The Operational Secrecy Capacity of Cognitive Radio MIMO Channel

L. Dong, S. Loyka, Y. Li

***Abstract*—Secure communications over cognitive radio (CR) MIMO channels is studied. The secrecy capacity, defined operationally as the maximum achievable secrecy rate subject to reliability and secrecy constraints, of CR MIMO wiretap channel is established under power and interference constraints, including a number of closed-form expressions, bounds and related properties. The secrecy capacity of this channel can be expressed as a minimax game between the transmitter (who selects the input covariance) and nature (who selects the noise covariance). Neither player can deviate from an optimal strategy without incurring a penalty.**

I. Introduction

Widespread use of wireless systems and services puts ever increasing demand on already overcrowded wireless spectrum. Cognitive radio (CR) approach has recently emerged as a solution to the spectrum scarcity problem by allowing secondary systems to use a resource when not in use by primary spectrum holder or when interference to primary users is not significant [1]-[3].

Due to the broadcast nature of wireless channels, wireless systems are especially vulnerable to various security threads. This is especially true for CR systems due to their open architecture and shared use of the same spectrum by primary and secondary users. A number of possible threads have been identified and studied, including primary user emulation, spectrum sensing data falsification, jamming and eavesdropping [4].

Physical-layer security approach has emerged as a valuable complement to cryptography-based approaches [5]. In this approach, the secrecy of communications is ensured at the physical layer by exploiting the properties of wireless channels to "hide" transmitted information from eavesdropping. Using this approach in combination with multi-antenna (MIMO) systems offers significant opportunities for enhancing the secrecy of wireless communications. The wiretap MIMO channel has emerged as a popular model to establish information-theoretic limits to secure communications [6]-[8]. The key performance metric is the secrecy capacity, defined operationally as the maximum achievable rate subject to reliability (low error probability) and secrecy (low information leakage) criteria [5], which is a counterpart of the regular channel capacity (without the secrecy criterion). The secrecy capacity of the AWGN MIMO wiretap channel (WTC) has been established in [7][8], where in particular the optimality of Gaussian signaling has been shown while leaving the problem of optimal covariance matrix open. This problem has been solved for a number of special cases in e.g. [9][10].

In this paper, we extend the classical AWGN MIMO wiretap channel model to the cognitive radio setting by adding an interference constraint so that any signalling must ensure that the interference generated to the primary receiver (PR) does not exceed a ceratin threshold. This significantly changes the problem as the feasible set (of admissible transmit covariance matrices) is *not* isotropic anymore, so that known results (e.g. [6]-[10]) do not apply. It is not even clear whether Gaussian signaling is optimal in such setting to maximize the secrecy rate (recall that it was far from trivial in [7][8] to establish the optimality of Gaussian signalling and some key steps in the proofs exploited the isotropic nature of the feasible set). The key contribution of this paper is to establish the secrecy capacity of the cognitive radio MIMO wiretap channel with AWGN by rigorously demonstrating that Gaussian signalling is still optimal under the interference constraint. We emphasize that the secrecy capacity here is defined operationally as the maximum achievable secrecy rate (subject to reliability and secrecy constraints, in addition to the power and interference constraints), rather than formally as the difference of certain mutual information terms without demonstrating their operational significance (as in e.g. [11]). While the operational secrecy capacity of MISO CR channel was established in [12], the present paper establishes the secrecy capacity and a number of alternative closed-form expressions and properties for the full MIMO case.

Our approach is based on that in [6][7] and extends it to the cognitive radio settings. In particular, while it is rather straightforward to show that the lower and upper bounds to the secrecy capacity in [7] still hold, it is far more challenging to show the key saddle-point property in [7] holds for such non-isotropic set and the upper and lower bounds to the secrecy capacity coincide at saddle point, hence establishing the operational secrecy capacity. The secrecy capacity of the CR MIMO WTC can also be expressed as a minimax game between the transmitter (who selects the input covariance) and nature (who selects the noise covariance); neither player can deviate from the optimal strategy without incurring a penalty.

Notations: bold lower-case and capitals denote vectors and matrices respectively; $\mathbf{A}^{'}$ and $\mathbf{A}^{+}$ denote transpose and conjugate transpose; $\mathbf{A} \geq \mathbf{0}$ means positive semi-definite; $E\{\cdot\}$ is statistical expectation.

L. Dong is with the School of Electronics and Information, Northwestern Polytechnical University, Xian, China, e-mail: dlm nwpu@hotmail.com. This work was done when L.Dong was visiting the School of Electrical Engineering and Computer Science, University of Ottawa, Canada. It was supported by the China Scholarship Council (CSC).

S. Loyka is with the School of Electrical Engineering and Computer Science, University of Ottawa, Canada, e-mail: sergey.loyka@ieee.org.

Y. Li is with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, e-mail: ruikel@nwpu.edu.cn.

978-1-5090-6026-9/17/$31.00 ©2017 Crown

II. Cognitive Radio Wiretap MIMO Channel Model

Let us consider the standard AWGN wiretap channel model where a transmitter (Tx) sends confidential information to a legitimate receiver (Rx) while an eavesdropper (Ev) intercepts the transmission. The objective is to ensure reliable communications between the Tx and Rx (the reliability criterion) while keeping the Ev ignorant about transmitted information (the secrecy criterion). The secrecy capacity is the largest transmission rate subject to the reliability and secrecy criteria [5].

In the discrete-time AWGN MIMO channel model, the signals received by the Rx and the Ev can be expressed as

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \boldsymbol{\xi}_1, \quad \mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \boldsymbol{\xi}_2 \tag{1}$$

where $\mathbf{y}_1, \mathbf{y}_2$ are the respective received signals, $\mathbf{x}$ is the transmitted signal, $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ represent zero-mean unit-variance i.i.d. noise at the Rx and Ev end respectively; $\mathbf{H}_1, \mathbf{H}_2$ are the channel matrices collecting channel gains from the Tx to the Rx and Ev respectively. We assume that the Tx has m antennas, while the Rx and Ev have n_1 and n_2 antennas. In addition to this, following the CR model, there is a primary receiver (PR) whose received signal is

$$\mathbf{y}_3 = \mathbf{H}_3\mathbf{x} + \boldsymbol{\xi}_3 \tag{2}$$

where $\mathbf{H}_3$ and $\boldsymbol{\xi}_3$ are the channel matrix and noise of the PR. We assume that full channel state information (CSI) is available at all ends of the links.

In the CR setting, the transmission is subject to power and interference constraints, so that any Tx covariance matrix $\mathbf{R} = E\left\{\mathbf{x}\mathbf{x}^+\right\}$ must be in the following feasible set $S_\mathbf{R}$:

$$S_\mathbf{R} = \left\{\mathbf{R} : tr(\mathbf{R}) \le P_T,\ tr(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) \le P_I,\ \mathbf{R} \ge \mathbf{0}\right\}, \tag{3}$$

where $P_T,\ P_I$ are the maximum Tx and interference powers. The interference constraint $tr(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) \le P_I$ ensures that the total interference power at the PR does not exceed the threshold P_I so that its performance is not distorted. The secrecy capacity of the cognitive radio WTC is defined operationally as the largest achievable rate subject to the power, reliability and interference constraints simultaneously.

Without interference constraint (i.e. when $P_I = \infty$), the secrecy capacity C_s of the Gaussian MIMO WTC has been established in [6]-[8]:

$$C_s = \max_{\mathbf{R}\ge 0} R_-(\mathbf{R}) \text{ s.t. } tr\mathbf{R} \le P_T \tag{4}$$

where

$$R_-(\mathbf{R}) = \ln|\mathbf{I} + \mathbf{H}_1\mathbf{R}\mathbf{H}_1^+| - \ln|\mathbf{I} + \mathbf{H}_2\mathbf{R}\mathbf{H}_2^+| \tag{5}$$

It is the purpose of this paper is to extend this result and to establish the secrecy capacity of the cognitive radio Gaussian MIMO wiretap channel. This task is complicated by the fact that the interference constraint in (3) makes the set $S_\mathbf{R}$ non-isotropic in general while the feasible set in (4) is always isotropic and this isotropy was exploited in [7][8] while establishing the secrecy capacity. In particular, this is critical while establishing a saddle-point in [7] and the equality of upper and lower bounds under this saddle point.

Our proof follows the same general path as in [6][7] (developed without the interference constraint) by establishing lower and upper bounds to the capacity, existence of a saddle point and demonstrating that the best lower and upper bounds coincide, all under the new interference constraint. We emphasize that this establishes the operational secrecy capacity (defined as the largest achievable secrecy rate) rather than just information capacity defined formally via the difference of respective mutual information terms (as in (5)).

III. Secrecy Capacity of Cognitive Radio MIMO WTC

In this section, we establish the operational secrecy capacity of cognitive radio MIMO wiretap channel with AWGN. To this end, let $S_\mathbf{K}$ be a covariance matrix of the form

$$S_\mathbf{K} \triangleq \left\{\mathbf{K} : \mathbf{K} = \begin{bmatrix} \mathbf{I} & \mathbf{N} \\ \mathbf{N}^+ & \mathbf{I} \end{bmatrix},\ \mathbf{K} > \mathbf{0}\right\}, \tag{6}$$

where $\mathbf{N}$ is any matrix of appropriate size, and

$$R_+(\mathbf{R}, \mathbf{K}) \triangleq \ln|\mathbf{I} + \mathbf{K}^{-1}\mathbf{H}\mathbf{R}\mathbf{H}^+| - \ln|\mathbf{I} + \mathbf{H}_2\mathbf{R}\mathbf{H}_2^+|, \tag{7}$$

where $\mathbf{H} = [\mathbf{H}_1^+, \mathbf{H}_2^+]^+$ is an extended channel. It will be seen later that $\mathbf{K}$ is the covariance matrix of $\boldsymbol{\xi} = [\boldsymbol{\xi}_1^+, \boldsymbol{\xi}_2^+]^+$ and $\mathbf{N}$ is the covariance matrix of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ for an equivalent degraded channel (where these noise vectors are allowed to be correlated with each other), $\mathbf{N} = E\left\{\boldsymbol{\xi}_1\boldsymbol{\xi}_2^+\right\}$.

Theorem 1. *The operational secrecy capacity of Gaussian MIMO wiretap cognitive radio channel in* (1)-(3) *is*

$$C = \max_{\mathbf{R}\in S_\mathbf{R}} R_-(\mathbf{R}) = \max_{\mathbf{R}\in S_\mathbf{R}} \min_{\mathbf{K}\in S_\mathbf{K}} R_+(\mathbf{R}, \mathbf{K}) \tag{8}$$

Furthermore, the following saddle-point property holds:

$$\max_{\mathbf{R}\in S_\mathbf{R}} \min_{\mathbf{K}\in S_\mathbf{K}} R_+(\mathbf{R}, \mathbf{K}) = \min_{\mathbf{K}\in S_\mathbf{K}} \max_{\mathbf{R}\in S_\mathbf{R}} R_+(\mathbf{R}, \mathbf{K}) \tag{9}$$

so that

$$R_+(\mathbf{R}, \mathbf{K}^*) \le R_+(\mathbf{R}^*, \mathbf{K}^*) \le R_+(\mathbf{R}^*, \mathbf{K}) \tag{10}$$

where $(\mathbf{R}^*, \mathbf{K}^*)$ *is a saddle-point.*

Proof. The theorem is proved by a sequence of propositions below. □

While the first equality in (8) can also be established using the approach of [13], the max-min characterization in (8) as well as the saddle point properties in (9) and (10) cannot be established via that approach. The importance of the max-min characterization comes from the fact that the original maximization problem in (8) (1st equality) is not convex so that all powerful tools of convex optimization cannot be used, while the max-min problem in (8) is convex (in each variable) and hence the KKT conditions are sufficient for *global* optimality. Furthermore, numerical algorithms with guaranteed global convergence can be constructed as in [14].

We begin the proof by establishing lower and upper bounds to the secrecy capacity.

Proposition 1. *Let $p_{\mathbf{x}}$ be a probability distribution of input $\mathbf{x}$. The secrecy capacity C of cognitive radio MIMO WTC can be bounded as follows:*

$$\max_{p_{\mathbf{x}} \in \mathcal{P}} [I(\mathbf{x};\mathbf{y}_1) - I(\mathbf{x};\mathbf{y}_2)] \leqslant C \leqslant \max_{p_{\mathbf{x}} \in \mathcal{P}} I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2) \quad (11)$$

where $I(\mathbf{x};\mathbf{y}_{1(2)})$ is the mutual information between $\mathbf{x}$ and $\mathbf{y}_{1(2)}$, and $I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2)$ is the conditional mutual information when $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are jointly Gaussian and the covariance of $[\boldsymbol{\xi}_1^+, \boldsymbol{\xi}_2^+]^+$ is $\mathbf{K}$ (as in (6)*); $\mathcal{P}$ is the set of all distributions $p_{\mathbf{x}}$ that satisfy the power and interference constraints:*

$$\mathcal{P} = \{p_{\mathbf{x}} : E\left\{|\mathbf{x}|^2\right\} \leqslant P_T,\ E\left\{|\mathbf{H}_3\mathbf{x}|^2\right\} \leqslant P_I\}. \quad (12)$$

Proof. See Appendix. □

Note that this proposition does not require $\mathbf{x}$ to be Gaussian. The following proposition establishes the optimality of Gaussian inputs.

Proposition 2. *For each $\mathbf{K} > \mathbf{0}$, the distribution of $\mathbf{x}$ maximizing $I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2)$ in* (11) *is Gaussian.*

Proof. See Appendix. □

Since Gaussian input maximizes the upper bound, $I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2)$ under such input can be expressed as

$$\begin{aligned} I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2) &= \ln|\mathbf{I} + \mathbf{K}^{-1}\mathbf{H}\mathbf{R}\mathbf{H}^+| - \ln|\mathbf{I} + \mathbf{H}_2\mathbf{R}\mathbf{H}_2^+| \\ &= R_+(\mathbf{R},\mathbf{K}) \end{aligned} \quad (13)$$

and $\max_{p_{\mathbf{x}}}$ can be replaced by $\max_{\mathbf{R}}$ on the both sides of (11) (still preserving the inequalities), giving

$$\max_{\mathbf{R}} R_-(\mathbf{R}) \leq C \leq \max_{\mathbf{R}} R_+(\mathbf{R},\mathbf{K}) \quad (14)$$

which holds for any $\mathbf{K}$ and hence

$$\max_{\mathbf{R}} R_-(\mathbf{R}) \leq C \leq \min_{\mathbf{K}} \max_{\mathbf{R}} R_+(\mathbf{R},\mathbf{K}) \quad (15)$$

We further establish the existence of a saddle-point in the minimax problem above, which is essential to establish Propositions 4 and 5 below, on which Proposition 6 depends.

Proposition 3. *The max-min problem in* (8) *has a saddle point solution as in* (9) *and* (10)*.*

Proof. See Appendix. □

Armed with the saddle-point solution, we further establish that it also solves the following entropy maximization problem.

Proposition 4. *Let $h(\mathbf{y})$ be the differential entropy of $\mathbf{y}$ and let $\mathbf{Z}_{12}^*$ be the optimal MMSE weight matrix to estimate $\mathbf{y}_1$ from $\mathbf{y}_2$ at saddle-point $(\mathbf{R}^*,\mathbf{K}^*)$:*

$$\mathbf{Z}_{12}^* = (\mathbf{N}^* + \mathbf{H}_1\mathbf{R}^*\mathbf{H}_2^+)(\mathbf{I} + \mathbf{H}_2\mathbf{R}^*\mathbf{H}_2^+)^{-1} \quad (16)$$

Then,

$$\arg\max_{\mathbf{R} \in S_R} h(\mathbf{y}_1 - \mathbf{Z}_{12}^*\mathbf{y}_2) = \arg\max_{\mathbf{R} \in S_R} R_+(\mathbf{R},\mathbf{K}^*) \quad (17)$$

where $h(\cdot)$ is evaluated under $\mathbf{K} = \mathbf{K}^*$.

Proof. See Appendix. □

Using this Proposition, we establish the following property of the saddle-point, which is needed to prove Proposition 6.

Proposition 5. *If $\mathbf{H}_1 \neq \mathbf{Z}_{12}^*\mathbf{H}_2$, the saddle point $(\mathbf{R}^*,\mathbf{K}^*)$ satisfies*

$$\mathbf{N}^{*+}\mathbf{H}_1\mathbf{S}^* = \mathbf{H}_2\mathbf{S}^* \quad (18)$$

for any full column-rank matrix $\mathbf{S}^*$ such that $\mathbf{S}^*\mathbf{S}^{*+} = \mathbf{R}^*$ (the columns of $\mathbf{S}^*$ are the scaled eigenvectors of $\mathbf{R}^*$ corresponding to strictly-positive eigenvalues).

Proof. See Appendix. □

Using Gaussian $\mathbf{x}$ in (11) and following Proposition 3 and (15), one obtains:

$$R_-(\mathbf{R}^*) \leq C \leq R_+(\mathbf{R}^*,\mathbf{K}^*) \quad (19)$$

The final step is to show that these bounds coincide.

Proposition 6. *The saddle point solution $(\mathbf{R}^*,\mathbf{K}^*)$ satisfies*

$$R_+(\mathbf{R}^*,\mathbf{K}^*) = \begin{cases} R_-(\mathbf{R}^*), & \mathbf{H}_1 \neq \mathbf{Z}_{12}^*\mathbf{H}_2 \\ 0, & \mathbf{H}_1 = \mathbf{Z}_{12}^*\mathbf{H}_2 \end{cases} \quad (20)$$

Proof. follows from Propositions 4 and 5 above - see Appendix for details. □

Combining (19) and (20), Theorem 1 follows. While we considered the case of non-singular $\mathbf{K}$ only, the singular case can be established in a similar way with somewhat more lengthy arguments (using pseudo-inverse instead of the inverse and related projection on the active sub-space only).

IV. Appendix

A. Proof of Proposition 1

The upper bound in (11) is obtained via a genie-aided channel in which the Rx observes $\mathbf{y}_2$ in addition to $\mathbf{y}_1$. Such channel has a larger capacity than the original one and it is degraded at the same time, making the analysis much simpler (since the original Wyner's construction of the converse applies). Furthermore, one can always choose the noises $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ to be jointly Gaussian and correlated with each other (since the secrecy capacity depends on the marginal distributions, not the joint one [5]) and select their cross-covariance as to minimize the upper bound [6]. The details follow below.

Suppose there is a $(2^{nR}, n)$ code for the channel, where n denotes the length of transmission time interval, consists of a message w uniformly distributed over the set $\{1, 2, ..., 2^{nR}\}$, an encoder that maps the message w to the transmitted vector sequence $\{\mathbf{x}(t)\}_{t=1}^n$, and a decoder that maps the received sequence $\{\mathbf{y}(t)\}_{t=1}^n$ to a message estimate $\hat{w}$. Let

$$\begin{aligned} \mathbf{X}^n &= [\mathbf{x}(1), \mathbf{x}(2), ..., \mathbf{x}(n)], \\ \mathbf{Y}_i^n &= [\mathbf{y}_i(1), \mathbf{y}_i(2), ..., \mathbf{y}_i(n)],\ i = 1, 2, \end{aligned} \quad (21)$$

denote the transmitted and received sequence matrix from time interval 1 to n respectively. The reliability and secrecy criteria are as follows: for every $\epsilon > 0$ and n sufficiently large,

$$\Pr(w \neq \hat{w}) \leqslant \epsilon,\ n^{-1}I(w;\mathbf{Y}_2^n) \leqslant \epsilon, \quad (22)$$

while the power and interference constraints are

$$\frac{1}{n}\sum_{i=1}^{n} E\left\{|\mathbf{x}(i)|^2\right\} \leqslant P_T, \tag{23}$$

$$\frac{1}{n}\sum_{i=1}^{n} E\left\{|\mathbf{H}_3\mathbf{x}(i)|^2\right\} \leqslant P_I. \tag{24}$$

We note that (22) implies, from Fano's inequality,

$$n^{-1} I(w;\mathbf{Y}_1^n) \geqslant R - \epsilon_F \tag{25}$$

where $\epsilon_F \rightarrow 0$ as $\epsilon \rightarrow 0$. By combining (22) and (25), we obtain for any $\epsilon' = \epsilon_F + \epsilon > 0$,

$$\begin{aligned} R - \epsilon' &\leqslant n^{-1}[I(w;\mathbf{Y}_1^n) - I(w;\mathbf{Y}_2^n)] \\ &\leqslant \frac{1}{n}\sum_{i=1}^{n} I(\mathbf{x}(i);\mathbf{y}_1(i)|\mathbf{y}_2(i)) \qquad (26) \\ &\leqslant I(\bar{\mathbf{x}}_q;\bar{\mathbf{y}}_{1q}|\bar{\mathbf{y}}_{2q}) \qquad (27) \\ &\leq \max_{p_{\mathbf{x}}\in\mathcal{P}} I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2) \qquad (28) \end{aligned}$$

where (26) and (27) are obtained via the same steps as in [6, Appendix I]. In (27), q is a time-sharing random variable with uniform distribution: $p_q = 1/n$, and $\bar{\mathbf{x}}_q$ is the composite (time-shared) random input whose distribution is the average of those of $\mathbf{x}(1),..,\mathbf{x}(n)$:

$$p_{\bar{\mathbf{x}}_q} = n^{-1}\sum_{i=1}^{n} p_{\mathbf{x}(i)}. \tag{29}$$

Note that, since $\{\mathbf{x}(1),...,\mathbf{x}(n)\}$ satisfy the power and interference constrained in (23) and (24), so is $\bar{\mathbf{x}}_q$, resulting in the desired upper bound in (28). Since this holds for any $\epsilon' > 0$, the upper bound in (11) follows.

To establish the lower bound in (11), note that it is an achievable rate, which follows from the Csiszar-Korner formula (see e.g. [5][6]) by setting $\mathbf{u} = \mathbf{x}$ and using Gaussian input subject to the constraints in (23) and (24), where where $\mathbf{u}$ is the auxiliary random variable. Since Proposition 2 demonstrates that Gaussian input is optimal for the upper bound, one can also use such input for the lower bound.

B. Proof of Proposition 2

After some manipulations, $I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2)$ can be expressed as

$$I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2) = h(\mathbf{y}_1|\mathbf{y}_2) - h(\boldsymbol{\xi}_1|\boldsymbol{\xi}_2) \tag{30}$$

Since the second term in (30) is independent of $\mathbf{x}$, it suffices to establish that $h(\mathbf{y}_1|\mathbf{y}_2)$ is maximized when $\mathbf{x}$ is Gaussian. While Gaussian distribution maximises the differential entropy under covariance constraint, it is not necessarily so for conditional entropy, since it is a difference of 2 differential entropies. To this end, we need the following Lemma of Thomas [15].

Lemma 1. *Let $z_1,..,z_k$ be a set of arbitrary zero-mean random variables with covariance matrix $\mathbf{R}$. Let S be any subset of $\{1,2,...,k\}$ and $\bar{S}$ be its complement. Then*

$$h(\mathbf{z}_S|\mathbf{z}_{\bar{S}}) \leqslant h(\mathbf{z}_S^*|\mathbf{z}_{\bar{S}}^*) \tag{31}$$

where $(z_1^*,..,z_k^*) \sim N(0,\mathbf{R})$, i.e. Gaussian with the same mean and covariance.

Since $\mathbf{R}$ in this Lemma is arbitrary, the power and interference constraints can be accommodated. Applying this inequality to $h(\mathbf{y}_1|\mathbf{y}_2)$ and maximizing the upper bound over $\mathbf{R} \in S_{\mathbf{R}}$, one concludes that Gaussian input achieves the upper bound in (11), since, under such input, $\mathbf{y}_1,\mathbf{y}_2$ are also Gaussian.

C. Proof of Proposition 3

It was shown in [7] that $R_+(\mathbf{R},\mathbf{K})$ is concave in $\mathbf{R}$ for any fixed $\mathbf{K}$ and convex in $\mathbf{K}$ for any fixed $\mathbf{R}$. Since the feasible set $S_{\mathbf{R}}$ in (3) and $S_{\mathbf{K}}$ are convex (as an intersection of convex sets representing each constraint individually), von Neumann mini-max theorem applies (see e.g. [16]), from which (9) and hence (10) follow.

D. Proof of Proposition 4

First, we note that one cannot use the respective result from [7] directly since our feasible set $S_{\mathbf{R}}$ is different from that in [7] (in particular, it is *not* isotropic) so that the respective KKT conditions and other steps in [7] are to be modified accordingly. Nevertheless, we demonstrate here that this important property does hold for our cognitive radio scenario with an extra interference constraint.

After some manipulations, $h(\mathbf{y}_1 - \mathbf{Z}_{12}^*\mathbf{y}_2)$ can be expressed as

$$h(\mathbf{y}_1 - \mathbf{Z}_{12}^*\mathbf{y}_2) = \ln|\mathbf{I} + \mathbf{B}_1 + \mathbf{B}_2\mathbf{R}\mathbf{B}_2^+| + n_1\ln(2\pi e) \tag{32}$$

where

$$\begin{aligned} \mathbf{B}_1 &= \mathbf{Z}_{12}^*\mathbf{Z}_{12}^{*+} - \mathbf{Z}_{12}^*\mathbf{N}^{*+} - \mathbf{N}^*\mathbf{Z}_{12}^{*+}, \\ \mathbf{B}_2 &= \mathbf{H}_1 - \mathbf{Z}_{12}^*\mathbf{H}_2 \end{aligned} \tag{33}$$

and where the last term can be neglected. Since $h(\mathbf{y}_1 - \mathbf{Z}_{12}^*\mathbf{y}_2)$ is concave in $\mathbf{R}$, the feasible set $S_{\mathbf{R}}$ is convex and Slater's condition holds, the KKT conditions are both necessary and sufficient for optimality of the LHS of (17), which take the following form:

$$\mathbf{B}_2^+[\mathbf{I} + \mathbf{B}_1 + \mathbf{B}_2\mathbf{R}\mathbf{B}_2^+]^{-1}\mathbf{B}_2 + \mathbf{M}_1 - \lambda_1\mathbf{I} - \lambda_2\mathbf{H}_3^+\mathbf{H}_3 = 0 \tag{34}$$

$$\mathbf{M}_1\mathbf{R} = 0,\ \lambda_1(tr(\mathbf{R}) - P_T) = 0,\ \lambda_2(tr(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) - P_I) = 0, \tag{35}$$

$$\mathbf{M}_1 \geq \mathbf{0},\ \lambda_1 \geqslant 0,\ \lambda_2 \geqslant 0 \tag{36}$$

where $\mathbf{M}_1$ is a Lagrange multiplier responsible for the positive semi-definite constraints $\mathbf{R} \geq \mathbf{0}$, λ_1 and λ_2 are Lagrange multiplier responsible for the total power $tr(\mathbf{R}) \leqslant P_T$ and interference $tr(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) \leqslant P_I$ constraints.

Likewise, since $R_+(\mathbf{R},\mathbf{K}^*)$ is concave in $\mathbf{R}$, the KKT conditions are both necessary and sufficient for the optimality of the RHS of (17). After some manipulations (using matrix inversion Lemma etc.), they take the following form:

$$\mathbf{B}_1^+\mathbf{A}_{12}^{-1}\mathbf{B}_1 + \mathbf{M}_2 - \lambda_3\mathbf{I} - \lambda_4\mathbf{H}_3^+\mathbf{H}_3 = 0, \tag{37}$$

$$\mathbf{M}_2\mathbf{R} = 0,\ \lambda_3(tr(\mathbf{R}) - P_T) = 0,\ \lambda_4(tr(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) - P_I) = 0, \tag{38}$$

$$\mathbf{M}_2 \geq \mathbf{0},\ \lambda_3 \geqslant 0,\ \lambda_4 \geqslant 0 \tag{39}$$

where $\mathbf{M}_2,\ \lambda_3,\ \lambda_4$ are Lagrange multipliers and

$$\mathbf{A}_{12} = \mathbf{I} + \mathbf{H}_1\mathbf{R}\mathbf{H}_1^+ - \mathbf{Z}_{12}(\mathbf{N}^{*+} + \mathbf{H}_2\mathbf{R}\mathbf{H}_1^+), \tag{40}$$

$$\mathbf{Z}_{12} = (\mathbf{N}^* + \mathbf{H}_1\mathbf{R}\mathbf{H}_2^+)(\mathbf{I} + \mathbf{H}_2\mathbf{R}\mathbf{H}_2^+)^{-1}. \tag{41}$$

After some manipulations, $\mathbf{A}_{12}$ can be further expressed as

$$\mathbf{A}_{12} = \mathbf{I} + \mathbf{B}_1 + \mathbf{B}_2\mathbf{R}\mathbf{B}_2^+. \tag{42}$$

so that the condition in (37) takes the form:

$$\mathbf{B}_1^+(\mathbf{I} + \mathbf{B}_1 + \mathbf{B}_2\mathbf{R}\mathbf{B}_2^+)^{-1}\mathbf{B}_1 + \mathbf{M}_2 - \lambda_3\mathbf{I} - \lambda_4\mathbf{H}_3^+\mathbf{H}_3 = 0 \tag{43}$$

By comparing (34)-(36) to (38), (39) and (43), it is clear that any solution of the 1st set of KKT conditions also solves the 2nd one and hence optimal $\mathbf{R}$ are the same, as desired.

E. Proof of Proposition 5

Following 2nd inequality in (10) and the steps of the proof in [7, Lemma 3], one obtains

$$\mathbf{B}_2\mathbf{S}^*\mathbf{S}^{*+}(\mathbf{N}^{*+}\mathbf{H}_1 - \mathbf{H}_2)^+ = \mathbf{0} \tag{44}$$

Using Proposition 4,

$$\begin{aligned}\mathbf{R}^* &= \arg\max_{\mathbf{R}\in S_{\mathbf{R}}}\ R_+(\mathbf{R},\mathbf{K}^*)\\ &= \arg\max_{\mathbf{R}\in S_{\mathbf{R}}}\ h(\mathbf{y}_1 - \mathbf{Z}_{12}^*\mathbf{y}_2)\\ &= \arg\max_{\mathbf{R}\in S_{\mathbf{R}}}\ \ln|\mathbf{I} + \mathbf{H}_e\mathbf{R}\mathbf{H}_e^+| \end{aligned} \tag{45}$$

where $\mathbf{H}_e = (\mathbf{I} + \mathbf{B}_1)^{-1/2}\mathbf{B}_2$. Using 1st inequality in (10) and following judiciously the steps of the proof of [7, Lemma 4], one verifies that $\mathbf{H}_e\mathbf{S}^*$ is of full column-rank and hence so is $\mathbf{B}_2\mathbf{S}^*$, from which it follows that

$$\mathbf{S}^{*+}(\mathbf{N}^{*+}\mathbf{H}_1 - \mathbf{H}_2)^+ = \mathbf{0}. \tag{46}$$

and hence the desired result.

F. Proof of Proposition 6

We consider 1st the case of $\mathbf{H}_1 \neq \mathbf{Z}_{12}^*\mathbf{H}_2$ and establish $R_+(\mathbf{R}^*,\mathbf{K}^*) = R_-(\mathbf{R}^*)$. To this end, take Gaussian $\mathbf{x}$ and use the chain rule to obtain

$$R_+(\mathbf{R}^*,\mathbf{K}^*) = I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2) = R_-(\mathbf{R}^*) + I(\mathbf{x};\mathbf{y}_2|\mathbf{y}_1) \tag{47}$$

Using (30), one can express $I(\mathbf{x};\mathbf{y}_2|\mathbf{y}_1)$ as

$$I(\mathbf{x},\mathbf{y}_2|\mathbf{y}_1) = h(\mathbf{y}_2 - \mathbf{Z}_{21}\mathbf{y}_1) - \ln|\mathbf{I} - \mathbf{N}^{*+}\mathbf{N}^*| \tag{48}$$

where

$$\mathbf{Z}_{21} = (\mathbf{N}^{*+} + \mathbf{H}_2\mathbf{R}^*\mathbf{H}_1^+)(\mathbf{I} + \mathbf{H}_1\mathbf{R}^*\mathbf{H}_1^+)^{-1} \tag{49}$$

denotes the MMSE matrix of estimating $\mathbf{y}_2$ from $\mathbf{y}_1$ and $\mathbf{N}^+$ also represents the MMSE matrix of estimating $\boldsymbol{\xi}_2$ from $\boldsymbol{\xi}_1$.

At a saddle point $(\mathbf{R}^*,\mathbf{K}^*)$, one obtains:

$$h(\mathbf{y}_2 - \mathbf{Z}_{21}\mathbf{y}_1) = \ln|\mathbf{I} + \mathbf{H}_2\mathbf{R}^*\mathbf{H}_2^+ - \mathbf{Z}_{21}(\mathbf{N}^* + \mathbf{H}_1\mathbf{R}^*\mathbf{H}_2^+)|$$

$$= \ln|\mathbf{I} + \mathbf{H}_2\mathbf{R}^*\mathbf{H}_2^+ - \mathbf{N}^{*+}(\mathbf{I} + \mathbf{H}_1\mathbf{S}^*\mathbf{S}^{*+}\mathbf{H}_1^+)\mathbf{N}^*| \tag{50}$$

$$= \ln|\mathbf{I} + \mathbf{N}^{*+}\mathbf{H}_1\mathbf{S}^*\mathbf{S}^{*+}\mathbf{H}_1^+\mathbf{N}^* - \mathbf{N}^{*+}(\mathbf{I} + \mathbf{H}_1\mathbf{S}^*\mathbf{S}^{*+}\mathbf{H}_1^+)\mathbf{N}^*| \tag{51}$$

$$= \ln|\mathbf{I} - \mathbf{N}^{*+}\mathbf{N}^*|$$

where (50) and (51) are obtained via $\mathbf{N}^{*+}\mathbf{H}_1\mathbf{S}^* = \mathbf{H}_2\mathbf{S}^*$ from Proposition 5. Therefore, using (47) and (48),

$$I(\mathbf{x},\mathbf{y}_2|\mathbf{y}_1) = R_+(\mathbf{R}^*,\mathbf{K}^*) - R_-(\mathbf{R}^*) = 0. \tag{52}$$

Thus, $R_+(\mathbf{R}^*,\mathbf{K}^*) = R_-(\mathbf{R}^*)$ as desired.

Finally, we show that $R_+(\mathbf{R}^*,\mathbf{K}^*) = 0$ if $\mathbf{H}_1 = \mathbf{Z}_{12}^*\mathbf{H}_2$. To this end, note that

$$\mathbf{y}_1 - \mathbf{Z}_{12}^*\mathbf{y}_2 = \boldsymbol{\xi}_1 - \mathbf{Z}_{12}^*\boldsymbol{\xi}_2. \tag{53}$$

so that

$$R_+(\mathbf{R}^*,\mathbf{K}^*) = h(\boldsymbol{\xi}_1 - \mathbf{Z}_{12}^*\boldsymbol{\xi}_2) - h(\boldsymbol{\xi}_1 - \mathbf{N}^*\boldsymbol{\xi}_2). \tag{54}$$

Substituting $\mathbf{H}_1 = \mathbf{Z}_{12}^*\mathbf{H}_2$ into (16) and after some manipulations, one obtains

$$\mathbf{Z}_{12}^*(\mathbf{I} + \mathbf{H}_2\mathbf{R}^*\mathbf{H}_2^+) = (\mathbf{N}^* + \mathbf{Z}_{12}^*\mathbf{H}_2\mathbf{R}^*\mathbf{H}_2^+). \tag{55}$$

so that $\mathbf{Z}_{12}^* = \mathbf{N}^*$ and hence $R_+(\mathbf{R}^*,\mathbf{K}^*) = 0$, as desired.

REFERENCES

[1] S. Haykin et al (Eds.), Cognitive Radio, Part 1: Practical perspectives, and Part 2: Fundamental Issues, Proceedings of the IEEE, v. 97, n. 4 and 5, Apr. and May 2009.

[2] Q. Zhang et al (Eds.), Special Issue on Cooperative Communication and Signal Processing in Cognitive Radio Systems, IEEE JSTSP, v. 5, n.1, Feb. 2011.

[3] Special Issue on Cognitive Radio, IEEE Journal on Selected Areas in Communications (Cognitive Radio Series), v. 30, No. 10, Nov. 2012.

[4] A. G. Fragkiadakis et al, A Survey on Security Threats and Detection Techniques in Cognitive Radio Networks, IEEE Commun. Surv. & Tut., vol.15, no.1 , pp. 428-445, Feb. 2013.

[5] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[6] A. Khisti, G.W. Wornell, Secure Transmission With Multiple Antennas—Part I: The MISOME Wiretap Channel, IEEE Trans. Info. Theory, v. 56, No. 7, July 2010.

[7] A. Khisti, G.W. Wornell, Secure Transmission With Multiple Antennas—Part II: The MIMOME Wiretap Channel, IEEE Trans. Info. Theory, v. 56, No. 11, Nov. 2010.

[8] F. Oggier, B. Hassibi, The Secrecy Capacity of the MIMO Wiretap Channel, IEEE Trans. Info. Theory, v. 57, No. 8, Aug. 2011.

[9] S. Loyka, C.D. Charalambous, Rank-Deficient Solutions for Optimal Signaling over Wiretap MIMO Channels, IEEE Trans. Communications, v. 64, n. 6, pp. 2400 2411, June 2016.

[10] S. Loyka, C.D. Charalambous, Optimal Signaling for Secure Communications Over Gaussian MIMO Wiretap Channels, IEEE Transactions on Information Theory, v. 62, n. 12, pp. 7207 7215, Dec. 2016.

[11] L. Zhang et al, On the relationship between the multi-antenna secrecy communications and cognitive radio communications, IEEE Trans. Commun., vol.58, no.6, pp. 1877-1886, Jun. 2010.

[12] Y. Pei et al, Secure communication over MISO cognitive radio channels, IEEE Trans. Wireless Comm., vol.9, no.4, pp. 1494-1502, Apr. 2010.

[13] T. Liu and S. Shamai (Shitz), *"A note on the secrecy capacity of the multi-antenna wiretap channel,"* IEEE Trans. Inf. Theory., vol. 55, no. 6, pp. 25472553, Jun. 2009.

[14] S. Loyka, C. D. Charalambous, *"An Algorithm for Global Maximization of Secrecy Rates in Gaussian MIMO Wiretap Channels,"* IEEE Trans. Commun., vol. 63, no. 6, pp. 2288-2299, June. 2015.

[15] J. A. Thomas, *"Feedback can at most double Gaussian multiple access channel capacity,"* IEEE Trans. Inf. Theory., vol. 33, no. 5, pp. 711-716, Sep. 1987.

[16] S. Boyd and L. Vandenberghe, *"Convex Optimization,"* Cambridge University Press, 2004.