{I} + \mathbf{\Theta} + \mathbf{\Upsilon}^* \mathbf{\Phi} \mathbf{\Upsilon}^T \right)^{-1} \mathbf{\Theta} \right. \\ &\times \left. \left( \sigma^2 \mathbf{I} + \mathbf{\Theta} + \mathbf{\Upsilon} \mathbf{\Phi} \mathbf{\Upsilon}^H \right)^{-1} \mathbf{\Theta}^{1/2} \right| \\ &- \frac{1}{2} \log_2 \left| \mathbf{I} - \left( \sigma^2 \mathbf{I} + \mathbf{\Theta} \right)^{-2} \mathbf{\Theta}^2 \right|, \end{aligned} \quad (61)$$

where  $\mathbf{\Upsilon} = \mathbf{U}^H \mathbf{V}_p$  and  $\mathbf{\Theta}$  is a diagonal matrix containing the eigenvalues of  $\mathbf{K}$ . In the following, we will use  $\prec$ ,  $\prec_{\log}$ ,  $\prec_w$ , and  $\prec^w$  to denote majorization, log-majorization, weak-submajorization, and weak-supermajorization, respectively. We refer the reader to [39] for the definitions.

The third term in (61) can be lower-bounded as follows. Using [39, 9.H.1.a] we have  $\lambda(\mathbf{Y} \mathbf{Z}) \prec_{\log} \lambda(\mathbf{Y}) \odot \lambda(\mathbf{Z})$  for positive semidefinite Hermitian matrices  $\mathbf{Y}$  and  $\mathbf{Z}$ , which implies, by [39, 5.A.2.b],  $\lambda(\mathbf{Y} \mathbf{Z}) \prec_w \lambda(\mathbf{Y}) \odot \lambda(\mathbf{Z})$ , where  $\odot$  denotes the Hadamard or element-wise product. Furthermore, using [39, Eq. 1.A.13b] we have  $\mathbf{1} - \lambda(\mathbf{Y} \mathbf{Z}) \prec^w \mathbf{1} - \lambda(\mathbf{Y}) \odot \lambda(\mathbf{Z})$ . By [39, 3.A.7], we then have

$$\log_2 |\mathbf{I} - \mathbf{Y} \mathbf{Z}| \geq \sum_{i=1}^N \log_2 [1 - \lambda_i(\mathbf{Y}) \lambda_i(\mathbf{Z})]. \quad (62)$$

Applying the above inequality to the third term in (61) we obtain the lower bound

$$\begin{aligned} &\log_2 \left| \mathbf{I} - \mathbf{\Theta}^{1/2} \left( \sigma^2 \mathbf{I} + \mathbf{\Theta} + \mathbf{\Upsilon}^* \mathbf{\Phi} \mathbf{\Upsilon}^T \right)^{-1} \mathbf{\Theta} \right. \\ &\times \left. \left( \sigma^2 \mathbf{I} + \mathbf{\Theta} + \mathbf{\Upsilon} \mathbf{\Phi} \mathbf{\Upsilon}^H \right)^{-1} \mathbf{\Theta}^{1/2} \right| \\ &\geq \sum_{i=1}^N \log_2 \left[ 1 - \lambda_i \left( \mathbf{\Theta} \left( \sigma^2 \mathbf{I} + \mathbf{\Theta} + \mathbf{\Upsilon} \mathbf{\Phi} \mathbf{\Upsilon}^H \right)^{-1} \right)^2 \right]. \end{aligned} \quad (63)$$

Plugging (63) into (61) we can obtain, after some manipulations,

$$\begin{aligned} R_p(\mathbf{K}, \mathbf{U}) &\geq \frac{1}{2} \log_2 \left| \sigma^2 \mathbf{I} + 2\mathbf{\Theta} + \mathbf{\Upsilon} \mathbf{\Phi} \mathbf{\Upsilon}^H \right| \\ &+ \frac{1}{2} \log_2 \left| \sigma^2 \mathbf{I} + \mathbf{\Phi} \right| - \log_2 \left| \sigma^2 \mathbf{I} + \mathbf{\Theta} \right| \\ &- \frac{1}{2} \log_2 \left| \mathbf{I} - \left( \sigma^2 \mathbf{I} + \mathbf{\Theta} \right)^{-2} \mathbf{\Theta}^2 \right|. \end{aligned} \quad (64)$$

The first term in (64) can be lower bounded using (55) as  $\frac{1}{2} \log_2 \left| \sigma^2 \mathbf{I} + 2\mathbf{\Theta} + \mathbf{\Upsilon} \mathbf{\Phi} \mathbf{\Upsilon}^H \right| \geq \frac{1}{2} \log_2 \left| \sigma^2 \mathbf{I} + 2\mathbf{\Theta} + \mathbf{\Phi} \right|$ , which yields

$$\begin{aligned} R_p(\mathbf{K}, \mathbf{U}) &\geq \frac{1}{2} \log_2 \left| \sigma^2 \mathbf{I} + 2\mathbf{\Theta} + \mathbf{\Phi} \right| + \frac{1}{2} \log_2 \left| \sigma^2 \mathbf{I} + \mathbf{\Phi} \right| \\ &- \log_2 \left| \sigma^2 \mathbf{I} + \mathbf{\Theta} \right| - \frac{1}{2} \log_2 \left| \mathbf{I} - \left( \sigma^2 \mathbf{I} + \mathbf{\Theta} \right)^{-2} \mathbf{\Theta}^2 \right| \\ &= \frac{1}{2} \log_2 \left| \mathbf{I} + \left( \sigma^2 \mathbf{I} + 2\mathbf{\Theta} \right)^{-1} \mathbf{\Phi} \right| + \frac{1}{2} \log_2 \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{\Phi} \right|. \end{aligned} \quad (65)$$

Notice that the above lower bound can be achieved by setting  $\mathbf{\Upsilon} = \mathbf{I}$ , i.e.,  $\mathbf{U} = \mathbf{V}_p$ . Therefore, (60) can be equivalently expressed as

$$\min_{\mathbf{K} \in \mathbb{K}(t)} \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) \geq \min_{\substack{\theta_i \leq t \\ \theta_i \geq 0}} R(\mathbf{V}_p \mathbf{\Theta} \mathbf{V}_p^H, \mathbf{V}_p). \quad (66)$$

Since  $R(\mathbf{V}_p \mathbf{\Theta} \mathbf{V}_p^H, \mathbf{V}_p)$  is convex in  $\theta_i$ , and the feasible set is non-empty (Slater's condition), the optimal solution can be found through the KKT conditions similarly as in [32], which yields

$$\theta_i^* = \frac{1}{2} \left[ \sqrt{\phi_i \left( \frac{1}{4} \phi_i + \mu \right)} - \left( \frac{1}{2} \phi_i + \sigma^2 \right) \right]^+, \quad (67)$$

with  $\mu$  such that  $\text{Tr}(\mathbf{\Theta}^*) = t$ , with  $\mathbf{\Theta}^* = \text{diag}(\theta_1^*, \dots, \theta_N^*)$ . Let us now consider the upper bound

$$\begin{aligned} \min_{\mathbf{K} \in \mathbb{K}(t)} \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{K}, \mathbf{F}_p) &\leq \max_{\mathbf{F}_p \in \mathbb{U}^N} R_p(\mathbf{V}_p \mathbf{\Theta}^* \mathbf{V}_p^H, \mathbf{F}_p) = \\ &\max_{\substack{\mathbf{W} \in \mathbb{U}^N \\ \mathbf{W} = \mathbf{W}^T}} \log_2 \left| \mathbf{I} + \begin{bmatrix} \sigma^2 \mathbf{I} + \mathbf{\Theta}^* & (\mathbf{\Theta}^*)^{1/2} \mathbf{W}^H (\mathbf{\Theta}^*)^{1/2} \\ (\mathbf{\Theta}^*)^{1/2} \mathbf{W} (\mathbf{\Theta}^*)^{1/2} & \sigma^2 \mathbf{I} + \mathbf{\Theta}^* \end{bmatrix}^{-1} \right. \\ &\times \left. \begin{bmatrix} \mathbf{\Phi} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi} \end{bmatrix} \right|, \end{aligned} \quad (68)$$

where  $\mathbf{W} = \mathbf{V}_p^T \mathbf{F}_p^* \mathbf{F}_p^H \mathbf{V}_p$  and  $\mathbf{\Theta}^*$  is the optimal solution to the right-hand side of (66), which is given by (67). Now we will show that the upper bound in (68) is maximized by

$\mathbf{W} = \mathbf{I}$  for  $t \geq \bar{t}$ . To this end, we may express (we henceforth express  $\Theta^*$  as  $\Theta$  to avoid cumbersome notation)

$$\begin{aligned} R_p(\mathbf{V}_p \Theta \mathbf{V}_p^H, \mathbf{F}_p) &= \log_2 \left| \mathbf{I} + (\sigma^2 \mathbf{I} + \Theta)^{-1} \Phi \right| \\ &+ \frac{1}{2} \log_2 \left| \mathbf{I} - \Theta^{1/2} (\sigma^2 \mathbf{I} + \Theta + \Phi)^{-1} \Theta^{1/2} \mathbf{W} \Theta^{1/2} \right. \\ &\times (\sigma^2 \mathbf{I} + \Theta + \Phi)^{-1} \Theta^{1/2} \mathbf{W}^H \left. - \frac{1}{2} \log_2 \left| \mathbf{I} - \Theta^{1/2} \right. \right. \\ &\times (\sigma^2 \mathbf{I} + \Theta)^{-1} \Theta^{1/2} \mathbf{W} \Theta^{1/2} (\sigma^2 \mathbf{I} + \Theta)^{-1} \Theta^{1/2} \mathbf{W}^H \left. \right|. \end{aligned} \quad (69)$$

The above expression is maximized for  $\mathbf{W} = \mathbf{I}$  (and therefore  $\mathbf{F}_p = \mathbf{V}_p$ ) if

$$\begin{aligned} \log_2 \left| \mathbf{I} - \mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H \right| - \log_2 \left| \mathbf{I} - \mathbf{D}_2 \mathbf{W} \mathbf{D}_2 \mathbf{W}^H \right| \\ \leq \log_2 \left| \mathbf{I} - \mathbf{D}_1^2 \right| - \log_2 \left| \mathbf{I} - \mathbf{D}_2^2 \right|, \end{aligned} \quad (70)$$

or, equivalently,

$$\begin{aligned} \log_2 \left| \mathbf{I} - \mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H \right| - \log_2 \left| \mathbf{I} - \mathbf{D}_1^2 \right| \\ \leq \log_2 \left| \mathbf{I} - \mathbf{D}_2 \mathbf{W} \mathbf{D}_2 \mathbf{W}^H \right| - \log_2 \left| \mathbf{I} - \mathbf{D}_2^2 \right|, \end{aligned} \quad (71)$$

where  $\mathbf{D}_1 = \Theta(\sigma^2 \mathbf{I} + \Theta + \Phi)^{-1}$  and  $\mathbf{D}_2 = \Theta(\sigma^2 \mathbf{I} + \Theta)^{-1}$ . Recall that  $\Theta$  is now a function of  $\Phi$  and  $\mu$  following (67). Equivalently, we may use (67) to express  $\Phi$  in terms of  $\Theta$  and  $\mu$  as

$$\phi_i(\mu) = \frac{(\sigma^2 + 2\theta_i)^2}{\mu - (\sigma^2 + 2\theta_i)}, \quad (72)$$

which makes  $\mathbf{D}_1$  a function of  $\mu$  as well, but not  $\mathbf{D}_2$ . Hence we may write  $\mathbf{D}_1(\mu)$ . It turns out that  $\phi_i(\infty) = 0 \forall i$ , so that  $\mathbf{D}_1(\infty) = \mathbf{D}_2$ . Therefore, in order to show that (71) holds for any symmetric and unitary matrix  $\mathbf{W}$ , we have to show

$$\begin{aligned} \arg \max_{\mu > \mu_{\min}} \{f(\mu) = \log_2 \left| \mathbf{I} - \mathbf{D}_1(\mu) \mathbf{W} \mathbf{D}_1(\mu) \mathbf{W}^H \right| \\ - \log_2 \left| \mathbf{I} - \mathbf{D}_1(\mu)^2 \right|\} = \infty, \end{aligned} \quad (73)$$

with  $\mu_{\min} = \sigma^2 + 2\theta_1$ , so that  $0 \leq \phi_i(\mu) < \infty$  for  $\mu > \mu_{\min}$ . Let us first consider the derivative of  $f(\mu)$ , which is (we drop the dependence of  $\mathbf{D}_1$  on  $\mu$  to simplify notation)

$$\begin{aligned} \frac{\partial f(\mu)}{\partial \mu} &= \text{Tr} \left[ 2 (\mathbf{I} - \mathbf{D}_1^2)^{-1} \mathbf{D}_1 \frac{\partial \mathbf{D}_1}{\partial \mu} - (\mathbf{I} - \mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H)^{-1} \right. \\ &\times \left. \left( \frac{\partial \mathbf{D}_1}{\partial \mu} \mathbf{W} \mathbf{D}_1 \mathbf{W}^H + \mathbf{D}_1 \mathbf{W} \frac{\partial \mathbf{D}_1}{\partial \mu} \mathbf{W}^H \right) \right], \end{aligned} \quad (74)$$

where

$$\left[ \frac{\partial \mathbf{D}_1}{\partial \mu} \right]_{ii} = \frac{\theta_i(\sigma^2 + 2\theta_i)^2}{[(\sigma^2 + \theta_i)(\mu - \sigma^2 - 2\theta_i) + (\sigma^2 + 2\theta_i)^2]^2}. \quad (75)$$

Consider the upper bound

$$\begin{aligned} \text{Tr} \left[ (\mathbf{I} - \mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H)^{-1} \frac{\partial \mathbf{D}_1}{\partial \mu} \mathbf{W} \mathbf{D}_1 \mathbf{W}^H \right] \\ = \sum_{i=1}^N \lambda_i \left[ (\mathbf{I} - \mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H)^{-1} \frac{\partial \mathbf{D}_1}{\partial \mu} \mathbf{W} \mathbf{D}_1 \mathbf{W}^H \right] \\ \leq \sum_{i=1}^N \lambda_i \left[ (\mathbf{I} - \mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H)^{-1} \right] \odot \lambda_i \left( \frac{\partial \mathbf{D}_1}{\partial \mu} \mathbf{W} \mathbf{D}_1 \mathbf{W}^H \right) \end{aligned} \quad (76)$$

$$\leq \sum_{i=1}^N \lambda_i \left[ (\mathbf{I} - \mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H)^{-1} \right] \odot \lambda_i \left( \frac{\partial \mathbf{D}_1}{\partial \mu} \right) \odot \lambda_i (\mathbf{D}_1) \quad (77)$$

$$\leq \sum_{i=1}^N \lambda_i \left[ (\mathbf{I} - \mathbf{D}_1^2)^{-1} \mathbf{D}_1 \right] \odot \lambda_i \left( \frac{\partial \mathbf{D}_1}{\partial \mu} \right), \quad (78)$$

where (76) is due to  $\lambda(\mathbf{A}\mathbf{B}) \prec_w \lambda(\mathbf{A}) \odot \lambda(\mathbf{B})$ , and (77) is obtained by using  $\mathbf{x} \prec_w \mathbf{y} \Rightarrow \mathbf{x} \odot \mathbf{u} \prec_w \mathbf{y} \odot \mathbf{u}$ , for  $x_1 \geq \dots \geq x_N$ ,  $y_1 \geq \dots \geq y_N$ , and  $u_1 \geq \dots \geq u_N \geq 0$  [39, 3.H.3.b]. Finally, (78) is obtained as follows. First we notice

$$\lambda_i \left[ (\mathbf{I} - \mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H)^{-1} \right] = \frac{1}{1 - \lambda_i(\mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H)}. \quad (79)$$

The second derivative of  $\frac{1}{1-x}$  is  $\frac{2(1-x)}{(1-x)^3} > 0$  for  $0 \leq x < 1$ , hence it is convex. Additionally, it is increasing in  $x$ , which implies that  $\sum_{i=1}^N \frac{1}{1-x_i}$  is increasing and Schur-convex [39, 3.B.2]. This, along with  $\lambda(\mathbf{D}_1 \mathbf{W} \mathbf{D}_1 \mathbf{W}^H) \prec_w \lambda(\mathbf{D}_1^2)$ , yields (78). Combining (78) with (74) we have

$$\begin{aligned} \frac{\partial f(\mu)}{\partial \mu} &\geq 2 \sum_{i=1}^N \left\{ \lambda_i \left[ (\mathbf{I} - \mathbf{D}_1^2)^{-1} \mathbf{D}_1 \frac{\partial \mathbf{D}_1}{\partial \mu} \right] \right. \\ &\left. - \lambda_i \left[ (\mathbf{I} - \mathbf{D}_1^2)^{-1} \mathbf{D}_1 \right] \odot \lambda_i \left( \frac{\partial \mathbf{D}_1}{\partial \mu} \right) \right\}. \end{aligned} \quad (80)$$

The above lower bound is equal to zero if the elements of the diagonal matrices  $\mathbf{D}_1$  and  $\frac{\partial \mathbf{D}_1}{\partial \mu}$  are similarly ordered. The elements in  $\frac{\partial \mathbf{D}_1}{\partial \mu}$  are in decreasing order if

$$\frac{\partial^2 \mathbf{D}_1}{\partial \mu \partial \theta_i} \geq 0 \forall i \Rightarrow \mu \geq \mu_0 = \frac{\theta_1(\sigma^2 + 2\theta_1)^2}{\sigma^4 + 2\theta_1^2 + 5\theta_1\sigma^2}. \quad (81)$$

Notice that  $\mu_{\min} > \mu_0$ , which means that the elements in  $\frac{\partial \mathbf{D}_1}{\partial \mu}$  are always arranged in decreasing order. Therefore, the elements in  $\mathbf{D}_1$  are in decreasing order for  $\mu \geq \bar{\mu}$ , with  $\bar{\mu}$  such that

$$\begin{aligned} \frac{\theta_1}{\sigma^2 + \theta_1 + \phi_1(\bar{\mu})} &= \frac{\theta_2}{\sigma^2 + \theta_2 + \phi_2(\bar{\mu})} \\ \Rightarrow \bar{\mu} &= \frac{1}{\sigma^2} (\sigma^2 + 2\theta_1) (\sigma^2 + 2\theta_2). \end{aligned} \quad (82)$$

Thus we can state that  $f(\mu)$  is non-decreasing in the interval  $\mu \geq \bar{\mu}$ , which implies

$$\arg \max_{\mu \geq \bar{\mu}} f(\mu) = \infty. \quad (83)$$

Since the lower and upper bounds coincide for  $\mu \geq \bar{\mu}$ , we obtain (28)–(30). For  $\mu < \bar{\mu}$  we cannot show that the lower and upper bounds are equal, yielding (31) and concluding the proof.

## REFERENCES

- [1] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals*, Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [2] S. Javed, O. Amin, S. S. Ikki, and M. S. Alouini, "Asymmetric hardware distortions in receive diversity systems: Outage performance analysis," *IEEE Access*, vol. 5, pp. 4492–4504, Feb. 2017.
- [3] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, pp. 5101–5125, Nov. 2011.
- [4] G. Tauböck, "Complex-valued random vectors and channels: Entropy, divergence, and capacity," *IEEE Trans. Inf. Theory*, vol. 58, pp. 2729–2744, May 2012.
- [5] W. H. Gerstacker, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. Commun.*, vol. 51, pp. 1512–1523, Sept. 2003.
- [6] S. Buzzi, M. Lops, and S. Sardellitti, "Widely linear reception strategies for layered space-time wireless communications," *IEEE Trans. Signal Process.*, vol. 54, pp. 2252–2262, June 2006.
- [7] P. Chevalier and F. Picon, "New insights into optimal widely linear array receivers for the demodulation of BPSK, MSK, and GMSK signals corrupted by noncircular interferences—application to SAIC," *IEEE Trans. Signal Process.*, vol. 54, pp. 870–883, Mar. 2006.
- [8] R. Schober, W. H. Gerstacker, and L. H.-J. Lampe, "Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA," *IEEE Trans. Signal Process.*, vol. 52, pp. 746–756, Mar. 2004.
- [9] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, pp. 1293–1302, July 1993.
- [10] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*, Wiley-Interscience, 2006.
- [11] 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, pp. 4552–4565, Sept. 2010.
- [12] Z. K. M. Ho and E. Jorswieck, "Improper Gaussian signaling on the two-user SISO interference channel," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3194–3203, Sept. 2012.
- [13] 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, pp. 2899–2913, June 2013.
- [14] S. Lagen, Adrian A., and J. Vidal, "Coexisting linear and widely linear transceivers in the MIMO interference channel," *IEEE Trans. Signal Process.*, vol. 64, pp. 652–664, Feb. 2016.
- [15] C. Lameiro, I. Santamaría, and P. J. Schreier, "Rate region boundary of the SISO Z-interference channel with improper signaling," *IEEE Trans. Commun.*, vol. 65, pp. 1022–1034, Mar. 2017.
- [16] H. D. Nguyen, R. Zhang, and S. Sun, "Improper signaling for symbol error rate minimization in K-user interference channel," *IEEE Trans. Commun.*, vol. 63, pp. 857–869, Mar. 2015.
- [17] M. Gaafar, M. G. Khafagy, O. Amin, and M. S. Alouini, "Improper Gaussian signaling in full-duplex relay channels with residual self-interference," in *Proc. IEEE Int. Conf. Commun.*, May 2016, pp. 1–7.
- [18] M. Gaafar, O. Amin, A. Ikhlef, A. Chaaban, and M. S. Alouini, "On alternate relaying with improper Gaussian signaling," *IEEE Commun. Lett.*, vol. 20, pp. 1683–1686, Aug. 2016.
- [19] C. Lameiro, I. Santamaría, and P. J. Schreier, "Benefits of improper signaling for underlay cognitive radio," *IEEE Wireless Comm. Lett.*, vol. 4, pp. 22–25, Feb. 2015.
- [20] C. Lameiro, I. Santamaría, and P. J. Schreier, "Analysis of maximally improper signalling schemes for underlay cognitive radio," in *Proc. IEEE Int. Conf. Commun.*, June 2015.
- [21] O. Amin, W. Abediseid, and M. S. Alouini, "Underlay cognitive radio systems with improper Gaussian signaling: Outage performance analysis," *IEEE Trans. Wireless Commun.*, vol. 15, pp. 4875–4887, July 2016.
- [22] M. Gaafar, O. Amin, W. Abediseid, and M. S. Alouini, "Underlay spectrum sharing techniques with in-band full-duplex systems using improper Gaussian signaling," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 235–249, Jan. 2017.
- [23] O. Amin, W. Abediseid, and M. S. Alouini, "Overlay spectrum sharing using improper Gaussian signaling," *IEEE J. Sel. Areas Commun.*, vol. 35, pp. 50–62, Jan. 2017.
- [24] W. Hedhly, O. Amin, and M. S. Alouini, "Interweave cognitive radio with improper Gaussian signaling," in *Proc. IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.
- [25] C. Hellings and W. Utschick, "Improper signaling versus time-sharing in the SISO Z-interference channel," *IEEE Commun. Lett.*, vol. 21, pp. 2432–2435, Nov. 2017.
- [26] C. Hellings and W. Utschick, "Improper signaling versus time-sharing in the two-user Gaussian interference channel with TIN," *arXiv:1808.01611 [cs.IT]*, Aug. 2018.
- [27] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, pp. 894–914, May 2009.
- [28] C. X. Wang, F. Haider, X. Gao, X. H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, pp. 122–130, Feb. 2014.
- [29] L. Zhang, Y. C. Liang, and Y. Xin, "Joint beamforming and power allocation for multiple access channels in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 38–51, Jan. 2008.
- [30] R. Zhang and Y. C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, pp. 88–102, Feb. 2008.
- [31] J. M. Moualeu, W. Hamouda, and F. Takawira, "Cognitive coded cooperation in underlay spectrum-sharing networks under interference power constraints," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 2099–2113, Mar. 2017.
- [32] C. Lameiro, W. Utschick, and I. Santamaría, "Interference-temperature limit for cognitive radio networks with MIMO primary users," in *Proc. Asilomar Conf. Signals Syst. Computers*, Nov. 2014.
- [33] C. Lameiro, W. Utschick, and I. Santamaría, "Spatial interference shaping for underlay MIMO cognitive networks," *Signal Process.*, vol. 134, pp. 174–184, 2017.
- [34] C. Lameiro, I. Santamaría, W. Utschick, and P. J. Schreier, "Maximally improper interference in underlay cognitive radio networks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2016.
- [35] C. Lameiro, I. Santamaría, and P. J. Schreier, "Performance analysis of maximally improper signaling for multiple-antenna systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2018.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, 2004.
- [37] H. Yu and V. K. N. Lau, "Rank-constrained Schur-convex optimization with multiple trace/log-det constraints," *IEEE Trans. Signal Process.*, vol. 59, pp. 304–314, Jan. 2011.
- [38] J. C. Bezdek and R. J. Hathaway, "Some notes on alternating optimization," in *Advances in Soft Computing — AFSS 2002*, Nikhil R. Pal and Michio Sugeno, Eds., Berlin, Heidelberg, 2002, pp. 288–300, Springer Berlin Heidelberg.
- [39] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, Springer Series in Statistics. Springer New York, New York, NY, 2011.