RESEARCH
Open Access

Multi-antenna transmission for underlay and overlay cognitive radio with explicit message-learning phase

Ricardo Blasco-Serrano[1*], Jing Lv[2], Ragnar Thobaben[1], Eduard Jorswieck[2] and Mikael Skoglund[1]

Abstract

We consider the coexistence of a multiple-input multiple-output secondary system with a multiple-input single-output primary link with different degrees of coordination between the systems. First, for the uncoordinated underlay cognitive radio scenario, we fully characterize the optimal parameters that maximize the secondary rate subject to a primary rate constraint for a transmission strategy that combines rate splitting and interference cancellation. Second, we establish a model for the coordinated overlay cognitive radio scenario that consists of a message-learning phase followed by a communication phase. We then propose a transmission strategy that combines techniques for cooperative communication and for the classical cognitive radio channel. We optimize our system to maximize the rate of communication for the secondary users under a primary-user rate constraint and find efficient algorithms to compute the optimal system parameters. Finally, we compare both cognitive radio strategies to assess their relative merits and to evaluate the effect of the message-learning phase. We observe that for closely located transmitters, the overlay strategy outperforms the underlay strategy. In this situation, learning the primary message is very beneficial for the secondary systems, especially if they are interference-limited rather than power-limited. The situation is reversed when the distance between the transmitters is large. In either case, we observe that there is room for significant improvement if the transmitter implements both strategies and decides adaptively which one to use according to the channel conditions. We conclude our work with a discussion on the extension to the coexistence with multiple-input multiple-output primaries.

Keywords: Cognitive radio; Underlay; Overlay; Multiple antennas; Message learning; Cooperation

1 Introduction

The scarcity of available spectrum for accommodating new services in combination with the underutilization of currently allocated spectrum has fueled research on alternative visions on communications over the last decade. It has been suggested that new, unlicensed (i.e. secondary) users could utilize portions of the spectrum licensed to primary users as long as the latter are not significantly affected. In this context, the concept of cognitive radio, with its promise of reconfigurability and adaptability to varying conditions, has emerged as a strong candidate for

implementing communication systems that make a more efficient use of the spectrum.

Three major cognitive radio paradigms that consider different degrees of interaction between primary and secondary users have been identified: interweave, underlay and overlay [1,2]. Interweave cognitive radio is conceptually the simplest one: the secondary devices sense the environment to detect the presence of primary users and transmit opportunistically only when these are silent. Underlay cognitive radio goes one step further and permits communication between secondary users as long as the disturbance created to the primary system is below some predefined threshold. Clearly, in this case, the secondary terminals need not only assess whether primary users are transmitting or not but also how much interference they will create and whether this will disrupt the primary communication. Finally, the overlay paradigm

*Correspondence: ricardo.blasco@ee.kth.se
[1] ACCESS Linnaeus Centre, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden
Full list of author information is available at the end of the article

allows for a tight interaction between primary and secondary systems. Of course, this comes not only at the price of a higher degree of sophistication of the secondary terminals but also requires flexibility in the primary system. Nevertheless, in all three cases, it is necessary to assess the impact of the presence of secondary users on primary systems. Several measures have been discussed in the literature for this purpose, for example, the probabilities of miss detection and interference for interweave cognitive radio or, more in general, soft- and peak-power-shaping interference temperature constraints [3,4]. An alternative is to consider directly the degradation suffered by the primary users, for example, in terms of the loss in rate [5].

Research on the physical layer has focused on establishing basic models for the different cognitive radio scenarios, deriving their fundamental limits, and designing practical transceivers that come close to these limits. From an information theoretic point of view, two channel models have been considered for the three cognitive radio paradigms: the Gaussian interference channel [6,7] and the cognitive radio channel [8-10]. As described before, in the cases of interweave and underlay cognitive radio, there is no cooperation between primary and secondary systems. This is precisely the situation described by the interference channel. The interweave cognitive radio paradigm corresponds to time sharing in the interference channel [6], with a sharing parameter that is fixed by the activity of the primary users. In this case, the challenge lies almost exclusively in sensing accurately the primary activity, a topic that lies outside the scope of this paper (see, e.g. [11] and references therein). Therefore, interweave cognitive radio scenarios will not be considered here. On the other hand, in the case of underlay cognitive radio, primary and secondary systems can transmit at the same time and thus the scenario is richer from the point of view of the communication strategies that can be used. This is well characterized by the interference channel if one places some additional restrictions on the model. For example, one usually restricts the communication strategies used by the primary user pairs to consist of point-to-point codes and single-user decoding.

In contrast, overlay cognitive radio scenarios are not described properly by the interference channel. The main reason for this is that the interference channel does not allow for any active cooperation between the user pairs. With the aim of overcoming this limitation, the cognitive radio channel was introduced in [9]. This model extends the interference channel by assuming that the secondary transmitter has non-causal knowledge of the primary message. This additional knowledge allows for asymmetric cooperation in the sense that the secondary transmitter can help the primary users to carry their

communication. In addition, it can combat the interference that the primary signal creates on the secondary receiver by means of interference cancellation or dirty-paper coding. This asymmetric cooperation was key for establishing the capacity of the cognitive radio channel with weak interference [8,9].

A usual system design criterion is to maximize the rate of transmission for the secondary users while ensuring a minimum quality of service (QoS) for the primary users. A key observation is that multiple transmit antenna techniques are a powerful and efficient way of controlling the disturbance created by the secondary users [12]. Unfortunately, the use of such techniques often leads to complex matrix optimization problems. This has motivated the use of tools from optimization theory for the design of transceivers. For example, convex optimization tools were used in [13] to study underlay cognitive radio models with single-user decoders. An underlay scenario with rate splitting and multiple-user decoding was considered in [14]. The problem of distributed beamforming and rate allocation in decentralized cognitive radio networks was treated in [15]. In a more general framework, the set of efficient strategies for multiple-input single-output (MISO) interference networks was characterized in [16,17] in terms of beamformers. The extension of the cognitive radio channel to the multiple-input multiple-output (MIMO) case was introduced in [18]. Overlay cognitive radio strategies for this channel with partial channel state information were considered in [19]. Optimal beamforming for the coexistence of a MIMO secondary user with a MISO primary user with non-causal knowledge of the primary message was considered in [20]. We studied the coexistence of a MISO secondary system with a single-input single-output primary system in [21] for different levels of channel state information, and considered linear precoding strategies in [22].

A comparison of the results for underlay and overlay cognitive radio channel models suggests that the additional knowledge of the primary message at the secondary transmitter in the cognitive radio channel leads to significantly higher achievable rates [21]. However, a critical point is how the secondary transmitter can acquire such knowledge in practice. Clearly, requiring the secondary transmitter to learn actively the primary message before communicating will lead to an inevitable loss in rate for the secondary users, especially under practical constraints such as half duplex communication. Some authors have motivated practical scenarios in which the primary message is obtained causally. For example, the secondary users may overhear a primary automatic repeat request (ARQ) session and use their resources during the repetition phases to help the

primaries finish their transmission earlier or to exploit the inefficiencies of the ARQ protocol [23,24]. Similarly, in [25], the secondary system acquires the primary message and uses it to help the primary system finish the transmission earlier and then use the channel during the idle period. However, these schemes do not fully exploit the possibilities of overlay cognitive radio, in particular the possibility of interaction between primary and secondary systems. The use cooperative communication techniques [26-28] as an enabling technology for cognitive radio networks was surveyed in [29]. They were considered in [30] for single-antenna overlay cognitive radio and evaluated in terms of outage probabilities. The optimal secondary power allocation and phase split in a two-phase spectrum sharing scenario was considered in [31]. In [32], the authors studied beamforming and power allocation for the coexistence of a primary single-input single-output (SISO) user with a secondary single-input multiple-output or MISO that acquired the message in a causal fashion. However, as opposed to the work presented here, their work focused only on the second phase of communication, without considering explicitly the first, learning phase. In [33], beamforming and power allocation were studied for a system, where the secondary users relay the primary signal in an amplify-and-forward fashion, and the performance of the proposed system was compared to an underlay cognitive radio scheme. The use of cooperative relaying mechanisms for spectrum sensing and secondary user transmission in cognitive radio systems was described in [34,35].

1.1 Contributions and outline

We study physical-layer aspects of cognitive radio communications in a scenario, where a MISO primary system coexists with a half-duplex MIMO secondary system. We consider two approaches: on one hand, an underlay cognitive radio model without any cooperation between primary and secondary systems. On the other hand, an overlay cognitive radio model that allows for causal cooperation between the systems. Our goal is to compare both strategies and assess the potential advantages of each of them under conditions that are more realistic than the original cognitive radio channel model in [8,9]. In particular, we require that the primary message be learned causally by the secondary system.

We emphasize that this paper deals with idealized models. In particular, the overlay scenario requires a high degree of cooperation between primary and secondary systems. Similarly, quite often, the terminals have access to larger portion of channel state information than in practical systems. In spite of this idealization, we have decided to take this approach to quantify the benefits of having coordinated primary and secondary system

(through the message-learning phase) in a quite general way, as compared to the more *ad hoc* approaches in [23-25]. Moreover, these systems are, at least in theory, implementable, unlike the less realistic scenarios where the secondaries have non-causal knowledge of the primary messages.

This paper extends our previous work on the coexistence of a SISO primary system with a MISO secondary link for underlay [14] and overlay systems [36] to the case of coexisting MISO primary and MIMO secondary systems. The addition of multiple antennas at the primary transmitter and secondary receiver results in a model that is richer and substantially more complex. In particular, for the overlay scenario, the new model allows not only for MIMO communication between secondary users but also for MIMO inter-transmitter communication. Moreover, this new channel configuration represents a departure from the interference network (e.g. [17]) as it also incorporates aspects from cooperative communications. Finally, note that the convex optimization framework developed in [13] for underlay cognitive radio is not directly applicable to the strategies presented here because they result in non-convex problems.

The main contributions of this paper refer to the coexistence of a MIMO secondary link with a MISO primary system. They are the following: First (Section 3), we consider an underlay strategy that includes rate splitting and interference decoding at the secondary and characterize completely the set of transmission parameters that maximize the secondary rate subject to a constraint on the primary rate. Second (Section 4), we establish a transmission strategy for cognitive radio communication over an extended channel model that consists of an initial learning phase, followed by a communication phase. This strategy combines elements from cooperative communications and communication over a non-causal cognitive radio channel that exploit the special properties of the extended cognitive radio channel model. In addition, we characterize the set of parameters that maximize the rate of the secondary users under a primary rate constraint and formulate simple algorithms to find such parameters. Third (Section 5), using a simple geometrical model, we evaluate numerically the performance of the strategies and compare them to establish the regions in which each of them outperforms the other. To our knowledge, this is one of the few studies that try to quantify the advantages of the information-theoretic cognitive radio channel models under realistic conditions (i.e. without assuming non-causal knowledge of the primary message). Finally (Section 6), we discuss the extension of all these contributions to MIMO-MIMO coexistence scenarios. The last part (Section 7) concludes our work. For clarity of exposition, we present the proofs of all the results in the 'Appendices' Section.

2 Preliminaries

2.1 Notation

Column vectors and matrices are represented in lower case and upper case boldface letters, respectively. $|\cdot|$ is the absolute value of a scalar or the determinant of a matrix, $\|\cdot\|$ is the Frobenius norm of a vector or matrix, and $(\cdot)^H$ stands for Hermitian transpose. The trace of a square matrix is denoted by $\mathrm{tr}\{\cdot\}$. $\Pi_X \triangleq X\left(X^H X\right)^{-1} X^H$ denotes the orthogonal projection operator onto the column space of X, and $\Pi_X^\perp \triangleq I - \Pi_X$, where I is the identity matrix, denotes the orthogonal projection operator onto the orthogonal complement of the column space of X. The notation $X \succeq 0$ denotes that the matrix X is positive semidefinite. All logarithms in this paper are taken to the base of 2, and all rates are expressed in bits.

2.2 System model

We consider a MISO primary system with $N_{T,1}$ transmit antennas that is willing to share its channel with a half-duplex MIMO secondary system with $N_{T,2}$ antennas at the transmitter and $N_{R,2}$ antennas at the receiver. Our goal is to compare basic communication strategies for underlay and overlay cognitive radio without assuming non-causal knowledge of the primary message at the secondary transmitter. For this purpose, we introduce the following two channel models.

2.2.1 Underlay cognitive radio

We use the Gaussian MIMO/MISO interference channel as a model to study the conflict between a primary and a secondary link in underlay cognitive radio. Each of the transmitters sends a signal that is observed by the intended receiver in the presence of interference (from the other transmitter) as well as white Gaussian noise. The tth received sample from the matched-filtered complex baseband model is

$$y_1(t) = h_{11}^H x_1(t) + h_{21}^H x_2(t) + n_1(t) \tag{1}$$

$$y_2(t) = H_{12}^H x_1(t) + H_{22}^H x_2(t) + n_2(t), \tag{2}$$

where $x_1(t)$ and $x_2(t)$ are the $N_{T,1} \times 1$ and $N_{T,2} \times 1$ signal vectors sent by the primary and secondary transmitters, respectively, h_{i1} is the $N_{T,i} \times 1$ vector of the channel gains from transmitter i $\in \{1,2\}$ to receiver 1, and H_{i2} is the

$N_{T,i} \times N_{R,2}$ matrix of channel gains from transmitter i $\in \{1,2\}$ to receiver 2. The scalar $y_1(t)$ and the vector $y_2(t)$ are the observations at the receivers, which are corrupted by the noise processes $n_1(t)$ and $n_2(t)$, respectively.

2.2.2 Overlay cognitive radio

Our model for communication with half-duplex devices in an overlay cognitive radio environment is illustrated in Figure 1 and consists of two phases. In the first phase, the primary transmitter broadcasts its message to both its intended receiver and the secondary transmitter. The tth received sample from the matched-filtered complex baseband model in this phase is

$$y_1^{(1)}(t) = h_{11}^H x_1^{(1)}(t) + n_1^{(1)}(t) \tag{3}$$
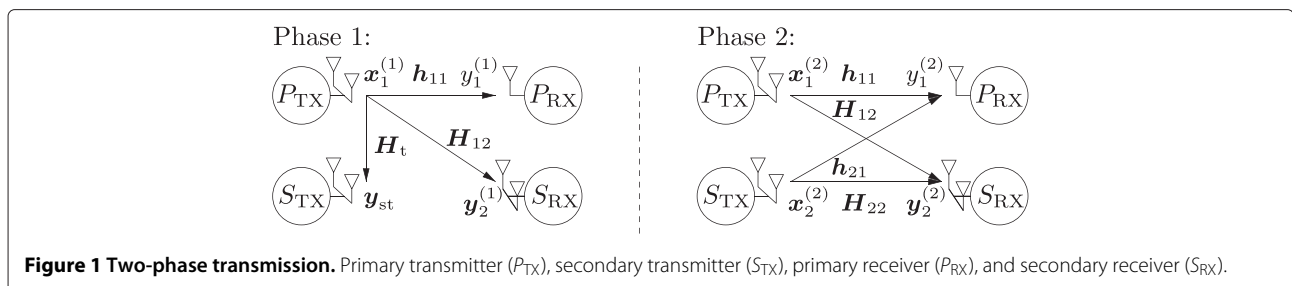
$$y_{st}(t) = H_t^H x_1^{(1)}(t) + n_{st}(t), \tag{4}$$

where $x_1^{(1)}(t)$ is the $N_{T,1} \times 1$ signal vector sent by the primary transmitter, h_{11} is the $N_{T,1} \times 1$ vector of channel coefficients between primary transmitter and receiver, and H_t is the $N_{T,1} \times N_{T,2}$ matrix of channel coefficients between both transmitters. The scalar $y_1^{(1)}(t)$ and the $N_{T,2} \times 1$ vector $y_{st}(t)$ are the observations at the primary receiver and secondary transmitter, respectively, which are corrupted by the noise processes $n_1^{(1)}(t)$ and $n_{st}(t)$, respectively. Note that, in principle, the secondary receiver can also obtain its own observation $y_2^{(1)}$ of the primary signal. However, as we shall see, this does not provide any gain for the transmission strategy proposed in Section 4.1.

The second phase corresponds to the set-up which is known as the cognitive radio channel. In this phase, the secondary transmitter can make use of the knowledge of the primary message (obtained in a causal fashion in the first phase). The model in this phase is

$$y_1^{(2)}(t) = h_{11}^H x_1^{(2)}(t) + h_{21}^H x_2^{(2)}(t) + n_1^{(2)}(t) \tag{5}$$

$$y_2^{(2)}(t) = H_{12}^H x_1^{(2)}(t) + H_{22}^H x_2^{(2)}(t) + n_2(t), \tag{6}$$

where $x_1^{(2)}(t)$ and $x_2^{(2)}(t)$ are the $N_{T,1} \times 1$ and $N_{T,2} \times 1$ signal vectors sent by the primary and secondary transmitters, respectively, h_{i1} is the $N_{T,i} \times 1$ vector of channel gains from transmitter i $\in \{1,2\}$ to receiver 1, and H_{i2} is the $N_{T,i} \times N_{R,2}$ matrix of channel gains from transmitter i $\in \{1,2\}$ to

Figure 1 Two-phase transmission. Primary transmitter (P_{TX}), secondary transmitter (S_{TX}), primary receiver (P_{RX}), and secondary receiver (S_{RX}).

receiver 2. The scalar $y_1^{(2)}(t)$ and the vector $\boldsymbol{y}_2^{(2)}(t)$ are the observations at the receivers, which are corrupted by the noise processes $n_1^{(2)}(t)$ and $\boldsymbol{n}_2(t)$, respectively.

The entire transmission is carried out over n channel uses; k channel uses are consumed during the first transmission phase, and $(n-k)$ channel uses during the second phase. The fraction of the channel uses in the first and the second phases is given by $\alpha = k/n$ and $1-\alpha$, respectively. We will assume that the channels remain constant during the duration of the two phases.

Noise and channel statistics For both underlay and overlay cognitive radio models, the noises at the receivers are modeled by independent circularly symmetric additive white Gaussian noise processes with unit variance: $n_1, n_1^{(1)}, n_1^{(2)} \sim \mathcal{CN}(0,1), \boldsymbol{n}_2, \boldsymbol{n}_{st} \sim \mathcal{CN}(0, \boldsymbol{I})$. In this paper, we assume that all nodes have perfect channel knowledge on all links. In order to evaluate the average behavior of our transmission strategies for different realizations of the channel coefficients, we will model the entries of $\boldsymbol{H}_t, \boldsymbol{h}_{11}, \boldsymbol{H}_{12}, \boldsymbol{h}_{21}$, and $\boldsymbol{H}_{22}$ as samples from independent circularly symmetric Gaussian processes with zero mean with appropriate variances.

3 Underlay cognitive radio
In this section, we introduce the transmission strategy that we consider for the underlay cognitive radio paradigm. Our goal is to maximize the communication rate of the secondary users while ensuring that the primary users have a minimum QoS, defined in terms of a minimum rate $R_1^\star$.

3.1 Underlay transmission strategy
We consider the extension to MIMO secondary systems of the underlay transmission strategy introduced in [14]. The primary transmitter is oblivious to the presence of the secondary users and broadcasts its single-stream signal with power P_1 using the covariance matrix $\boldsymbol{K}_1$ corresponding to the maximum-ratio transmit (MRT) beamformer, i.e. $\boldsymbol{K}_1 = P_1 \frac{\boldsymbol{h}_{11}\boldsymbol{h}_{11}^H}{\|\boldsymbol{h}_{11}\|^2}$. The primary receiver decodes the message in the presence of interference from the secondary system and noise. The secondary transmitter splits its message into two parts (i.e. rate splitting) using possibly different covariance matrices with possibly different powers for each of the parts: $\boldsymbol{K}_{2,1}$ and $\boldsymbol{K}_{2,2}$, respectively. The secondary receiver performs successive/interference decoding to recover the first part of the secondary message, then the primary message (i.e. the interference), and finally the second part of the secondary message.

The communication rate for the primary users is

$$R_1^{\text{und}} \triangleq \log\left(1 + \frac{\boldsymbol{h}_{11}^H \boldsymbol{K}_1 \boldsymbol{h}_{11}}{1 + \boldsymbol{h}_{21}^H (\boldsymbol{K}_{2,1} + \boldsymbol{K}_{2,2}) \boldsymbol{h}_{21}}\right), \qquad (7)$$

and the rate achieved by the secondary users is

$$R_2^{\text{und}} \triangleq \log \frac{\left| \boldsymbol{I} + \boldsymbol{H}_{22}^H (\boldsymbol{K}_{2,1} + \boldsymbol{K}_{2,2}) \boldsymbol{H}_{22} + \boldsymbol{H}_{12}^H \boldsymbol{K}_1 \boldsymbol{H}_{12} \right|}{\left| \boldsymbol{I} + \boldsymbol{H}_{22}^H \boldsymbol{K}_{2,2} \boldsymbol{H}_{22} + \boldsymbol{H}_{12}^H \boldsymbol{K}_1 \boldsymbol{H}_{12} \right|}$$
$$+ \log \left| \boldsymbol{I} + \boldsymbol{H}_{22}^H \boldsymbol{K}_{2,2} \boldsymbol{H}_{22} \right|. \qquad (8)$$

The first term in (8) corresponds to the part of the secondary message decoded in the presence of interference (both from primary transmitter and self-interference). The second term in (8) corresponds to the part of the secondary message recovered after decoding and subtracting the primary message. This adds the constraint that the secondary receiver must be able to decode the primary message as well. That is,

$$R_{1,2}^{\text{und}} \triangleq \log \frac{\left| \boldsymbol{I} + \boldsymbol{H}_{22}^H \boldsymbol{K}_{2,2} \boldsymbol{H}_{22} + \boldsymbol{H}_{12}^H \boldsymbol{K}_1 \boldsymbol{H}_{12} \right|}{\left| \boldsymbol{I} + \boldsymbol{H}_{22}^H \boldsymbol{K}_{2,2} \boldsymbol{H}_{22} \right|}. \qquad (9)$$

In addition, we have the constraint on the QoS for the primary user, i.e. $R_1^{\text{und}} \geq R_1^\star$. Note that by setting appropriately $\boldsymbol{K}_{2,1}$ and $\boldsymbol{K}_{2,2}$, we obtain the extreme cases, where the secondary receiver decodes first the primary message or does not decode it at all.

We remark that we do not make any assumption on the rank of the matrices $\boldsymbol{K}_{2,1}$ or $\boldsymbol{K}_{2,2}$. Basic considerations on the number of transmit/receive antennas required for multiple-stream transmission apply here, too (see e.g. [37]).

3.2 Problem formulation
The problem of finding the covariance matrices $\boldsymbol{K}_{2,1}$ and $\boldsymbol{K}_{2,2}$ that maximize the secondary rate under the aforementioned constraints is expressed as

$$\max_{\boldsymbol{K}_{2,1}, \boldsymbol{K}_{2,2}} R_2^{\text{und}} \qquad (10a)$$

subject to:

$$R_1^{\text{und}} \geq R_1^\star, \qquad (10b)$$

$$R_{1,2}^{\text{und}} \geq R_1^\star, \qquad (10c)$$

$$\text{tr}\{\boldsymbol{K}_{2,1} + \boldsymbol{K}_{2,2}\} \leq P_2, \qquad (10d)$$

$$\boldsymbol{K}_{2,1} \succeq 0, \boldsymbol{K}_{2,2} \succeq 0, \qquad (10e)$$

where it is implicitly assumed that (10c) applies only if $\boldsymbol{K}_{22} \neq \boldsymbol{0}$. Note that this problem is not concave due to the constraints (10b) and (10c). Constraint (10b) can easily be transformed into a linear constraint. However, dealing with (10c) is more involved.

3.3 Optimal transmission parameters
The following proposition characterizes the solution to (10). This extends the result in [14] to MIMO secondaries.

Proposition 1. *The optimization problem in* (10) *admits the following solution:*

Case 1: If

$$R_1^\star \geq \log \left| I + H_{12}^H K_1 H_{12} \right|, \tag{11}$$

then decoding the primary message at the secondary receiver is not possible at all. Without interference decoding, we have that $K_{2,2} = 0$, and $K_{2,1}$ is the covariance matrix that maximizes

$$\log \frac{\left| I + H_{22}^H K_{2,1} H_{22} + H_{12}^H K_1 H_{12} \right|}{\left| I + H_{12}^H K_1 H_{12} \right|} \tag{12}$$

subject to the corresponding constraints. This is equivalent to solving the following concave problem:

$$\max_{\Phi} \log \left| I + H_{22}^H \Phi H_{22} + H_{12}^H K_1 H_{12} \right| \tag{13a}$$

subject to:

$$\operatorname{tr}\{h_{21}^H \Phi h_{21}\} \leq P_{int}^{und}, \tag{13b}$$

$$\operatorname{tr}\{\Phi\} \leq P_2, \tag{13c}$$

$$\Phi \succeq 0, \tag{13d}$$

where

$$P_{int}^{und} \triangleq \frac{h_{11}^H K_1 h_{11}}{2^{R_1^\star} - 1} - 1. \tag{14}$$

Case 2: If

$$R_1^\star \leq \log \frac{\left| I + H_{22}^H \Sigma^\star H_{22} + H_{12}^H K_1 H_{12} \right|}{\left| I + H_{22}^H \Sigma^\star H_{22} \right|}, \tag{15}$$

where $\Sigma^\star$ is the covariance matrix that solves the concave problem

$$\max_{\Sigma} \log \left| I + H_{22}^H \Sigma H_{22} \right| \tag{16a}$$

subject to:

$$\operatorname{tr}\{h_{21}^H \Sigma h_{21}\} \leq P_{int}^{und}, \tag{16b}$$

$$\operatorname{tr}\{\Sigma\} \leq P_2, \tag{16c}$$

$$\Sigma \succeq 0, \tag{16d}$$

with P_{int}^{und} as defined in (14), then it is possible to decode the interference directly, without using rate splitting. Thus, the optimal covariance matrices are $K_{2,1} = 0$ and $K_{2,2} = \Sigma^\star$.

Case 3: In all other cases, i.e. if

$$\log \frac{\left| I + H_{22}^H \Sigma^\star H_{22} + H_{12}^H K_1 H_{12} \right|}{\left| I + H_{22}^H \Sigma^\star H_{22} \right|}$$
$$< R_1^\star < \log \left| I + H_{12}^H K_1 H_{12} \right|, \tag{17}$$

the problem is solved by $K_{2,1} = \gamma \Delta^\star$ and $K_{2,2} = (1 - \gamma) \Delta^\star$, where $\gamma \in [0, 1]$ is chosen such that $R_{1,2}^{und} = R_1^\star$, and $\Delta^\star$ is the matrix that solves the

following concave problem

$$\max_{\Delta} \left| I + H_{22}^H \Delta H_{22} + H_{12}^H K_1 H_{12} \right| \tag{18a}$$

subject to:

$$\operatorname{tr}\{h_{21}^H \Delta h_{21}\} \leq P_{int}^{und}, \tag{18b}$$

$$\operatorname{tr}\{\Delta\} \leq P_2, \tag{18c}$$

$$\Delta \succeq 0, \tag{18d}$$

with P_{int}^{und} as defined in (14).

Proof. The proof is provided in Appendix 1. □

Remark 1. In all three cases, the solution can be efficiently obtained using convex optimization tools [38].

Remark 2. The preceding results for case 3 reveal that the same covariance matrix (up to a scaling factor) is used for both parts of the secondary message when using rate splitting. For the case of beamformers, which are optimal for MISO secondaries (see e.g. [17] or [39]), this means that it suffices to consider the same beamformer for both parts of the secondary message (cf. [14]).

4 Overlay cognitive radio with explicit message-learning phase

In this section, we introduce the transmission strategy that we consider for the overlay cognitive radio paradigm. Our goal is again to maximize the communication rate of the secondary users while ensuring that the primary users have a minimum QoS, defined in terms of a minimum rate $R_1^\star$.

4.1 Overlay transmission strategy

Our strategy for overlay cognitive radio combines cooperative communication techniques, in particular decode-and-forward (DF) [26-28], with communication for cognitive radio channels [8,9]. The strategy makes full use of the potential of overlay cognitive radio by establishing active asymmetric cooperation between the users. The protocol establishes transmission of the primary message in two phases. Moreover, the primary transmitter chooses the system parameters as to maximize the system efficiency while ensuring that its message is reliably communicated. The secondary transmitter, which only broadcasts during the second phase, not only sends its own message but also acts as a relay for the message of the primary users. In addition to this, some degree of cooperation in the process of channel estimation is required so that the transmitters obtain the relevant channel state information.

Let $R_1^\star$ be the target rate of the primary users. In the first phase, of relative duration α, the primary transmitter

broadcasts its message using the $N_{T,1}$ antennas with transmit covariance matrix $K_1^{(1)} \succeq 0$. The primary receiver and secondary transmitter listen to this transmission. Consider the rates

$$R_1^{(1)} \triangleq \alpha \log \left(1 + h_{11}^H K_1^{(1)} h_{11} \right), \tag{19}$$

$$R_t \triangleq \alpha \log \left| I + H_t^H K_1^{(1)} H_t \right|, \tag{20}$$

and let $P_1^{(1)}$ denote the power spent by the primary transmitter in the first phase, i.e. $P_1^{(1)} \triangleq \text{tr}\{K_1^{(1)}\}$. Expressions (19) and (20) correspond to the rates from the primary transmitter to the primary receiver and to the secondary transmitter in the first phase, respectively.

If the channel H_t is significantly better than h_{11} (e.g. $\text{tr}\{H_t^H H_t\} \gg \|h_{11}\|^2$), then the secondary transmitter will need less redundancy to decode the message. In particular, if

$$R_1^{(1)} < R_1^\star \leq R_t, \tag{21}$$

then the secondary transmitter can decode the primary message but the primary receiver cannot. Although it cannot decode, the primary receiver has collected useful observations of the primary signal. Roughly speaking, it only needs additional redundancy to resolve its uncertainty and be able to decode [26].

Once the secondary transmitter is able to decode, the system can switch to the second phase. The second phase has the duration $1 - \alpha$ and consists of two simultaneous transmissions. On one hand, primary and secondary transmitters cooperate to resolve the uncertainty of the primary receiver. They act as one single virtual transmitter that uses a virtual covariance matrix

$$K_{\text{co}} = \begin{bmatrix} K_1^{(2)} & \Psi \\ \Psi^H & \frac{1}{1-\alpha} K_r \end{bmatrix} \tag{22}$$

to send the remaining part of the primary message over the extended channel $h_{\text{ext}}^H = [h_{11}^H, h_{21}^H]$ that consists of the concatenation of both channels to the primary receiver. The sub-matrices $K_1^{(2)}$ and K_r correspond to actual the covariance matrices used by each transmitter, while the sub-matrix Ψ corresponds to correlation of the signals sent by each transmitter, so that they add constructively at the receiver (cf. [18], Eq. (3)). Note that while they act coordinately, each transmitter has an independent power constraint (i.e. on $\text{tr}\{K_1^{(2)}\}$ and $\text{tr}\{K_r\}$, respectively): the primary transmitter uses the power left after the first phase, while the secondary uses only a fraction

of its available power. Simultaneously with this cooperative transmission, the secondary transmitter employs the remaining power and a different covariance matrix K_p for private communication to the secondary receiver. Moreover, it can use the knowledge of the primary message to predict the interference that the secondary receiver will experience and precode against it using dirty paper coding. Using this strategy, the rates

$$R_1^{(2)} \triangleq (1 - \alpha) \log \left(1 + \frac{h_{\text{ext}}^H K_{\text{co}} h_{\text{ext}}}{1 + \frac{1}{1-\alpha} h_{21}^H K_p h_{21}} \right), \tag{23}$$

$$R_2 \triangleq (1 - \alpha) \log \left| I + \frac{1}{1 - \alpha} H_{22}^H K_p H_{22} \right|, \tag{24}$$

are achievable for transmitting information about the primary message and the secondary message during the second phase. The factor $\frac{1}{1-\alpha}$ in front of the matrices K_p and K_r scales up the power to take into account the duration of the second phase.

Using DF relaying arguments (see e.g. [26,27]), it is possible to show that the rate

$$R_1 \triangleq R_1^{(1)} + R_1^{(2)} \tag{25}$$

is achievable for the primary users. Note that at this point, we do not make any assumption on the rank of the covariance matrices. In particular, K_p can incorporate multiple streams, subject to the usual constraints [37].

Remark 3. We stress that it is necessary that $R_t \geq R_1^\star$ to start the second phase. However, enforcing $R_t = R_1^\star$ does not necessarily yield the largest secondary rate. As we will see, it is sometimes better to extend 'artificially' the duration of the first phase.

Remark 4. The requirement of decoding the primary message at the secondary transmitter in combination with the use of dirty paper coding during the second phase renders ineffective the direct observation $y_2^{(1)}$ of the primary message obtained by the secondary receiver obtained during the first phase, that is, the rate (24) is already free from interference.

4.2 Problem formulation
We are interested in finding the choice of phase splitting α, covariance matrices $K_1^{(1)}, K_1^{(2)}, K_p$ and K_r, and the correlation matrix Ψ that maximize the secondary rate R_2 while ensuring a target rate $R_1^\star$ for the primary user pair under average power constraints P_1 and P_2 at the primary

and secondary transmitters, respectively. This is formulated mathematically as

$$\max_{\substack{\alpha, K_1^{(1)}, K_1^{(2)} \\ K_p, K_r, \Psi}} (1 - \alpha) \log \left| I + \frac{1}{1 - \alpha} H_{22}^H K_p H_{22} \right| \qquad (26a)$$

subject to:

$$R_t \geq R_1^\star, \qquad (26b)$$

$$R_1 \geq R_1^\star, \qquad (26c)$$

$$\alpha \operatorname{tr}\{K_1^{(1)}\} + (1 - \alpha) \operatorname{tr}\{K_1^{(2)}\} \leq P_1, K_1^{(1)} \succeq 0, K_1^{(2)} \succeq 0, \qquad (26d)$$

$$\operatorname{tr}\{K_p + K_r\} \leq P_2, K_p \succeq 0, K_r \succeq 0, K_{co} \succeq 0, \qquad (26e)$$

$$0 < \alpha < 1. \qquad (26f)$$

We characterize the solution to (26) in the following section.

4.3 Optimal transmission parameters

The problem in (26) is not convex; in particular, dealing with constraint (26c) is problematic. An exhaustive search over the 6 parameters seems unfeasible too. Our approach is to study the properties of the optimal parameters through a series of propositions. Then, we use them to reduce the optimization problem to a simpler search over a small set of bounded real-valued parameters and to find efficient algorithms to calculate the numerical values of the system parameters.

4.3.1 Characterization of the solution

As it was discussed in Section 4.1, our transmission strategy is reasonable only if the secondary transmitter can decode the primary message earlier than the primary receiver. This condition appears in the characterization of the solution to (26) and is captured by the following definition:

Definition 1 (Cooperation condition). Let

$$K^{\mathrm{WF}}(\sigma) \triangleq \underset{\Sigma \succeq 0: \operatorname{tr}\{\Sigma\} \leq \sigma}{\arg\max} \log \left| I + H_t^H \Sigma H_t \right| \qquad (27)$$

for some $\sigma \in \mathbb{R}^+$. We say that the cooperation condition is satisfied if

$$\log \left(1 + h_{11}^H K^{\mathrm{WF}}(P_1) h_{11} \right) < \log \left| I + H_t^H K^{\mathrm{WF}}(P_1) H_t \right|. \qquad (28)$$

The matrix $K^{\mathrm{WF}}(\sigma)$ corresponds to the waterfilling (WF) solution with power constraint σ. Note that if the cooperation condition is not satisfied, the primary receiver may decode the message earlier than the secondary transmitter when the transmission is optimized for the latter. In addition, we assume that $K^{\mathrm{WF}}(\sigma)$ is never proportional to the MRT covariance matrix $\frac{h_{11} h_{11}^H}{\|h_{11}\|^2}$.

This technical condition simply ensures that the transmission between transmitters is never strictly co-linear with h_{11} because this case would virtually turn the primary transmitter into a single-antenna transmitter.

The first observation that we make regarding the solution to (26) concerns the power used by the transmitters. Over the two phases, the primary transmitter uses all its available power. Note that this power is in general distributed unequally over the phases. Similarly, the secondary transmitter also exhausts all its power, distributing it between the two simultaneous transmissions: cooperation and private communication. This is stated in the following proposition.

Proposition 2. *The optimal transmission strategy in* (26) *makes use of all the available power at the primary and secondary transmitters, that is,*

1. $\operatorname{tr}\{K_p + K_r\} = P_2$,
2. $\alpha \operatorname{tr}\{K_1^{(1)}\} + (1 - \alpha) \operatorname{tr}\{K_1^{(2)}\} = P_1$.

Proof. The proof is provided in Appendix 2. □

Our second observation is that the presence of the secondary transmitter always pushes the primary system to the limit of decodability as described by the following proposition:

Proposition 3. *The set of parameters that solves the optimization problem in* (26) *satisfies*

$$R_1^{(1)} + R_1^{(2)} = R_1^\star \qquad (29)$$

(i.e. constraint (26c) *with equality) if the cooperation condition is satisfied.*

Proof. The proof is provided in Appendix 3. □

This result is a consequence of the tight interaction between users allowed in overlay cognitive radio scenarios. On one hand, the secondary system makes use of its resources in the way that maximizes the rate R_2. At the same time, the primary transmitter cooperates towards this goal by distributing its resources between the two phases in the way that R_2 is maximized. For example, it may choose a covariance matrix $K_1^{(1)}$ that makes the first phase shorter if this is beneficial in terms of secondary rate.

We can make a similar observation with respect to the communication between transmitters in the first phase.

Proposition 4. *The set of parameters that solve the optimization problem in* (26) *satisfies*

$$R_{\mathrm{t}} = R_1^\star \tag{30}$$

(*i.e. constraint* (26b) *with equality) unless the optimal covariance matrix* $\boldsymbol{K}_1^{(1)}$ *is proportional to the orthogonal projector onto* $\boldsymbol{h}_{11}$, *that is, proportional to* $\frac{\boldsymbol{h}_{11}\boldsymbol{h}_{11}^H}{\|\boldsymbol{h}_{11}\|^2}$.

Proof. The proof is provided in Appendix 4.   □

This result can be interpreted in terms of the duration of the phases. In the cases where (30) holds, the system switches from first phase to second phase as soon as the secondary transmitter can decode the primary message. However, (30) is not always satisfied; hence, this is not true in general. In fact, it is sometimes beneficial to extend 'artificially' the first phase in order to achieve a larger secondary rate. For example, if the primary transmitter only has one antenna, then we cannot find non-trivial conditions that ensure $R_{\mathrm{t}} = R_1^\star$. The reason for this is that with only one antenna, there is no way to distinguish directions, i.e. we always transmit in the direction to the primary receiver. Similarly, it was observed in [27] in the context of DF for single-antenna Gaussian relay channels that the optimal split of phases has to be found numerically.

Although Proposition 4 only gives a partial characterization of the covariance matrix $\boldsymbol{K}_1^{(1)}$, it turns out to be very useful when it comes to finding its value numerically. Combined with Proposition 2, it allows us to derive Algorithm 1 that efficiently finds $\boldsymbol{K}_1^{(1)}$ given the optimal values of the phase split α and the power used by the primary in the first phase (i.e. $P_1^{(1)} \triangleq \mathrm{tr}\{\boldsymbol{K}_1^{(1)}\}$).

Algorithm 1 Find optimal covariance matrix

1: **procedure** OPTIMAL-COVARIANCE($\alpha, P_1^{(1)}$)
2:     $P_{\mathrm{f}} \leftarrow P_1^{(1)}$
3:     $\boldsymbol{K}_{\mathrm{f}} \leftarrow \underset{\boldsymbol{\Sigma} \succeq 0:\ \mathrm{tr}\{\boldsymbol{\Sigma}\} \leq P_{\mathrm{f}}}{\arg\max} \log\left|\boldsymbol{I} + \boldsymbol{H}_{\mathrm{t}}^H \boldsymbol{\Sigma} \boldsymbol{H}_{\mathrm{t}}\right|$
4:     **if** $\alpha \log\left|\boldsymbol{I} + \boldsymbol{H}_{\mathrm{t}}^H \boldsymbol{K}_{\mathrm{f}} \boldsymbol{H}_{\mathrm{t}}\right| < R_1^\star$ **then**
5:         $\boldsymbol{K}_1^{(1)} \leftarrow \emptyset$
6:     **else**
7:         $P_{\mathrm{h}} \leftarrow P_1^{(1)}$
8:         $\boldsymbol{K}_{\mathrm{h}} \leftarrow \frac{\boldsymbol{h}_{11}\boldsymbol{h}_{11}^H}{\|\boldsymbol{h}_{11}\|^2} P_{\mathrm{h}}$
9:         **if** $\alpha \log\left|\boldsymbol{I} + \boldsymbol{H}_{\mathrm{t}}^H \boldsymbol{K}_{\mathrm{h}} \boldsymbol{H}_{\mathrm{t}}\right| \geq R_1^\star$ **then**
10:            $\boldsymbol{K}_1^{(1)} \leftarrow \boldsymbol{K}_{\mathrm{h}}$
11:        **else**
12:            BISECTION($R_1^\star, P_1^{(1)}, \epsilon$)
13:        **end if**
14:    **end if**
15:    **return** $\boldsymbol{K}_1^{(1)}$
16: **end procedure**

Algorithm 2 Bisection method

1: **procedure** BISECTION($R_1^\star, P, \epsilon$)
2:     $P_{\mathrm{h,top}} \leftarrow P$
3:     $P_{\mathrm{h,bot}} \leftarrow 0$
4:     **while** true **do**
5:         $P_{\mathrm{h}} \leftarrow \frac{P_{\mathrm{h,top}} + P_{\mathrm{h,bot}}}{2}$
6:         $P_{\mathrm{f}} \leftarrow P - P_{\mathrm{h}}$
7:         $\boldsymbol{K}_{\mathrm{h}} \leftarrow P_{\mathrm{h}} \frac{\boldsymbol{h}_{11}\boldsymbol{h}_{11}^H}{\|\boldsymbol{h}_{11}\|^2}$
8:         $\boldsymbol{K}_{\mathrm{f}} \leftarrow \underset{\boldsymbol{\Sigma} \succeq 0:\mathrm{tr}\{\boldsymbol{\Sigma}\} \leq P_{\mathrm{f}}}{\arg\max} \log\left|\boldsymbol{I} + \boldsymbol{H}_{\mathrm{t}}^H(\boldsymbol{K}_{\mathrm{h}} + \boldsymbol{\Sigma})\boldsymbol{H}_{\mathrm{t}}\right|$
9:         gap $\leftarrow \alpha \log\left|\boldsymbol{I} + \boldsymbol{H}_{\mathrm{t}}^H(\boldsymbol{K}_{\mathrm{h}} + \boldsymbol{K}_{\mathrm{f}})\boldsymbol{H}_{\mathrm{t}}\right| - R_1^\star$
10:        **if** gap < 0 **then**
11:            $P_{\mathrm{h,top}} \leftarrow P_{\mathrm{h}}$
12:        **else if** $\epsilon <$ gap **then**
13:            $P_{\mathrm{h,bot}} \leftarrow P_{\mathrm{h}}$
14:        **else**
15:            $\boldsymbol{K}_1^{(1)} \leftarrow \boldsymbol{K}_{\mathrm{h}} + \boldsymbol{K}_{\mathrm{f}}$
16:            **break**
17:        **end if**
18:    **end while**
19:    **return** $\boldsymbol{K}_1^{(1)}$
20: **end procedure**

Algorithm 1 starts by verifying (line 4) if a solution to (26b) exists for the given level of power $P_1^{(1)}$ by allocating it freely, as in $\boldsymbol{K}_{\mathrm{f}}$, to maximize the expression in line 3. Provided that such solution exists, the algorithm verifies if MRT beamforming to the primary receiver (i.e. in the direction of $\boldsymbol{h}_{11}$, using the covariance matrix $\boldsymbol{K}_{\mathrm{h}}$) is sufficient for decoding at the secondary transmitter (26b) (line 9). If MRT does not satisfy (26b), then it uses the bisection method (Algorithm 2) to find the covariance matrix with largest component in the direction of $\boldsymbol{h}_{11}$ that satisfies (26b). The search finishes when the rate achieved for this choice of covariance matrix exceeds the target rate $R_1^\star$ by less than a predefined threshold ϵ. The maximization in Algorithm 1 (line 3) and in the bisection method (Algorithm 2, line 8) can be written as standard waterfilling problems, which can be efficiently approximated or solved exactly (see e.g. [40]). The following corollary establishes the the optimality of Algorithm 1.

Corollary 1. Given the optimal values of α and power $P_1^{(1)}$ used by the primary in the first phase, Algorithm 1 finds the optimal covariance matrix $\boldsymbol{K}_1^{(1)}$ if the cooperation condition is satisfied.

Proof. The proof is provided in Appendix 5.   □

Remark 5. Note that, by construction, if a call to Algorithm 1 results in the MRT covariance matrix for

some $(\alpha, P_1^{(1)})$, then it will also result in the MRT covariance matrix for any $(\alpha, \tilde{P}_1^{(1)})$ with $\tilde{P}_1^{(1)} > P_1^{(1)}$.

We conclude this section by characterizing the optimal covariance matrices used in the second phase.

Proposition 5. *The optimal covariance matrices in the second phase are given by*

$$K_1^{(2)} = P_1^{(2)} \frac{h_{11} h_{11}^H}{\|h_{11}\|^2}, \quad K_r = P_r \frac{h_{21} h_{21}^H}{\|h_{21}\|^2},$$

$$\Psi = \sqrt{P_1^{(2)} P_r} \frac{h_{11} h_{21}^H}{\|h_{11}\| \|h_{21}\|}, \tag{31}$$

and K_p is the solution to the following concave problem:

$$\max_{\Omega} (1 - \alpha) \log \left| I + \frac{1}{1 - \alpha} H_{22}^H \Omega H_{22} \right| \tag{32a}$$

subject to:

$$\text{tr}\{h_{21}^H \Omega h_{21}\} \leq P_{int}, \tag{32b}$$

$$\text{tr}\{\Omega\} = P_2 - P_r, \Omega \succeq 0, \tag{32c}$$

where

$$P_1^{(2)} \triangleq \frac{P_1 - \alpha \, \text{tr}\{K_1^{(1)}\}}{1 - \alpha}, \tag{33}$$

$$P_{int} \triangleq (1 - \alpha) \left(\frac{\left(\|h_{11}\| \sqrt{P_1^{(2)}} + \|h_{21}\| \sqrt{\frac{P_r}{1 - \alpha}} \right)^2}{2^{\frac{R_1^\star - R_1^{(1)}}{1 - \alpha}} - 1} - 1 \right), \tag{34}$$

for some $P_r \in [0, P_2]$ such that $P_{int} \geq 0$.

Proof. The proof is provided in Appendix 6. $\qquad\square$

The interpretation of the optimal values for $K_1^{(2)}$ and K_r is straightforward: they are adapted to their respective channels and combine coherently at the receiver. The matrix K_p used for the secondary communication is chosen to maximize the secondary rate without violating the interference constraint at the primary.

In the case of secondary MISO systems (i.e. h_{12} and h_{22} instead of H_{12} and H_{22}, respectively), there is no loss in restricting the covariance matrix K_p at the secondary transmitter to have rank 1, i.e. $K_p = (P_2 - P_r) w_p w_p^H$. The following corollary characterizes the optimal beamforming vector w_p.

Corollary 2. The optimal beamformer w_p is

$$w_p = \sqrt{\lambda} \frac{\Pi_{h_{21}} h_{22}}{\|\Pi_{h_{21}} h_{22}\|} + \sqrt{1 - \lambda} \frac{\Pi_{h_{21}}^\perp h_{22}}{\|\Pi_{h_{21}}^\perp h_{22}\|} \tag{35}$$

with

$$\lambda = \begin{cases} \lambda_{\text{MRT}} & \text{if } \lambda_{\text{MRT}} \leq \frac{P_{\text{int}}}{\|h_{21}\|^2 (P_2 - P_r)} \\ \frac{P_{\text{int}}}{\|h_{21}\|^2 (P_2 - P_r)} & \text{otherwise,} \end{cases} \tag{36}$$

$$\lambda_{\text{MRT}} = \frac{\left\| \Pi_{h_{21}} h_{22} \right\|^2}{\|h_{22}\|^2}, \tag{37}$$

with P_{int} as defined in (34), for some $P_r \in [0, P_2]$ such that $P_{\text{int}} \geq 0$.

Proof. The proof is provided in Appendix 7. $\qquad\square$

In the MISO case, we see more clearly that the beamformer w_p used for the secondary communication is chosen to be the one with largest projection over h_{22} that satisfies the interference constraint, which is determined by the projection over h_{21} [13,16].

4.3.2 An algorithm to find the optimal parameters
The results from the previous section allow us to reduce the solution to (26) to a search over three real-valued parameters: the phase split α, the power spent by the primary in the first phase (i.e. $P_1^{(1)} \triangleq \text{tr}\{K_1^{(1)}\}$), and the distribution of power between relaying and private communication at the secondary (e.g. $P_r = \text{tr}\{K_r\}$). Each of these parameters is defined in a closed and bounded interval. In contrast, solving (26) directly requires search over one real-valued parameter and five complex-valued matrices. We have summarized this simplified search in Algorithm 3, which we describe in the following:

To find the solution, we perform a search over the phase split α and the admissible power for the primary transmitter in the first phase $P_1^{(1)}$. Given these two values, the matrix $K_1^{(1)}$ is found using Algorithm 1, whereas $K_1^{(2)}$ is readily determined. To obtain the remaining matrices K_p, K_r and Ψ, we perform a search over the different splits of secondary power using the results in Proposition 5. The optimal choice of parameters is the one that yields the largest secondary rate R_2.

5 Numerical evaluation
5.1 Geometrical model
To present our results, we will use the simple geometrical model in Figure 2, in which the different nodes are placed on a plane. The relative positioning of the nodes is summarized by the distance between each pair of nodes. We model the block flat fading channel coefficient between two nodes as

$$h_{ij} = \frac{1}{\sqrt{d_{ij}^p}} \tilde{h}_{ij}, \tag{38}$$

where d_{ij} is the distance between them, p is the path loss exponent, and $\tilde{h}_{ij} \sim \mathcal{CN}(0, 1)$. In the case of channel

---

Algorithm 3 Algorithm to find the optimal parameters

1: **for** $\alpha \leftarrow [0, \ldots, 1]$ **do**
2:      **for** $P_1^{(1)} \leftarrow [0, \ldots, \frac{P_1}{\alpha}]$ **do**
3:         $P_1^{(2)} \leftarrow \frac{P_1 - \alpha P_1^{(1)}}{1 - \alpha}$
4:         $\boldsymbol{K}_1^{(1)} \leftarrow \text{OPTIMAL-COVARIANCE}(\alpha, P_1^{(1)})$
5:         $\boldsymbol{K}_1^{(2)} \leftarrow P_1^{(2)} \frac{\boldsymbol{h}_{11} \boldsymbol{h}_{11}^H}{\|\boldsymbol{h}_{11}\|^2}$
6:         **for** $P_{\mathrm{r}} \leftarrow [0, \ldots, P_2]$ **do**
7:            $\boldsymbol{K}_{\mathrm{r}} \leftarrow P_{\mathrm{r}} \frac{\boldsymbol{h}_{21} \boldsymbol{h}_{21}^H}{\|\boldsymbol{h}_{21}\|^2}$
8:            $\boldsymbol{\Psi} \leftarrow \sqrt{P_1^{(2)} P_{\mathrm{r}}} \frac{\boldsymbol{h}_{11} \boldsymbol{h}_{21}^H}{\|\boldsymbol{h}_{11}\| \|\boldsymbol{h}_{21}\|}$
9:            Obtain $\boldsymbol{K}_{\mathrm{p}}$
10:            Compute R_2
11:         **end for**
12:      **end for**
13: **end for**

---

vectors or matrices, each of the entries is independently modeled as in (38).

For convenience, we normalize all distances with respect to the distance between the primary users (i.e. $d_{11} = 1$). We will consider the square surface $\{(x, y) : x \in [0, 1], y \in [0, 1]\}$, and vary the position of the secondary nodes (relative to the primary nodes) over a regular square grid of size 11×11, that is, we will move the secondary transmitter and receiver over this grid, always parallel to the line between primary transmitter and receiver (as in Figure 2). The primary transmitter and receiver will be fixed at positions $(0, 0.5)$ (black filled circle) and $(1, 0.5)$ (black filled box), respectively.

In the plots, a pair of coordinates (x, y) identifies the position of the secondary transmitter. All our results consider $d_{22} = 1/4$ while the remaining distances d_{12}, d_{21} and d_{tt} vary as described before. This models a secondary middle-range communication in the presence of primary users.

5.2 Note on the strategies
The overlay strategy in Section 4.1 yields $R_2 = 0$ for some channel realizations. The reason for this is that
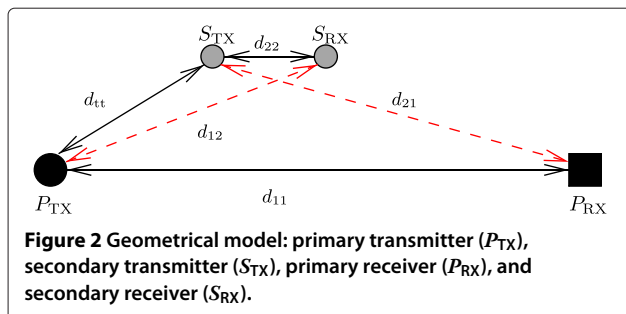
Figure 2 Geometrical model: primary transmitter (P_{TX}), secondary transmitter (S_{TX}), primary receiver (P_{RX}), and secondary receiver (S_{RX}).

constraint (26b) cannot always be fulfilled for $R_2 > 0$. In such a scenario, a cognitive radio system would switch to a different transmission strategy that can provide a non-zero secondary rate R_2. For example, it could switch to the underlay transmission mode presented here. In this way, the hybrid overlay-underlay strategy would never perform worse than the pure underlay strategy. However, including such a functionality in our experiments is against the nature of our work, which is to compare the underlay and overlay scenarios, and evaluate the effect of the learning phase. For this reason, we implement the strategies exactly as described in Sections 3.1 and 4.1.

5.3 Complexity of the strategies
The complexity of the underlay solution varies for the different cases in Proposition 1, which depend on the instantaneous channel conditions. For cases 1 and 2, the complexity is that of solving one concave problem ((13) and (16), respectively). For case 3, the complexity is that of solving two concave problems: (16) (to check the constraint) and (18), and finding the optimal split γ (e.g. using a loop or a bisection method). For MISO secondaries, the complexity can be lowered (e.g. using Remark 2 and [14]).

In contrast, Algorithm 3 finds the optimal overlay transmission parameters by searching over three-real valued parameters defined on a closed and bounded space. Up to a scaling factor that depends on the powers, the matrices $\boldsymbol{K}_1^{(2)}, \boldsymbol{K}_{\mathrm{r}}$ and $\boldsymbol{\Psi}$ can be determined before hand. The covariance matrix $\boldsymbol{K}_1^{(1)}$ needs to be determined for each pair $(\alpha, \mathrm{tr}\{\boldsymbol{K}\})$ using Algorithm 1. This algorithm relies on the waterfilling and bisection methods that can be implemented very efficiently (see e.g. [40]). In addition, note that Remark 5 can be used to minimize the number of calls to Algorithm 1. The optimal $\boldsymbol{K}_{\mathrm{p}}$ needs to be determined for each triple $(\alpha, P_1^{(1)}, P_{\mathrm{r}})$ by solving the concave problem in (32), which can also be implemented efficiently. Solving this last problem can be avoided in the case where $\boldsymbol{K}_{\mathrm{p}}$ has rank 1 using the results in Corollary 2.

When compared, it is clear that the complexity of solving the overlay problem is significantly larger than that of the underlay problem, in particular for the case where $\boldsymbol{K}_{\mathrm{p}}$ is not rank 1. Nevertheless, the solution to both problems reduces to solving concave problems, for which a large variety of efficient algorithms exist (see e.g. [38]).

5.4 Simulation results
We have performed extensive simulations of our underlay and overlay cognitive radio strategies to assess their individual performances and merits relative to each other. We show here results for a few representative cases and comment in the end on the differences for other system parameters.

In the results in Figures 3, 4 and 5 the transmitters are equipped with $N_{T,1} = N_{T,2} = 2$ antennas, and the receivers with one single antenna. In contrast, in Figure 6, we study the behavior for varying $N_{T,1}$ and $N_{T,2}$ and single-antenna receivers. In all cases, the path loss exponent is fixed to $p = 3$, and the primary power is set to $P_1 = 10$ dB. The secondary power is $P_2 = 1$ dB for the results in Figures 3 to 5 and variable for Figure 6. We assume that the primary system has a target rate $R_1^\star$ that corresponds to a fraction ρ of its instantaneous point-to-point Shannon capacity, that is,

$$R_1^\star = \rho \log(1 + \|\boldsymbol{h}_{11}\|^2 P_1). \qquad (39)$$

We refer to ρ as the load factor of the primary system. We consider $\rho = 0.75$ for Figures 3 to 5, and $\rho = 1$ for Figure 6. Every point in the plots represents the average over $5 \cdot 10^4$ independent realizations of the channels. We focus on the results for the overlay strategy and the comparison between the strategies because the results for the underlay strategy alone do not differ qualitatively from the single-antenna case in [14].

Figure 3 shows the average of the secondary rate R_2 (in bits per channel use, bpcu) achieved by our overlay cognitive radio strategy for $N_{T,1} = N_{T,2} = 2, N_{R,2} = 1$, $P_1 = 10$ dB, $P_2 = 1$ dB, $p = 3$ and $\rho = 0.75$. To set the numerical values in the figure in a context note that if the secondaries were alone in the scenario, the ergodic

capacity would be 6.96 bpcu. In comparison, the highest average secondary rate in Figure 3 is $R_2 = 6.29$ bpcu and is obtained when primary and secondary transmitters are closely located. This represents 90% of the aforementioned capacity. As one would expect, the average secondary rate becomes lower as the two transmitters are separated.

It is more interesting to look at the advantage in average rate over the underlay strategy. Figure 4 shows the ratio between the average of the secondary rate for overlay R_2 and the average of the secondary rate for underlay R_2^{und} for $N_{T,1} = N_{T,2} = 2, N_{R,2} = 1, P_1 = 10$ dB, $P_2 = 1$ dB, $p = 3$ and $\rho = 0.75$. The results are somewhat surprising in the sense that the largest-advantage region does not correspond to the largest-secondary-rate region, that is, the maximum in Figure 4 is not obtained for $(x, y) = (0, 0.5)$ but rather for $(x, y) \approx (0.4, 0.5)$. The reason for this is that for $(x, y) = (0, 0.5)$, the underlay strategy also benefits from closely located transmitters, thanks to the interference decoding functionalities. In fact, if one removes this functionality in the underlay transmission mode, the results change significantly. In that case, the overlay system is overwhelmingly better than the underlay strategy.

In addition, note that the advantage of the overlay system diminishes as the two transmitters are separated. In fact, in some regions, using the underlay strategy is better in terms of average secondary rate. The reason for
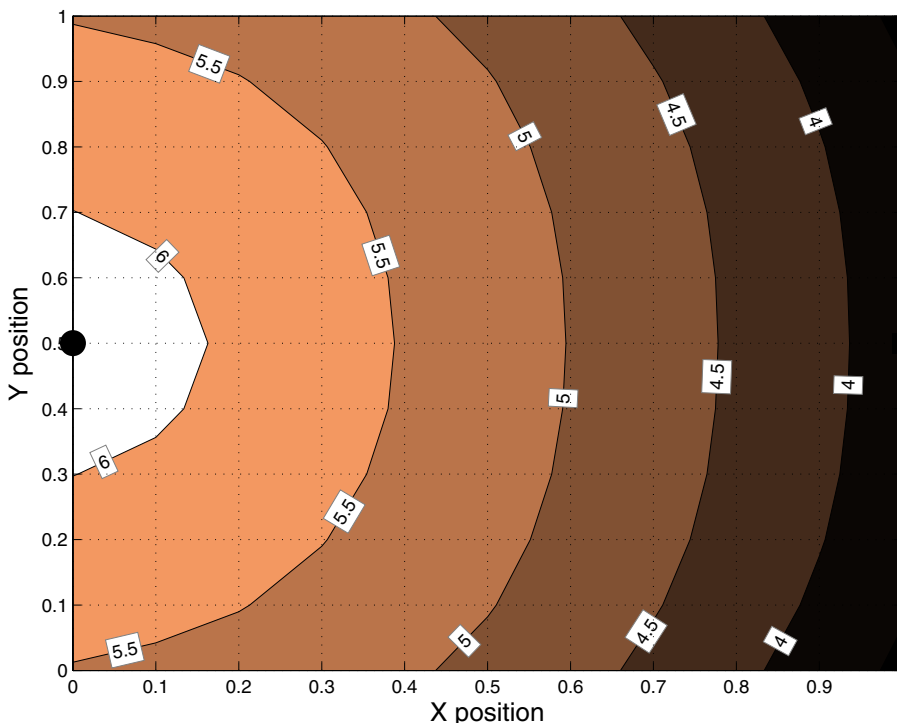

Figure 3 Overlay cognitive radio achievable rate R_2 averaged over channel realizations.
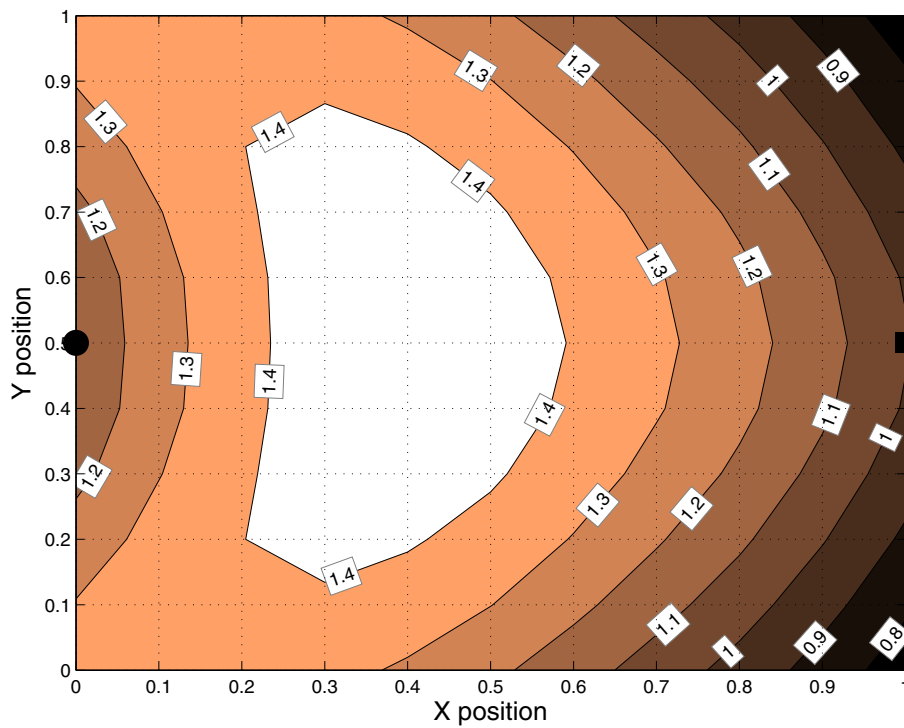
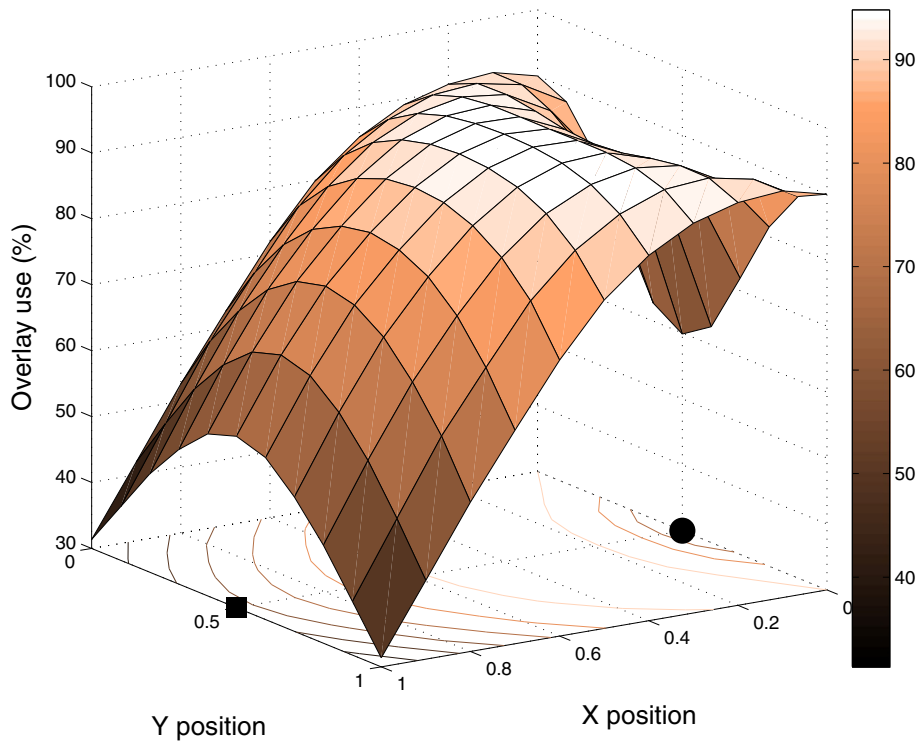**Figure 4 Ratio of the average of the rates R_2 (overlay) and R_2^{und} (underlay).**

Figure 5 Percentage of channel realizations yielding an overlay rate R_2 larger than the underlay rate R_2^{und}.
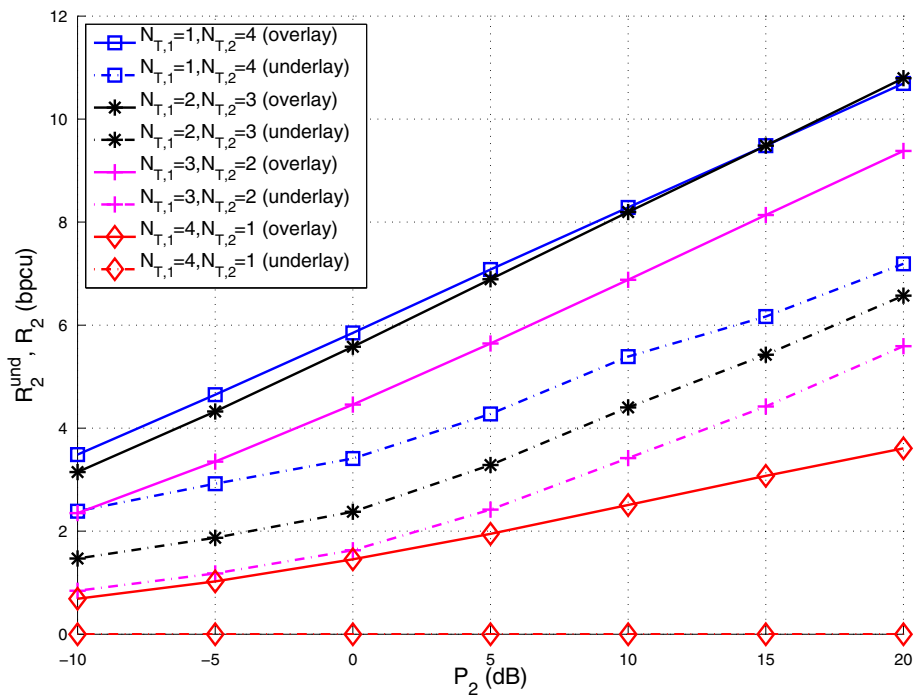
Figure 6 Averaged rates R_2^{und} (underlay) and R_2 (overlay) for different secondary powers and transmit antenna configurations.

this is simple: in these regions, the first phase is relatively long (e.g. $\alpha > 0.5$), and the higher sophistication of the secondary transmitter (i.e. dirty-paper coding, cooperative transmission) cannot compensate for the loss in secondary rate due to the passive first phase. Thus, the underlay approach, even if it has to transmit mainly in the zero-forcing direction to avoid interference, can make a more efficient use of the resources and provide a larger rate to the secondary users.

In order to implement a system that combines both strategies (as discussed in Section 5.2), it is desirable to know how often they outperform each other. This is shown in Figure 5, in terms of the percentage of channel realizations for which the overlay strategy yields a larger rate than the underlay strategy for $N_{T,1} = N_{T,2} = 2, N_{R,2} = 1, P_1 = 10$ dB, $P_2 = 1$ dB, $p = 3$ and $\rho = 0.75$. Again, we observe that the region with largest rate corresponding to the overlay strategy does not correspond exactly to the collocation of transmitters. In the figure, we observe that, except for a small region where overlay is better over 90% of the time, there is room for significant improvement if the system implements both strategies and chooses the best one in each block.

Regarding variations in the scenario, we have observed the following general trends. The secondary rate (Figure 3) increases with both the number of antennas and the secondary power as one would expect. More interestingly, as we increase the secondary power P_2 or the number

of antennas, the maximum in Figure 4 (i.e. the advantage of overlay in terms of average rate) increases its value and shifts its position towards the primary transmitter. The load factor ρ is the parameter that has the most impact: the largest advantages of the overlay strategy are obtained for high primary load factors. For example, if $\rho = 1$, the maximum advantage corresponds to a factor of approximately 2.55. In contrast, for small loads, the advantage might be too small to compensate for the additional complexity when compared to the underlay strategy; for example, in the case of a single-antenna primary system, we observed an advantage factor of just 1.15 (see [36]). Similar conclusions can be drawn for Figure 5: the maximum tends to move towards the primary transmitter as we increase the secondary power or the number of antennas and the region where overlay is better most of the time becomes larger. Finally, for larger path losses (e.g. $p = 4$), the results become more extreme: the positions of the maxima in Figures 3 to 5 remain the same, but their values are higher. In contrast, when the transmitters are separated, the underlay scheme yields a larger advantage than the one presented here.

Finally, Figure 6 shows the behavior of the underlay and overlay strategies in terms of the average of the rates and R_2^{und} and R_2, respectively, as a function of the secondary power P_2 for different transmit antenna configurations such that $N_{T,1} + N_{T,2} = 5$ and $N_{R,2} = 1$ for $P_1 = 10$ dB in a fully loaded system, i.e. $\rho = 1$, with path loss exponent

$p = 3$. The secondary transmitter is placed at position $(x, y) = (0.3, 0.5)$, i.e. on the line between the primary users. The main observation is that, in terms of secondary rate, it is better to deploy the antennas at the secondary transmitter rather than at the primary transmitter. In the underlay case, this is rather straightforward for the secondary system cannot benefit from the antennas at the primary. In the overlay case, this observation implies that the gains obtained via spatial diversity (i.e. larger $N_{\text{T},2}$) increase faster than those obtained by shortening the learning phase (i.e. larger $N_{\text{T},1}$). However, observe that increasing $N_{\text{T},2}$ suffers from a law of diminishing returns and that beyond a certain value the gains are minor. Regarding the changes in the behavior for varying secondary power P_2, we observe the following general trends. For very low P_2, all the strategies are power-constrained, and thus the gap between underlay and overlay vanishes. This effect is more pronounced for $\rho < 1$, where the primary can tolerate some interference. The gap between the strategies widens as P_2 increases, meaning, than when the secondary transmitter is no longer power limited, the use of spatial shaping alone fails to exploit the available resources. A special, extreme case is the underlay strategy with $N_{\text{T},2} = 1$: lacking spatial resources, it cannot make any use of a fully loaded primary channel, i.e. $R_2 = 0$ independently of P_2.

6 Coexistence with MIMO primary systems

The discussion in this paper has been restricted to the coexistence of a MIMO secondary system with a MISO primary link. The results presented here cannot be extended in their totality to the case of MIMO primaries neither for underlay nor for overlay. However, as we will see in this section, under some reasonable assumptions, they carry over to scenarios with MIMO primary systems.

In the case of underlay cognitive radio, it is important to emphasize the underlying assumption that the primary users are oblivious to the presence of secondary users. This effectively decouples the design of the optimal secondary transmitter from the primary transmit parameters. Moreover, note that the effect of the primary users enters the optimization in (10) through constraints (10b) and (10c). The validity of Lemma 1 which plays a fundamental role in dealing with the non-convexity of (10c) does not rely on any assumption about the primary transmit covariance matrix and thus applies to the primary MIMO case as well. In contrast, the simple transformation of (10b) into a linear constraint (i.e. (40b)) is no longer possible in the MIMO primary case. If, however, this constraint is replaced by a constraint that is linear or convex in $(\boldsymbol{K}_{21}, \boldsymbol{K}_{21})$, then the results in Proposition 1 remain valid. For example, one may define a constraint analog to (10b) by considering the worst-interference

direction in the span of $\boldsymbol{H}_{21}$. Alternatively, if the primary system uses single-stream transmission with fixed receiver beamformer, the results presented here remain valid.

In the case of the overlay cognitive radio strategy, the problem is more involved. In addition to a similar problem regarding constraint (26c), the transmit strategies of primary and secondary systems are necessarily coupled by the very nature of the extended cognitive radio channel (i.e. by the message-learning phase). Moreover, in the case of MIMO primaries, the optimization over the virtual joint covariance matrix $\boldsymbol{K}_{\text{co}}$ is more complex than in the case of MISO primaries, where beamforming was optimal, and thus $\boldsymbol{K}_{\text{co}}$ could be determined easily. This is issue is especially important when considering efficient algorithms to find the optimal parameters. Notwithstanding these considerations, the results in this paper remain valid if the primary system uses single-stream transmission with fixed receive beamformer, as in the case of underlay.

7 Conclusion

In this paper, we have studied the transmission strategies for underlay cognitive radio and overlay cognitive radio with an explicit learning phase, in which the secondary transmitter acquires the primary message. Our strategy for underlay uses interference decoding and exploits spatial resources using multi-antenna methods. For the overlay case, we have combined cooperative communication techniques (decode-and-forward relaying) with communication over a cognitive radio channel (cooperation and interference control at the primary receiver and interference pre-cancellation at the secondary transmitter) using multi-antenna methods. For both strategies, we have characterized the set of system parameters that maximize the secondary rate while ensuring a fixed rate for the primary system.

Finally, we have evaluated the performance of the strategies relative to each other in order to quantify the advantages and disadvantages of the degrees of coordination (i.e. uncoordinated for underlay vs. message-learning phase and cooperative communication for overlay). We have observed that for a wide range of channel conditions, when the primary and secondary transmitters are close to each other, the overlay strategy provides a significant advantage over the underlay strategy. This gain is particularly relevant for those scenarios where the secondary is interference-limited rather than power-limited. However, as the distance between transmitters becomes larger, this advantage vanishes and in fact at some point underlay starts outperforming overlay. Our analysis reveals that a combination of underlay and overlay strategies is necessary to exploit best the available resources, especially if the users in the system do not have fixed positions.

Appendices

Appendix 1

Proof of proposition 1

We first prove an auxiliary lemma that will be used in the proof of Proposition 1. Note that using simple manipulations, the optimization problem in (10) can be reformulated as

$$\max_{K_{2,1},K_{2,2}} \left| I + H_{22}^H(K_{2,1} + K_{2,2})H_{22} + H_{12}^H K_1 H_{12} \right| - R_{1,2}^{\mathrm{und}} \tag{40a}$$

subject to:

$$\mathrm{tr}\{h_{21}^H(K_{2,1} + K_{2,2})h_{21}\} \le P_{\mathrm{int}}^{\mathrm{und}}, \tag{40b}$$

$$R_{1,2}^{\mathrm{und}} \ge R_1^\star, \tag{40c}$$

$$\mathrm{tr}\{K_{2,1} + K_{2,2}\} \le P_2, \tag{40d}$$

$$K_{2,1} \succeq 0, K_{2,2} \succeq 0, \tag{40e}$$

with $P_{\mathrm{int}}^{\mathrm{und}}$ as defined in (14).

We will show now that when considering case 3, there is no loss of generality in restricting constraint (40c) to be an equality.

Lemma 1. *Any optimal point that falls within case 3 can be attained by a pair of covariance matrices $(\tilde{K}_{2,1}, \tilde{K}_{2,2})$, such that $\tilde{K}_{2,2}$ satisfies constraint* (40c) *with equality.*

Proof. Let $K_{2,1}$ and $K_{2,2}$ solve the optimization problem and assume that

$$R_{1,2}^{\mathrm{und}}(K_{2,2}) > R_1^\star, \tag{41}$$

where the notation $R_{1,2}^{\mathrm{und}}(K_{2,2})$ stresses out the dependency of $R_{1,2}^{\mathrm{und}}$ on $K_{2,2}$. Similarly, the notation $R_2^{\mathrm{und}}(K_{2,1}, K_{2,2})$ will stress out the dependency of R_2^{und} on $K_{2,1}$ and $K_{2,2}$.

First, we consider the case $K_{2,1} = 0$. Let $\Sigma^\star$ be the solution to problem (16) (in case 2) and recall that

$$R_{1,2}^{\mathrm{und}}(\Sigma^\star) < R_1^\star, \tag{42}$$

$$R_2^{\mathrm{und}}(0, K_{2,2}) \le R_2^{\mathrm{und}}(0, \Sigma^\star), \tag{43}$$

for case 3. Now, construct the new covariance matrix

$$\tilde{K}_{2,2} = \gamma K_{2,2} + (1 - \gamma)\Sigma^\star. \tag{44}$$

Note that for any $\gamma \in [0, 1]$, this matrix satisfies constrains (40b), (40d) and (40e), and

$$R_2^{\mathrm{und}}(0, K_{2,2}) \le R_2^{\mathrm{und}}(0, \tilde{K}_{2,2}), \tag{45}$$

by the concavity property of the log-determinant. $R_{1,2}^{\mathrm{und}}(\tilde{K}_{2,2})$ is a continuous function of γ that satisfies

$$R_{1,2}^{\mathrm{und}}|_{\gamma=1} = R_{1,2}^{\mathrm{und}}(K_{2,2}) > R_1^\star > R_{1,2}^{\mathrm{und}}(\Sigma^\star) = R_{1,2}^{\mathrm{und}}|_{\gamma=0}. \tag{46}$$

Thus, by choosing λ appropriately, we construct either an admissible matrix that yields a higher secondary rate or

a matrix yielding the same secondary rate, and such that (40c) is satisfied with equality.

We now consider the case $K_{2,1} \ne 0$. Construct the following two covariance matrices

$$\tilde{K}_{2,1} = (1 - \gamma)K_{2,1}, \tag{47}$$

$$\tilde{K}_{2,2} = K_{2,2} + \gamma K_{2,1} \tag{48}$$

for $\gamma \in [0, 1]$. Note that by construction, both $\tilde{K}_{2,1}$ and $\tilde{K}_{2,2}$ are positive semi-definite. Moreover, this choice of covariance matrices satisfies

$$\tilde{K}_{2,1} + \tilde{K}_{2,2} = K_{2,1} + K_{2,2}, \tag{49}$$

and thus the constraints (40b), (40d) and (40e) are satisfied, and the first term in the objective function (40a) remains unchanged. However, noting that

$$\frac{|A + B + C|}{|B + C|} \le \frac{|A + B|}{|B|} \tag{50}$$

for $A \succeq 0, C \succeq 0$ and $B \succ 0$, we see that

$$R_{1,2}^{\mathrm{und}}(K_{2,2}) \ge R_{1,2}^{\mathrm{und}}(\tilde{K}_{2,2}) \tag{51}$$

for any $\gamma \in [0, 1]$. Moreover, $R_{1,2}^{\mathrm{und}}(\tilde{K}_{2,2})$ is a non-increasing and continuous function of γ. If, for any $\gamma \in (0, 1]$, we have that

$$R_{1,2}^{\mathrm{und}}(K_{2,2}) > R_{1,2}^{\mathrm{und}}(\tilde{K}_{2,2}) \ge R_1^\star, \tag{52}$$

then we have contradicted our initial hypothesis. Otherwise, by the non-increasing property, the pair of matrices $\tilde{K}_{2,2} = K_{2,2} + K_{2,1}$ and $\tilde{K}_{2,1} = 0$ (i.e. $\gamma = 1$) must also be a valid solution. Thus, we can use the first part of the proof to show that there is no loss of generality in restricting (40c) to be an equality. $\square$

We now proceed to prove Proposition 1.

Proof of Proposition 1. The proof for case 1 follows from the fact that it is not possible for the secondary receiver to decode the primary message (for the case of equality in (11), any $K_{2,2} \ne 0$ would render decoding of the primary message impossible). Thus, the best that the transmitter can do is to choose the covariance matrix that maximizes (12). The formulation in (13) follows by noting that the denominator in (12) is independent from the covariance matrix.

The proof for case 2 follows easily by noting that the solution to (16) is the best the secondary system can do given the power and interference constraints.

To prove the solution for case 3, we make use of Lemma 1 to rewrite the optimization problem in (40) as

$$\max_{K_{2,1},K_{2,2}} \left| I + H_{22}^H(K_{2,1} + K_{2,2})H_{22} + H_{12}^H K_1 H_{12} \right| - R_1^\star \tag{53a}$$

subject to:

$$\text{tr}\{h_{21}^H(K_{2,1} + K_{2,2})h_{21}\} \le P_{\text{int}}^{\text{und}}, \tag{53b}$$

$$R_{1,2}^{\text{und}} = R_1^\star, \tag{53c}$$

$$\text{tr}\{K_{2,1} + K_{2,2}\} \le P_2, \tag{53d}$$

$$K_{2,1} \succeq 0, K_{2,2} \succeq 0. \tag{53e}$$

Note that only the first term in the objective function is relevant for the optimization. Moreover, except for (53c), the maximization only depends on $K_{2,1}, K_{2,2}$ through their sum, which we denote by $\mathbf{\Delta}$. The general solution $(K_{2,1}, K_{2,2})$ can be obtained by computing the optimal $\mathbf{\Delta}^\star$ disregarding constraint (53c) and then setting

$$K_{2,1} = \gamma \mathbf{\Delta}^\star, \tag{54}$$

$$K_{2,2} = (1 - \gamma)\mathbf{\Delta}^\star, \tag{55}$$

with $\gamma \in [0, 1]$, such that $R_{1,2}^{\text{und}} = R_1^\star$. Note that such γ must exist because $R_{1,2}^{\text{und}}$ is continuous in γ and

$$R_{1,2}^{\text{und}}|_{\gamma=1} < R_1^\star < R_{1,2}^{\text{und}}|_{\gamma=0}, \tag{56}$$

by assumption for case 3. □

Appendix 2
Proof of proposition 2
We shall make use of the following well-known Lemma in our arguments:

Lemma 2. *The function*

$$\beta \log \left| I + \frac{1}{\beta} B^H C B \right| \tag{57}$$

defined for $\beta \in (0, 1]$, any B and any $C \succeq 0$ (with appropriate dimensions) is strictly increasing in β.

Proof. We have that

$$\beta \log \left| I + \frac{1}{\beta} B^H C B \right| = \sum_{i=1}^{r} \beta \log \left(1 + \frac{\lambda_i}{\beta}\right), \tag{58}$$

where λ_i and r are the singular values and the rank of $B^H C B$, respectively. It is easy to check that the first derivative of each of the terms in the sum is positive for $\beta > 0$, proving that (57) is strictly increasing in β. □

Proof of proposition 2. First, we prove statement 1 by contradiction. Assume that the set of parameters that attains the optimum satisfies

$$\text{tr}\{K_p + K_r\} < P_2. \tag{59}$$

Consider two new covariance matrices

$$\tilde{K}_p = \gamma_p K_p, \tag{60}$$

$$\tilde{K}_r = \gamma_r K_r. \tag{61}$$

Since $R_1^{(2)}$ is a continuous function of both $\text{tr}\{K_p\}$ and $\text{tr}\{K_r\}$, we can find (sufficiently small) $\gamma_p > 1$ and $\gamma_r > 1$ that do not violate constraint (26d) and such that $R_1^{(2)}$ evaluated for $\tilde{K}_p$ and $\tilde{K}_r$ remains unchanged (and hence satisfy (26c)). However, using $\tilde{K}_p$ yields a larger secondary rate R_2, which contradicts our assumption that the set of parameters solved the optimization problem.

We now prove statement 2 also by contradiction. Assume that the optimal choice of parameters yields

$$\alpha \, \text{tr}\{K_1^{(1)}\} + (1 - \alpha) \, \text{tr}\{K_1^{(2)}\} < P_1, \tag{62}$$

where $K_1^{(1)}$ is the optimal choice of covariance matrix. Now, define the matrix $\tilde{K}_1^{(1)} = \gamma K_1^{(1)}$ for some $\gamma > 1$, such that

$$\alpha \, \text{tr}\{\tilde{K}_1^{(1)}\} + (1 - \alpha) \, \text{tr}\{K_1^{(2)}\} \le P_1. \tag{63}$$

This choice of matrix yields

$$\tilde{R}_1^{(1)} \triangleq \alpha \log \left(1 + h_{11}^H \tilde{K}_1^{(1)} h_{11}\right) \tag{64}$$

$$= \alpha \log \left(1 + \gamma h_{11}^H K_1^{(1)} h_{11}\right) \tag{65}$$

$$> \alpha \log \left(1 + h_{11}^H K_1^{(1)} h_{11}\right) \tag{66}$$

$$= R_1^{(1)} \tag{67}$$

and

$$\tilde{R}_t \triangleq \alpha \log \left| I + H_t^H \tilde{K}_1^{(1)} H_t \right| \tag{68}$$

$$= \alpha \log \left| I + \gamma H_t^H K_1^{(1)} H_t \right| \tag{69}$$

$$= \alpha \sum_{i=1}^{r} \log(1 + \gamma \lambda_i) \tag{70}$$

$$> \alpha \sum_{i=1}^{r} \log(1 + \lambda_i) \tag{71}$$

$$= \alpha \log \left| I + H_t^H K_1^{(1)} H_t \right| \tag{72}$$

$$= R_t, \tag{73}$$

where λ_i and r are the singular values and the rank of $H_t K_1^{(1)} H_t^H$, respectively. Thus, we have that

$$\tilde{R}_1^{(1)} + R_1^{(2)} > R_1^\star \tag{74}$$

$$\tilde{R}_t > R_1^\star, \tag{75}$$

and we can find a shorter duration of the first phase $\tilde{\alpha} < \alpha$ such that the rates, evaluated at $\tilde{\alpha}$, satisfy

$$\tilde{R}_1^{(1)}(\tilde{\alpha}) + R_1^{(2)}(\tilde{\alpha}) \ge R_1^\star, \tag{76}$$

$$\tilde{R}_t(\tilde{\alpha}) \ge R_1^\star. \tag{77}$$

At the same time, we have increased the secondary rate by Lemma 2, thus contradicting our hypothesis on the optimality of the set of parameters.

Appendix 3
Proof of Proposition 3
Assume that the set of parameters that attains the maximum in (26) satisfies

$$R_1^{(1)}(K_1^{(1)}) + R_1^{(2)} > R_1^\star, \tag{78}$$

$$R_t(K_1^{(1)}) \geq R_1^\star, \tag{79}$$

where $K_1^{(1)}$ is the optimal covariance matrix. The notation remarks the dependency of $R_1^{(1)}$ and R_t on the covariance matrix $K_1^{(1)}$. Let $\sigma^\star$ denote the power used by this covariance matrix, i.e. $\sigma^\star \triangleq \text{tr}\{K_1^{(1)}\}$. We divide the proof into two cases.

First, consider the case $K_1^{(1)} \neq K^{\text{WF}}(\sigma^\star)$ with $K^{\text{WF}}(\sigma^\star)$ as defined in (27). Both $R_1^{(1)}$ and R_t are continuous functions of the entries of the covariance matrix, and the log-det operator is concave on the set of Hermitian positive semi-definite matrices with bounded trace. Therefore, we can find a Hermitian positive semi-definite covariance matrix $\tilde{K}_{11}$, with $\|\tilde{K}_{11} - K_1^{(1)}\|$ small enough such that

$$R_1^{(1)}(\tilde{K}_{11}) + R_1^{(2)} > R_1^\star, \tag{80}$$

$$R_t(\tilde{K}_{11}) > R_1^\star. \tag{81}$$

Now, since $R_1^{(1)}$, R_t, and $R_1^{(2)}$ are all continuous in α, we can find a shorter duration for the first phase, i.e. $\tilde{\alpha} < \alpha$, such that the two constraints are still satisfied. However, by Lemma 2 in Appendix 2, shortening the first phase strictly increases the secondary rate R_2, contradicting our assumption on the optimality of the set of parameters.

In the case where $K_1^{(1)} = K^{\text{WF}}$, the rate R_t is already maximum. In this case, if either $K_1^{(2)} \neq 0$ or $K_r \neq 0$, we can use similar arguments to those used in the proof of Proposition 2 to arrive at a contradiction. In contrast, if $K_1^{(2)} = 0$ and $K_r = 0$, we cannot always ensure that (26c) is satisfied with equality. However, in the cases where we cannot reach a contradiction, we can use that $R_1^{(2)} = 0$ and $\text{tr}\{K_1^{(1)}\} = \frac{P_1}{\alpha}$ (cf. Proposition 2). Combined with the fact that the solution to (26) must satisfy $R_1^{(1)} \geq R_1^\star$, we can show that

$$\alpha \log\left(1 + h_{11}^H K^{\text{WF}} h_{11}\right) \geq \alpha \log\left|I + H_t^H K^{\text{WF}} H_t\right|, \tag{82}$$

thus violating the cooperation condition.

Appendix 4
Proof of Proposition 4
We prove the first part of the claim by contradiction. Assume that the optimal choice of parameters yields

$$R_t = \alpha \log\left|I + H_t^H K_1^{(1)} H_t\right| > R_1^\star, \tag{83}$$

where $K_1^{(1)}$ is the optimal covariance matrix. Note that we can express $K_1^{(1)}$ as

$$K_1^{(1)} = \Pi_{h_{11}} K_1^{(1)} + \Pi_{h_{11}}^\perp K_1^{(1)} \tag{84}$$

$$= \beta_1 \Sigma_1 + \beta_2 \Sigma_2, \tag{85}$$

where $\beta_1 = \|\Pi_{h_{11}} K_1^{(1)}\|, \beta_2 = \|\Pi_{h_{11}}^\perp K_1^{(1)}\|, \Sigma_1 = \beta_1^{-1} \Pi_{h_{11}} K_1^{(1)}$, and $\Sigma_2 = \beta_2^{-1} \Pi_{h_{11}}^\perp K_1^{(1)}$ for i $\in \{1, 2\}$ with $\beta_i > 0$. Otherwise, set $\Sigma_i = 0$ for i such that $\beta_i = 0$. Assuming $\beta_i > 0$ for i $\in \{1, 2\}$, both Σ_1 and Σ_2 have unit norm. Now, let

$$K_\parallel = \frac{h_{11} h_{11}^H}{\|h_{11}\|^2}. \tag{86}$$

Note that $K_\parallel = \Pi_{h_{11}}$. Thus, we have

$$\Pi_{h_{11}} K_\parallel = K_\parallel, \tag{87}$$

$$\Pi_{h_{11}}^\perp K_\parallel = 0. \tag{88}$$

Now define a new matrix

$$\tilde{K}_1^{(1)} = \gamma K_1^{(1)} + \epsilon K_\parallel = \gamma \beta_1 \Sigma_1 + \gamma \beta_2 \Sigma_2 + \epsilon K_\parallel, \tag{89}$$

where $\epsilon = (1 - \gamma)(\beta_1 + \beta_2)$. Note that $\tilde{K}_1^{(1)}$ is a valid choice of covariance matrix because it is the sum of positive semi-definite Hermitian matrices and satisfies $\text{tr}\{\tilde{K}_1^{(1)}\} = \text{tr}\{K_1^{(1)}\}$. Since the determinant is a continuous function of the entries of the matrix, and the logarithm is a continuous function of its argument, we can find $0 < \gamma < 1$ such that

$$\tilde{R}_t \triangleq \alpha \log\left|I + H_t^H \tilde{K}_1^{(1)} H_t\right| > R_1^\star. \tag{90}$$

This choice of $\tilde{K}_1^{(1)}$ yields

$$\tilde{R}_1^{(1)} \triangleq \alpha \log\left(1 + h_{11}^H \tilde{K}_1^{(1)} h_{11}\right) \tag{91}$$

$$= \alpha \log\left(1 + h_{11}^H \left(\gamma \beta_1 \Sigma_1 + \epsilon K_\parallel\right) h_{11}\right) \tag{92}$$

$$= \alpha \log\left(1 + \gamma \beta_1 h_{11}^H \Sigma_1 h_{11} + \epsilon h_{11}^H K_\parallel h_{11}\right) \tag{93}$$

$$\geq \alpha \log\left(1 + \gamma \beta_1 h_{11}^H \Sigma_1 h_{11} + \epsilon h_{11}^H \Sigma_1 h_{11}\right) \tag{94}$$

$$= \alpha \log\left(1 + (\beta_1 + \beta_2(1 - \gamma)) h_{11}^H \Sigma_1 h_{11}\right) \tag{95}$$

$$> \alpha \log\left(1 + \beta_1 h_{11}^H \Sigma_1 h_{11}\right) \tag{96}$$

$$= R_1^{(1)}. \tag{97}$$

The inequality in (94) is due to the fact that

$$h_{11}^H K_\parallel h_{11} \geq h_{11}^H \Sigma_1 h_{11} \geq 0. \tag{98}$$

The inequality in (96) follows if $\beta_2 > 0$ by the fact that $0 < \gamma < 1$. Hence, for this new choice of covariance matrix $\tilde{K}_1^{(1)}$, we have

$$\tilde{R}_1^{(1)} + R_1^{(2)} > R_1^\star, \tag{99}$$
$$\tilde{R}_t > R_1^\star. \tag{100}$$

Now, we can find a shorter duration of the first phase $\tilde{\alpha} < \alpha$, such that the rates evaluated at $\tilde{\alpha}$ satisfy

$$\tilde{R}_1^{(1)}(\tilde{\alpha}) + R_1^{(2)}(\tilde{\alpha}) \geq R_1^\star, \tag{101}$$
$$\tilde{R}_t(\tilde{\alpha}) \geq R_1^\star. \tag{102}$$

At the same time, we have increased the secondary rate by Lemma 2 in Appendix 2, thus contradicting our hypothesis on the optimality of the set of parameters.

Finally, note that $\beta_2 = 0$ implies that

$$\Pi_{h_{11}} K_1^{(1)} = K_1^{(1)}, \tag{103}$$

so that $K_1^{(1)}$ is a Hermitian rank-one covariance matrix. Therefore, we must have

$$K_1^{(1)} = \rho \frac{h_{11} h_{11}^H}{\|h_{11}\|^2}, \tag{104}$$

for some $\rho \in \mathbb{R}$. This concludes the proof.

Appendix 5
Proof of Corollary 1
Assume that $K_1^{(1)}$ is the optimal covariance matrix in (26), and let $\hat{K}_1^{(1)}$ be the output of Algorithm 1. Note that by construction of the algorithm $\text{tr}\{\hat{K}_1^{(1)}\} = \text{tr}\{K_1^{(1)}\}$. We divide the proof into two parts:

If $K_1^{(1)} = \rho \frac{h_{11} h_{11}^H}{\|h_{11}\|^2}$ for some $\rho \in \mathbb{R}$ (i.e. it corresponds to the MRT beamformer to receiver 1), then trivially $\hat{K}_1^{(1)} = K_1^{(1)}$ as this is the initial guess of the algorithm (lines 7 and 8) and it satisfies

$$\alpha \log \left| I + H_t^H \hat{K}_1^{(1)} H_t \right| \geq R_1^\star. \tag{105}$$

Thus, this is the output of the algorithm (lines 9 and 10).

For the case when $K_1^{(1)}$ does not correspond to the MRT beamformer, we prove the optimality of the algorithm by contradiction. Assume $\hat{K}_1^{(1)} \neq K_1^{(1)}$ and note that

$$\alpha \log \left| I + H_t^H K_1^{(1)} H_t \right| = R_1^\star, \tag{106}$$
$$\alpha \log \left| I + H_t^H \hat{K}_1^{(1)} H_t \right| = R_1^\star. \tag{107}$$

The equality in (106) comes from Proposition 4 and the fact that $K_1^{(1)}$ is the optimal covariance matrix. The equality in (107) is ensured by construction of the algorithm in the limit of arbitrary numerical precision in the bisection

method, i.e. $\epsilon \rightarrow 0$ (lines 9 to 17 in Algorithm 2). In addition, we have

$$h_{11}^H \hat{K}_1^{(1)} h_{11} > h_{11}^H K_1^{(1)} h_{11} \tag{108}$$

because by construction, Algorithm 1 finds the matrix with largest component in the direction of h_{11} that satisfies (26b) with equality. Thus,

$$R_1^{(1)}(\hat{K}_1^{(1)}) + R_1^{(2)} > R_1^\star, \tag{109}$$
$$R_t(\hat{K}_1^{(1)}) = R_1^\star. \tag{110}$$

We can now proceed as in Proposition 3 to contradict our initial hypothesis on the optimality of $K_1^{(1)}$. Thus, we must have $\hat{K}_1^{(1)} = K_1^{(1)}$ in this case as well.

Appendix 6
Proof of Proposition 5
The matrix K_{co} and its sub-matrices $K_1^{(2)}, K_r$ and Ψ only appear in the expression for $R_1^{(2)}$ through the expression

$$h_{\text{ext}}^H K_{\text{co}} h_{\text{ext}}. \tag{111}$$

It is easy to see that the optimal K_{co} has rank 1, i.e. $K_{\text{co}} = v_{\text{co}} v_{\text{co}}^H$. The vector v_{co} is chosen as to maximize the projection $v_{\text{co}}^H h_{\text{ext}}$ while satisfying the constraints on the traces of $K_1^{(2)}$ and K_r. Simple calculus shows that the optimal v_{co} is given, up to a common factor, by

$$v_{\text{co}} = \begin{bmatrix} \sqrt{P_1^{(2)}} \frac{h_{11}}{\|h_{11}\|} \\ \sqrt{P_r} \frac{h_{21}}{\|h_{21}\|} \end{bmatrix}. \tag{112}$$

The desired $K_1^{(2)}, K_r$ and Ψ are readily obtained from K_{co}.

Using these results, it is straightforward to establish the identity

$$h_{\text{ext}}^H K_{\text{co}} h_{\text{ext}} = \left(\|h_{11}\| \sqrt{P_1^{(2)}} + \|h_{21}\| \sqrt{\frac{P_r}{1-\alpha}} \right)^2. \tag{113}$$

From (113), we see that the effect of Ψ is to correlate the primary and secondary transmissions so that their signals add constructively at the receiver. Finally, given the matrices $K_1^{(2)}, K_r$ and Ψ, the characterization of K_p in terms of the concave problem in (32) follows immediately (see [20] as well).

Appendix 7
Proof of Corollary 2
The beamformer w_p appears both in the objective function (26a) and in constraint (26c) through $R_1^{(2)}$. First, note that if $P_{\text{int}} < 0$, the problem has no valid solution. For a given second phase (that is, given α and $P_1^{(2)}$), using Propositions 2 and 3, the optimization problem is reduced

to finding w_p and P_r and can be reformulated, for $P_{int} > 0$, as

$$\max_{P_r, w_p} \min \left\{ \frac{|h_{22}^H w_p|^2}{|h_{21}^H w_p|^2} P_{int}, |h_{22}^H w_p|^2 (P_2 - P_r) \right\}, \quad (114)$$

with $0 \leq P_r \leq P_2$, $\|w_p\| = 1$. For a fixed P_r, the objective function (114) is monotonically increasing in $|h_{22}^H w_p|^2$ and monotonically decreasing in $|h_{21}^H w_p|^2$. Thus, for given P_r, the optimal beamformer w_p can be parametrized as

$$w_p(\lambda) = \sqrt{\lambda} \frac{\Pi_{h_{21}} h_{22}}{\|\Pi_{h_{21}} h_{22}\|} + \sqrt{1 - \lambda} \frac{\Pi_{h_{21}}^\perp h_{22}}{\|\Pi_{h_{21}}^\perp h_{22}\|} \quad (115)$$

for some $\lambda \in [0, 1]$. Using this parametrization, we define $f(\lambda) \triangleq |h_{22}^H w_p(\lambda)|^2$ and note that $|h_{21}^H w_p|^2 = \lambda \|h_{21}\|^2$ to write, for fixed P_r, the optimization problem as

$$\max_\lambda f(\lambda) \min \left\{ \frac{P_{int}}{\lambda \|h_{21}\|^2}, P_2 - P_r \right\}. \quad (116)$$

The function $f(\lambda)$ is unimodal with maximum value attained for $\lambda = \lambda_{MRT}$, that is, when $w_p(\lambda)$ is in the direction of h_{22} (i.e. MRT). Thus, if

$$\lambda_{MRT} \leq \frac{P_{int}}{(P_2 - P_r) \|h_{21}\|^2}, \quad (117)$$

then $\lambda = \lambda_{MRT}$ yields the optimum value. Otherwise, the basic calculus shows that (114) is maximized for

$$\lambda = \frac{P_{int}}{(P_2 - P_r) \|h_{21}\|^2}. \quad (118)$$

Using this parametrization, we can find the optimal beamformer by varying P_r from P_2 to 0 to find the maximum value of (114).

For $P_{int} = 0$, the primary receiver is already at the limit of decodability and cannot tolerate any interference. Thus, the secondary transmitter must transmit in the ZF direction. This special case is already considered by our parametrization (i.e. setting $\lambda = 0$).

Author details
[1]ACCESS Linnaeus Centre, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. [2]Communications Theory, Communications Laboratory, Dresden University of Technology, Dresden, Germany.

References
1. S Srinivasa, S Jafar, Cognitive radios for dynamic spectrum access - the throughput potential of cognitive radio: a theoretical perspective. IEEE Commun. Mag. **45**(5), 73–79 (2007)
2. A Goldsmith, S Jafar, I Marić, S Srinivasa, Breaking spectrum gridlock with cognitive radios: an information theoretic perspective. Proc. IEEE **97**(5), 894–914 (2009)
3. S Haykin, Cognitive radio: brain-empowered wireless communications. IEEE J. Selected Areas Commun. **23**(2), 201–220 (2005)
4. G Scutari, D Palomar, J Pang, F Facchinei, Flexible design of cognitive radio wireless systems. IEEE Signal Proc. Mag. **26**(5), 107–123 (2009)
5. W Zhang, U Mitra, Spectrum shaping: a new perspective on cognitive radio-part I: coexistence with coded legacy transmission. IEEE Trans. Commun. **58**(6), 1857–1867 (2010)
6. A Carleial, Interferencechannels. IEEE Trans. Inf. Theory **24**, 60–70 (1978)
7. TS Han, K Kobayashi, A new achievable rate region for the interference channel. IEEE Trans. Inf. Theory **27**, 49–60 (1981)
8. W Wu, S Vishwanath, A Arapostathis, Capacity of a class of cognitive radio channels: interference channels with degraded message sets. IEEE Trans. Inf. Theory **53**(11), 4391–4399 (2007)
9. A Jovičić, P Viswanath, Cognitive radio: an information-theoretic perspective. IEEE Trans. Inf. Theory **55**(9), 3945–3958 (2009)
10. I Marić, RD Yates, G Kramer, Capacity of interference channels with partial transmitter cooperation. IEEE Trans. Inf. Theory **53**(10), 3536–3548 (2007)
11. S Haykin, D Thomson, J Reed, Spectrum sensing for cognitive radio. Proc. IEEE **97**(5), 849–877 (2009)
12. S Yiu, M Vu, V Tarokh, Interference and noise reduction by beamforming in cognitive networks. IEEE Trans. Commun. **57**(10), 3144–3153 (2009)
13. R Zhang, YC Liang, Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks. IEEE J. Selected Top. Signal Proc. **2**, 88–102 (2008)
14. J Lv, EA Jorswieck, Spatial shaping in cognitive system with coded legacy transmission. Paper presented at the international ITG workshop on smart antennas (WSA), Aachen, Germany, 24–25 February 2011.
15. A Tajer, N Prasad, X Wang, Beamforming and rate allocation in MISO cognitive radio networks. IEEE Trans. Signal Proc. **58**, 362–377 (2010)
16. E Jorswieck, E Larsson, D Danev, Complete characterization of the Pareto boundary for the MISO interference channel. IEEE Trans. Signal Proc. **56**(10), 5292–5296 (2008)
17. R Mochaourab, E Jorswieck, Optimal beamforming in interference networks with perfect local channel information. IEEE Trans. Signal Proc. **59**(3), 1128–1141 (2011)
18. S Sridharan, S Vishwanath, On the capacity of a class of MIMO cognitive radios. IEEE J. Sel. Top. Sign. Proces. **2**, 103–117 (2008)
19. U Salim, Achievable rate regions for cognitive radio Gaussian fading channels with partial CSIT. Paper presented at IEEE workshop on signal processing advances in wireless communications (SPAWC). San Francisco, CA, USA, 26–29 June 2011.
20. C Huppert, A Sezgin, On beamforming for overlay cognitive multiple antenna systems. Paper presented at IEEE 8th international symposium on wireless communication systems (ISWCS). Aachen, Germany, 6–9 November 2011.
21. J Lv, R Blasco-Serrano, E Jorswieck, R Thobaben, A Kliks, Optimal beamforming in MISO cognitive channels with degraded message sets. Paper presented at IEEE wireless communications and networking conference (WCNC), Paris, France, 1–4 April 2012.
22. J Lv, E Jorswieck, R Blasco-Serrano, R Thobaben, A Kliks, Linear precoding in MISO cognitive channels with degraded message sets. Paper presented at international ITG workshop on smart antennas (WSA), Dresden, Germany, 7–8 March 2012
23. RA Tannious, A Nosratinia, Cognitive radio protocols based on exploiting hybrid ARQ retransmissions. IEEE Trans. Wireless Commun. **9**(9), 2833–2841 (2010)
24. N Michelusi, O Simeone, M Levorato, P Popovski, M Zorzi, Optimal cognitive transmission exploiting redundancy in primary ARQ process. Paper presented at workshop on information theory and applications (ITA), La Jolla, CA, USA, 6–11 February 2011.
25. Y Li, P Wang, D Niyato, Optimal power allocation for secondary users in cognitive relay networks. Paper presented at IEEE wireless communications and networking conference (WCNC) Cancun, Quintana Roo, Mexico, 28–31 March 2011

26. T Cover, A El Gamal, Capacity theorems for the relay channel. IEEE Trans. Inf. Theory **25**(5), 572–584 (1979)

27. A Høst-Madsen, J Zhang, Capacity bounds and power allocation for wireless relay channels. IEEE Trans. Inf. Theory **51**(6), 2020–2040 (2005)

28. B Wang, J Zhang, A Høst-Madsen, On the capacity of MIMO relay channels. IEEE Trans. Inf. Theory **51**, 29–43 (2005)

29. KB Letaief, W Zhang, Cooperative communications for cognitive radio networks. Proc. IEEE **97**(5), 878–893 (2009)

30. Y Han, A Pandharipande, S Ting, Cooperative decode-and-forward relaying for secondary spectrum access. IEEE Trans. Wireless Commun. **8**(10), 4945–4950 (2009)

31. EH Shin, D Kim, Time and power allocation for collaborative primary-secondary transmission using superposition coding. IEEE Commun. Lett. **15**(2), 196–198 (2011)

32. L Li, FA Khan, M Pesavento, T Ratnarajah, Power allocation and beamforming in overlay cognitive radio systems. Paper presented at IEEE 73rd vehicular technology conference (VTC), Budapest, Hungary, 15–18 May 2011

33. R Manna, RH Louie, Y Li, B Vucetic, Cooperative spectrum sharing in cognitive radio networks with multiple antennas. IEEE Trans. Signal Proc. **59**(11), 5509–5522 (2011)

34. Y Zou, YD Yao, B Zheng, Cognitive transmissions with multiple relays in cognitive radio networks. IEEE Trans. Wireless Commun. **10**(2), 648–659 (2011)

35. Y Zou, YD Yao, B Zheng, Cooperative relay techniques for cognitive radio systems: spectrum sensing and secondary user transmissions. IEEE Commun. Mag. **50**(4), 98–103 (2012)

36. R Blasco-Serrano, J Lv, R Thobaben, E Jorswieck, A Kliks, M Skoglund, Comparison of underlay and overlay spectrum sharing strategies in MISO cognitive channels. Paper presented at 7th international ICST conference on cognitive radio oriented wireless networks (CROWNCOM), Stockholm, Sweden, 18–20 June 2012

37. D Tse, P Viswanath, *Fundamentals of Wireless Communication*. (Cambridge University Press, UK, 2005)

38. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press, UK, 2004)

39. E E Björnson, Jorswieck, Optimal resource allocation in coordinated multi-cell systems. Found. Trends Commun. Inf. Theory **9**(2–3), 113–381 (2013)

40. D Palomar, J Fonollosa, Practical algorithms for a family of waterfilling solutions. IEEE Trans. Signal Proc. **53**(2), 686–695 (2005)