Cognitive Beamforming for Multiple Secondary Data Streams With Individual SNR Constraints

Sheng-Ming Cai and Yi Gong, Senior Member, IEEE

Abstract

In this paper, we consider cognitive beamforming for multiple secondary data streams subject to individual signal-to-noise ratio (SNR) requirements for each secondary data stream. In such a cognitive radio system, the secondary user is permitted to use the spectrum allocated to the primary user as long as the caused interference at the primary receiver is tolerable. With both secondary SNR constraint and primary interference power constraint, we aim to minimize the secondary transmit power consumption. By exploiting the individual SNR requirements, we formulate this cognitive beamforming problem as an optimization problem on the Stiefel manifold. Both zero forcing beamforming (ZFB) and nonzero forcing beamforming (NFB) are considered. For the ZFB case, we derive a closed form beamforming solution. For the NFB case, we prove that the strong duality holds for the nonconvex primal problem and thus the optimal solution can be easily obtained by solving the dual problem. Finally, numerical results are presented to illustrate the performance of the proposed cognitive beamforming solutions.

Index Terms

Cognitive radio, transmit beamforming, interference constraint, Stiefel manifold, MIMO.

I. INTRODUCTION

Cognitive radio (CR) has received considerable attention over the past few years because of its potential to ease the current overcrowded frequency spectrum. Based on the current spectrum allocation policy, most frequency bands are allocated exclusively to specified services. However,
such policy results in underutilization of precious spectrum resources [1]. In the meantime, the demand of extra spectrum is increasing with the rapid growth of wireless applications. As a result, it is worth considering the idea of allowing other users to use the spectrum while guaranteeing the priority of authorized users. In a CR network, the spectrum can be shared with unauthorized or secondary users (SUs), provided that they do not cause harmful interference to the authorized or primary users (PUs) [2], [3]. SUs may transmit when they detect a spectrum hole in either time or frequency domain [4], [5]. Such schemes usually work when the spectrum is severely underutilized, otherwise SUs might not have sufficient opportunities to gain channel access. Therefore, the secondary throughput would be significantly constrained and the secondary system would suffer from a long latency.

Multiple-input multiple-output (MIMO) technology provides extra spatial dimensions for transmissions. Multiple antennas can be used to reduce the interference at PU and satisfy the demand for high data rate at SU by carefully designing transmit and receive beamforming [6]. As a result, SUs may access the primary spectrum without causing harmful interference, even if PUs are also using the spectrum at the same time. By assuming full channel knowledge known at SU, the capacity of a CR network is given in closed form in [7], when no interference is allowed at PU. For the case of nonzero interference power constraint, the expressions of the secondary signal-to-noise ratio (SNR) and the interference power received at PU usually result in quadratically constrained quadratic programming problems and these problems may not be directly solved by convex tools, especially when there is a rank constraint. Semidefinite programming (SDP) relaxation can be used to convert such problem to a convex optimization problem by dropping the rank constraint and generate a local optimum [8]. It is shown in [9] that under certain conditions, a new solution can be generated from the one obtained by SDP relaxation without ruining the constraints or changing the objective function, and hence the solution is optimal. However, in general scenarios, the obtained local optimum may not be feasible for the original problem because usually its rank does not meet the requirement. As a result, approximation approaches such as the randomization procedure are used to generate a feasible solution [10], [11].

Under individual SNR constraints, the downlink transmission where each user has a single data stream is studied in [12]. In [13], the study is extended to multiple data streams. However, there exists interference between any two data streams even if they are for the same user. In [14],
the authors studied the transmit power minimization problem with individual SNR requirements and used joint decoding to remove the interference. An iterative algorithm is proposed therein to solve the problem, but it is not clear whether the iterative algorithm can converge to the global/local optimum. The cognitive transmission with multiple antennas equipped at secondary transmitter (ST) and secondary receiver (SR) is studied in [15], where the secondary throughput is maximized subject to the secondary power constraint and primary interference power constraint. It is shown therein that the secondary transmit beamforming problem can be converted to an optimization problem with unitary constraint, and then an algorithm is proposed to compute the beamforming matrix such that a local optimum can be obtained.

In this paper, we study the problem of secondary transmit beamforming with multiple secondary data streams subject to individual SNR constraints. We are interested in the beamforming design at ST so as to minimize its transmit power under both its own per data stream SNR constraint at SR and the interference-power constraint at the primary receiver (PR). We use zero forcing beamforming (ZFB) to deal with the case when no interference is allowed at PR. If a positive interference power constraint is allowed, we design nonzero forcing beamforming (NFB). We formulate the secondary power minimization problem as an optimization problem on the Stiefel manifold [16]. We show that SDP relaxation can achieve the global optimum when there is a single data stream but it may not be suitable for the multiple data streams scenario. For the multiple data streams, we derive a closed form solution for the ZFB case. As for the NFB case, we analyze the associated dual problem and provide the sufficient condition for strong duality. As a result, the global optimum can be obtained by solving the dual problem efficiently.

The rest of this paper is organized as follows. In Section II, we describe the system model and formulate the main optimization problem. The secondary beamforming feasibility test is also provided in this section. In Section III, we prove that with a single secondary data stream, SDP relaxation can lead to the global optimum. The case of multiple secondary data streams is considered in Section IV, where we firstly derive the closed-form solution for the ZFB problem. We then prove that for the NFB problem, the strong duality holds as well as the optimality of the solution of the dual problem. The numerical results are provided in Section V to illustrate the performance of the proposed secondary beamforming solutions. This paper is concluded in Section VI.

Notations: Scalar is denoted by a lower-case letter, while vector is denoted by bold-face lower-
case letter and matrix is denoted by bold-face upper-case letter. \( I_p \) denotes the \( p \times p \) identity matrix. For a matrix \( S \), \( \text{tr}(S) \), \( \text{rank}(S) \), \( S^H \), and \( S_{ij} \) denote its trace, rank, Hermitian matrix, and the entry at the \( i \)-th row and the \( j \)-the column, respectively. \( \text{diag}(s_1, s_2, \cdots, s_n) \) denotes a diagonal matrix with diagonal elements given by \( s_1, s_2, \cdots, s_n \). For a matrix \( S \), \( S \succeq 0 \) denotes that \( S \) is positive semidefinite. The complex Stiefel manifold \( St(n, p) \) is the set \( St(n, p) = \{ V \in \mathbb{C}^{n \times p} : V^H V = I_p \} \), where \( n \geq p \). For a Hermitian matrix \( S \in \mathbb{C}^{n \times n} \), the eigenvalue decomposition (EVD) is represented as \( S = U^H \Sigma U \), where \( U \) is a unitary matrix and \( \Sigma \) is a diagonal matrix. For a Hermitian matrix \( S \in \mathbb{C}^{n \times n} \) with \( \text{rank}(S) = r \leq n \), \( S^{-1} \) denotes the Moore-Penrose pseudoinverse of \( S \), i.e., \( SS^{-1}S = S \) and \( S^{-1}SS^{-1} = S^{-1} \). It is obtained from the following decomposition of \( S \): If \( S \) is written as \( S = U^H \Sigma U \), where \( U^H \in St(n, r) \), \( \Sigma \in \mathbb{R}^{r \times r} \) is a full rank diagonal matrix, then \( S^{-1} = U^H \Sigma^{-1} U \). Likewise, \( S^{1/2} \) and \( S^{-1/2} \) are found as \( S^{1/2} = U^H \Sigma^{1/2} U \) and \( S^{-1/2} = U^H \Sigma^{-1/2} U \), respectively.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

We consider a multi-antenna CR network in the presence of primary transmission, where there are a single pair of ST and SR, supporting multiple secondary data streams, and the primary users and the secondary users share the same bandwidth for transmission in an overlay approach. In particular, both the interference caused by secondary transmission experienced at PR and the interference caused by primary transmission experienced at SR are considered in this paper. We consider narrowband transmission for both primary and secondary users where multiple antennas are equipped at the primary transmitter (PT), PR, ST and SR. The numbers of antennas equipped at PT, PR, ST and SR are denoted as \( p, q, m \) and \( n \), respectively. Let \( G_x \in \mathbb{C}^{n \times p} \), \( H_x \in \mathbb{C}^{q \times m} \) and \( H \in \mathbb{C}^{n \times m} \) denote the channel matrices (all assumed to be full rank) of the PT \( \rightarrow \) SR link, ST \( \rightarrow \) PR link and ST \( \rightarrow \) SR link, respectively. It is assumed that \( H \in \mathbb{C}^{n \times m} \) and \( H_x \in \mathbb{C}^{q \times m} \) are known at ST, and \( G_x \in \mathbb{C}^{n \times p} \) is known at SR. Under this assumption, subject to its own SNR constraint at each data stream, ST is able to adjust its beamforming matrix based on the channel knowledge so as to optimally balance between minimizing its own transmit power and avoiding interferences at PR. In a fading environment, there are cases where it is difficult for ST and SR to obtain perfect knowledge of the instantaneous channels. In such cases, the results
obtained in this paper provide a performance upper-bound for the considered secondary transmit beamforming problem.

Letting $s_p \in \mathbb{C}^{p \times 1}$ denote the transmitted primary signal with zero mean and variance $P_p$, and $s_s \in \mathbb{C}^{m \times 1}$ denote the transmitted secondary signal, the received signal at SR can be written as

$$y = Hs_s + G_x s_p + z$$  \hspace{1cm} (1)

where $z \in \mathbb{C}^{n \times 1}$ represent the additive Gaussian noise with zero mean and unit variance at SR. The second term on the right-hand side of (1) represents the interference from the primary transmission. Therefore, the interference-plus-noise covariance matrix at SR is given by

$$W = E[(G_x s_p + z)(G_x s_p + z)^H] = P_p G_x G_x^H + I_n.$$ 

Letting $T \in \mathbb{C}^{m \times d}$ denote the secondary transmit beamforming matrix, the transmitted secondary signal can be represented as

$$s_s = Td_s$$

where $d_s \in \mathbb{C}^{d \times 1}$ denotes the secondary data, modeled as a random vector with $d \leq \min(m, n)$ denoting the number of secondary data streams and $E[d_s d_s^H] = I_d$.

It can be easily shown that the eigenvalues of $T^H H^H W^{-1} HT$ represent the SNR$^1$ values of secondary data streams at SR, after proper receive beamforming that maximizes the SNR of each data stream, as shown in Appendix A. Letting $M \triangleq H^H W^{-1} H$, we can use EVD to decompose $T^H M T$ as $T^H M T = U^H \Sigma U$, where the diagonal entries of $\Sigma$ now represent the SNR of each secondary data stream. We can always choose a unitary matrix and post-multiply it to $T$ to get a new $T$ such that $T^H M T$ is diagonal with the same eigenvalues, i.e.

$$T^H M T = \Sigma.$$ \hspace{1cm} (2)

In order to protect the primary communication, the interference power experienced at PR should not exceed a certain threshold. The peak interference power constraint can then be written as

$$\text{tr}(T^H H_x^H H_x T) \leq \xi$$

where the value of $\xi$ represents the maximum tolerable interference power at PR. As $\xi$ increases, ST has higher flexibility to design the transmit beamforming matrix. If $\xi$ is sufficiently large,

$^1$In this paper, it actually refers to the signal-to-interference-plus-noise ratio.
ST can communicate to SR as if PR is absent. For a certain $\xi$, it is possible that the underlying channel conditions fail to support the secondary QoS requirement with a certain $d$. In this case, ST may have to reduce $d$ or relax its QoS requirement to be able to transmit. Given the secondary QoS requirement and primary interference power constraint, we can test if there is a feasible secondary beamforming solution.

Clearly, the number of secondary data streams $d$ should not be greater than $\min(m, n)$. If the number of ST antennas, $m$, is strictly larger than the number of PR antennas, $q$, then there are $m - q$ available degrees of freedom or spatial dimensions for secondary transmission without causing any interference at PR, which can be realized by placing $T$ in the null space of $H_x$. On the other hand, if the PR can tolerate a nonzero interference (i.e., $\xi > 0$), the number of supported secondary data streams can be greater than $m - q$, depending on the value of $\xi$ and the underlying channel condition. Therefore, the secondary system can support at least $m - q$ secondary data streams. In this paper, we consider individual SNR requirements for all the secondary data streams. Let $\rho_i$ denote the $i$-th data stream’s SNR requirement, where $i = 1, \cdots, d$. Without loss of generality, we assume $\rho_1 \geq \cdots \geq \rho_d$.

**B. Problem Formulation**

Our objective is to minimize the secondary sum transmit power while satisfying both the secondary per data stream SNR constraint and the primary interference power constraint. Such problem is formulated as

$$(\text{P1}) : \quad \min_T \tr(T^H T)$$

subject to

$$\tr(T^H H_x^H H_x T) \leq \xi$$

$$\Sigma \succeq \diag(\rho_1, \rho_2, \cdots, \rho_d).$$

*Lemma 1*: Any $T$ satisfying (2) can be expressed as

$$T = M^{-1/2} V \Sigma^{1/2}$$

where $V \in \text{St}(m, d)$.

*Proof*: Please refer to Appendix B.

*Lemma 2*: The inequality constraint in (3c) can be replaced with its equality constraint, i.e., $\Sigma = \diag(\rho_1, \cdots, \rho_d)$.
Proof: With \((4)\), the objective function in \((3a)\) and the constraint in \((3b)\) can be respectively rewritten as

\[
\text{tr}(T^HT) = \text{tr}((\Sigma^{1/2})^HVH(M^{-1/2})^HM^{-1/2}V\Sigma^{1/2})
\]

\[= \text{tr}(\Sigma VH^HM^{-1}V) \tag{5}\]

and

\[
\text{tr}(T^HH_HH_xT) = \text{tr}(\Sigma VH^HM_xV) \leq \xi \tag{6}\]

where \(M_x \triangleq M^{-1/2}H_xH_xM^{-1/2}\).

Since it always holds that

\[
\text{tr}(\Sigma VH^HM^{-1}V) \geq \text{tr}(\text{diag}(\rho_1, \cdots, \rho_d)V^HM^{-1}V)
\]

\[
\text{tr}(\Sigma VH^HM_xV) \geq \text{tr}(\text{diag}(\rho_1, \cdots, \rho_d)V^HM_xV)
\]

we can set \(\Sigma = \text{diag}(\rho_1, \cdots, \rho_d)\) without affecting the optimal solution of (P1).

With (5), (6), and Lemma 2, the original problem (P1) can be reformulated as

\[
\text{(P2)}: \quad \min_{V \in St(m, d)} \text{tr}(\Sigma VH^HM^{-1}V) \tag{7a}
\]

\[
\text{s.t. } \text{tr}(\Sigma VH^HM_xV) \leq \xi. \tag{7b}
\]

C. Access Feasibility

Before solving (P2), we need to perform a feasibility test. Note that by letting \(\Sigma = \text{diag}(\rho_1, \cdots, \rho_d)\) and expressing \(T\) in the form of \((4)\), we have already met the secondary per data stream SNR requirement. However, for a given primary interference power constraint \((\xi)\) and underlying channel condition \((M_x)\), \((7b)\) is not always satisfied. If no \(V \in St(m, d)\) satisfies the interference constraint \((7b)\), we say that the secondary transmission is not feasible.

We define

\[
\xi_0 \triangleq \min_{V \in St(m, d)} \text{tr}(\Sigma VH^HM_xV). \tag{8}\]

If the actual interference power constraint \(\xi \geq \xi_0\), we can find a feasible \(V\) to generate the secondary transmit beamforming matrix to satisfy both the secondary SNR and primary interference power constraint, otherwise ST should keep silent.

In order to find \(\xi_0\), we provide the following lemma first.
Lemma 3: Given a diagonal matrix \( \Delta = \text{diag}(\delta_1, \cdots, \delta_u) \in \mathbb{C}^{u \times u} \) (\( \delta_1 \geq \cdots \geq \delta_u \)) and a Hermitian matrix \( \Omega \in \mathbb{C}^{v \times v} \) (\( v \geq u \)) with \( \omega_i, i = 1, \cdots, v, \) being its the eigenvalues (\( \omega_1 \geq \cdots \geq \omega_v \)) and \( \varphi_i \) being its eigenvector corresponding to \( \omega_i, \) for a matrix \( \Theta \in \text{St}(v, u) \), we have the following inequality:

\[
\text{tr}(\Delta \Omega^H \Theta) \geq \sum_{i=1}^{u} \delta_i \omega_{v-i+1}
\]

(9)

where the equality holds if \( \Theta \) is constructed as \( \Theta = [\varphi_v, \cdots, \varphi_{v-u+1}] \).

Proof: Please refer to Appendix C.

With Lemma 3, we can show that \( \xi_0 \) is given by

\[
\xi_0 = \sum_{i=1}^{d} \rho_i x_{m-i+1}
\]

(10)

where \( x_i \) denotes the \( i \)-th eigenvalue of \( M_x \) and \( x_1 \geq \cdots \geq x_m \). Since rank(\( M_x \)) = \( q \), we have

\[
\begin{cases}
  x_i > 0, & i \leq q \\
  x_i = 0, & i > q.
\end{cases}
\]

(11)

It is clear from (10) and (11) that when \( m-q \geq d \), \( \xi_0 = 0 \) and therefore secondary access (via ZFB or NFB) is always possible; when \( m-q < d \), \( \xi_0 \) will always be greater than zero and therefore only NFB is possible.

In Fig. 1, we illustrate the secondary access probability, i.e., the probability of \( \xi_0 \leq \xi \), with different number of secondary data streams, where identical SNR requirements are considered for multiple-stream cases. It is assumed that each entry of the involved channel matrices is an i.i.d. random variable, distributed as \( \mathcal{CN}(0, 1) \). With \( m = n = 5 \) and \( p = q = 2 \), Fig. 1 shows that the cases of \( d = 1, 2 \) and 3 lead to 100\% access probability since \( m-q \geq d \) in these cases. On the other hand, when \( d > 3 \), the access probability, increasing with a lower secondary SNR requirement, heavily depends on the number of secondary streams and the primary interference constraint \( \xi \). It is obvious that in this case more data streams or lower primary interference constraint results in a lower secondary access probability.

III. The Case of Single Data Stream

In this section, we show that SDP relaxation can be used to find the optimal secondary beamforming solution for the single data stream case.
By defining $X \triangleq TT^H$ and dropping the rank constraint ($\operatorname{rank}(X) = \operatorname{rank}(T) = d$), (P1) can be reformulated as a relaxed problem

\begin{align}
\text{(P3)} : & \quad \min_{\mathbf{X} \succeq 0} \operatorname{tr}(X) \\
& \quad \text{s.t. } \operatorname{tr}(H_x^H H_x X) \leq \xi \\
& \quad \det(XM - \rho_i I_n) = 0, \forall i = 1, \ldots, d.
\end{align}

The secondary per data SNR requirement is reflected in (12c), since $XM$ shares the same eigenvalues with $T^HMT$ and these eigenvalues denote the required secondary SNR values.

When there is a single data stream, constraint (12c) can be equivalently rewritten as $\operatorname{tr}(XM) = \rho_1$. It turns out that all the constraints together with the objective function are convex and thus we can apply any convex optimization algorithm to solve the problem. Let $X^*$ denote the solution for (P3). Since the problem is relaxed, $X^*$ leads to the optimal beamforming vector if $\operatorname{rank}(X^*) = 1$, or another rank-one Hermitian matrix can be generated from $X^*$ with all the optimization objective function and constraints unchanged [10]. Otherwise, there will be a
nonzero gap between the solutions of the relaxed problem (P3) and the original problem (P1).

**Theorem 1:** When \( d = 1 \), there exists a rank-one solution of the relaxed problem that is optimal.

**Proof:** The dual of the relaxed single data stream problem is given by

\[
\begin{align*}
\min_{\mu_1, \mu_2 \in \mathbb{R}} & \quad \mu_1 \xi + \mu_2 \rho_1 \\
\text{s.t.} & \quad (I + \mu_1 H_x^H H_x + \mu_2 M) \succeq 0 \\
& \quad \mu_1 \geq 0
\end{align*}
\]

where \( \mu_1 \) and \( \mu_2 \) are the Lagrange multipliers associated with (12b) and \( \text{tr}(XM) = \rho_1 \), respectively. Since the relaxed problem is convex, strong duality holds, i.e., the original and dual problems lead to the same solution. As a result, the following complementary conditions establish

\[
\begin{align*}
\text{tr}(X^*(I + \mu_1^* H_x^H H_x + \mu_2^* M)) &= 0 \\
\mu_1^*(\text{tr}(X^* H_x^H H_x) - \xi) &= 0 \\
\mu_2^*(\text{tr}(X^* M) - \rho_1) &= 0
\end{align*}
\]

where \( X^* \) and \( (\mu_1^*, \mu_2^*) \) are the primal and dual optimal solutions, respectively [17].

If \( \text{rank}(X^*) = 1 \), \( X^* \) can be decomposed as \( X^* = tt^H \), and hence \( t \in \mathbb{C}^{m \times 1} \) is the optimal transmit beamforming vector.

If \( \text{rank}(X^*) > 1 \), based on the rank-one decomposition theory [20], we can find a vector \( t \in \mathbb{C}^{m \times 1} \) such that the following equations

\[
\begin{align*}
\text{tr}(tt^H) &= \text{tr}(X^*) \\
\text{tr}(tt^H H_x^H H_x) &= \text{tr}(X^* H_x^H H_x) \\
\text{tr}(tt^H M) &= \text{tr}(X^* M)
\end{align*}
\]

hold simultaneously [9, Section V.B].

It can be directly concluded from (15b) and (15c) that matrix \( tt^H \) satisfies conditions (14b) and (14c). That is

\[
\begin{align*}
\mu_1^*(\text{tr}(tt^H H_x^H H_x) - \xi) &= 0 \\
\mu_2^*(\text{tr}(tt^H M) - \rho_1) &= 0.
\end{align*}
\]
On the other hand, since
\[
\text{tr}(tt^H (I + \mu_1^* H_x^H H_x + \mu_2^* M)) \\
= \text{tr}(tt^H) + \mu_1^* \text{tr}(tt^H H_x^H H_x) + \mu_2^* \text{tr}(tt^H M) \\
= \text{tr}(X^*) + \mu_1^* \text{tr}(X^* H_x^H H_x) + \mu_2^* \text{tr}(X^* M) \\
= \text{tr}(X^*(I + \mu_1^* H_x^H H_x + \mu_2^* M)) \\
= 0
\]
it follows that matrix $tt^H$ satisfies condition (14a). Therefore, $tt^H$ satisfies all the three complementary conditions in (14), and thus is the optimal rank-one solution of (P3). As a result, $t$ is the optimal secondary transmit beamforming vector resulting in zero duality gap.

For the multiple data streams case, since constraint (12c) is nonconvex, to the best of our knowledge, no applicable reformulation or relaxation on this constraint set can be found. As a result, the SDP relaxation might not be feasible for the considered multiple data streams case. In the next section, we reformulate the cognitive beamforming problem for multiple secondary data streams to a new problem on the Stiefel manifold and solve it effectively.

IV. The Case of Multiple Data Streams

A. Zero Forcing Beamforming

In the ZFB scenario, no interference is allowed at PR. According to (10) and (11), this scenario is possible only when $m - q \geq d$. Therefore, the primary interference constraint (7b) is rewritten as

\[
H_x T = 0.
\]  

(16)

The ZFB constraint (16) requires that the transmit beamforming matrix $T$ should be projected to the null space of $H_x$. By substituting (4) and the singular value decomposition (SVD) of $H_x = U_1 A_1 V_1$ into (16), where $U_1$ is a $q \times q$ unitary matrix, $A_1$ is a positive definite diagonal matrix, and $V_1^H \in St(m, q)$, we have

\[
H_x T = U_1 A_1 V_1 M^{-1/2} V \Sigma^{1/2} = 0.
\]  

(17)

It is clear to see that matrices $U_1$ and $A_1$ are full-rank square matrices. We remove them by left multiplying the corresponding inverse matrices and then drop the SNR requirement matrix $\Sigma$. 
It thus follows that
\[ V_1 M^{-1/2} V = 0. \] (18)

Now we denote the SVD of \( V_1 M^{-1/2} \) as \( V_1 M^{-1/2} = U_2 A_2 V_2 \), where \( U_2 \) is a \( q \times q \) unitary matrix, \( A_2 \) is a positive definite diagonal matrix, and \( V_2^H \in St(m, q) \). Similarly, (18) establishes when
\[ V_2 V = 0. \] (19)

The \( m \)-dimensional space that \( V \) lies in can be separated into two subspaces via the projected-channel SVD [7]: one is perpendicular to \( V_2 \) (by multiplying \( (I_m - V_2^H V_2) \)) and the other is parallel to \( V_2 \) (by multiplying \( V_2^H V_2 \)). Because \( \text{rank}(V_2^H V_2) = \text{rank}(V_2) = q \) and \( V_2 (I_m - V_2^H V_2) = 0 \), we have \( \text{rank}(I_m - V_2^H V_2) = m - q \). From (19), we can apply the subspace separation on \( V \) to get
\[
V = [(I_m - V_2^H V_2) + V_2^H V_2] V \\
= (I_m - V_2^H V_2) V. \] (20)

Therefore, we can rewrite the ZFB problem as

\begin{align*}
(P4) : \quad & \min_{V \in St(m, d)} \text{tr}(\Sigma V^H M^{-1} V) \\
& \text{s.t. } V = (I_m - V_2^H V_2) V. \tag{21a}
\end{align*}

The problem is now reduced to finding a Stiefel manifold matrix \( V \) with a signal subspace constraint.

By substituting constraint (21b) into the objective function in (21a), we have
\[
\text{tr}(\Sigma V^H M^{-1} V) \\
= \text{tr}(\Sigma V^H (I_m - V_2^H V_2)^H M^{-1} (I_m - V_2^H V_2) V) \\
= \text{tr}(\Sigma V^H R V) \tag{22}
\]

where \( R \triangleq (I_m - V_2^H V_2)^H M^{-1} (I_m - V_2^H V_2) \).

It can be easily shown that \( \text{rank}(R) = \text{rank}(I_m - V_2^H V_2) = m - q \), \( R \) can be decomposed as
\[
R = V_R A_R V_R^H \tag{23}
\]
where \( \mathbf{V}_R \in \text{St}(m, m - q) \), \( \Lambda_R \) is an \((m - q) \times (m - q)\) diagonal positive definite matrix with entries in non-decreasing order. Since \( \mathbf{V}_2 \mathbf{R} = \mathbf{0} \), we have \( \mathbf{V}_2 \mathbf{V}_R = \mathbf{0} \). With \( \mathbf{V}_R \), the optimal Stiefel manifold matrix \( \mathbf{V} \) is given in the following theorem.

**Theorem 2:** For the ZFB problem (P4), the solution is

\[
\mathbf{V}^* = \mathbf{V}_R \begin{bmatrix} I_d \\ 0 \end{bmatrix}_{(m-q) \times d}.
\]

1) **Proof of Theorem 2:** We first consider the case of \( d = m - q \). In this case, \( \mathbf{V} \) is an \( m \times (m - q) \) matrix on the Stiefel manifold. Since \( \mathbf{V} \) and \( \mathbf{V}_R \) are of the same size and perpendicular to \( \mathbf{V}_2 \), \( \mathbf{V} \) and \( \mathbf{V}_R \) are in the same signal subspace and there exists a unitary matrix \( \mathbf{Q} \) that satisfies

\[
\mathbf{V} = \mathbf{V}_R \mathbf{Q}.
\]  

(24)

By substituting (24) into (4), the transmit beamforming matrix is given by

\[
\mathbf{T} = \mathbf{M}^{-1/2} \mathbf{V}_R \mathbf{Q} \Sigma^{1/2}.
\]  

(25)

Therefore finding the optimal \( \mathbf{T} \) is equivalent to finding the optimal \( \mathbf{Q} \). By substituting (22) and (24) into (P4), the problem can be reformulated as

\[
\min_{\mathbf{Q} \in \text{St}(m-q,m-q)} \text{tr}(\mathbf{Q}^H \Lambda_R \mathbf{Q}).
\]  

(26)

By applying Lemma 3 and considering the structure of \( \Lambda_R \), the optimal \( \mathbf{Q} \) can be easily constructed as \( \mathbf{Q}^* = \mathbf{I}_{m-q} \).

If \( d < m - q \), we can treat this case as if there are still \( m - q \) secondary data streams, among which \( m - q - d \) data streams have zero-valued SNR requirements. In other words, the SNR constraint matrix is still an \((m - q) \times (m - q)\) diagonal matrix, but given by \( \text{diag}(\rho_1, \cdots, \rho_d, 0, \cdots, 0) \). The corresponding \( \mathbf{V} \) is still an \( m \times (m - q) \) matrix on the Stiefel manifold. As a result, the result for the case of \( d = m - q \) can be directly applied and the optimal transmit beamforming matrix can be shown as

\[
\mathbf{T}^* = \mathbf{M}^{-1/2} \mathbf{V}_R \mathbf{I}_{m-q} \text{diag}(\rho_1^{1/2}, \cdots, \rho_d^{1/2}, 0, \cdots, 0)
\]

\[
= \mathbf{M}^{-1/2} \mathbf{V}_R \begin{bmatrix} I_d \\ 0 \end{bmatrix}_{(m-q) \times d} \Sigma^{1/2}.
\]  

(27)

This completes the proof for Theorem 2.
Remarks: In (27), $\Sigma^{1/2}$ is used to allocate power to the secondary data streams, $V_R$ projects the data streams to suitable signal subspaces in order to avoid interference at PR, $Q^* = [I_d \ 0]^T$ is used to select the optimal dimension among the possible ones, and $M^{-1/2}$ is used to handle the interference together with noise at SR.

2) Other Choices of $Q^*$: When the number of distinct SNR requirements is less than $d$, there exist multiple data streams that have identical SNR requirement, and therefore there will be multiple possible forms of $Q^*$ as well as $V^*$. Suppose $\Sigma$ has $K$ distinct SNR values with

$$\Sigma = \text{diag}(\rho'_1, \ldots, \rho'_1, \rho'_2, \ldots, \rho'_2, \ldots, \rho'_K, \ldots, \rho'_K)$$

where $\sum_{k=1}^{K} n_k = d$ and $\rho'_1 > \rho'_2 > \cdots > \rho'_{n_K}$. It turns out that the optimal $Q$ is a block diagonal matrix in the form of

$$Q^* = \begin{bmatrix} \text{diag}(Q^*_1, \ldots, Q^*_K) & 0 \\ 0 & Q^*_K \end{bmatrix} \in St(m - q, d)$$

(28)

where $Q^*_k, k = 1, \ldots, K$, can be any $n_k \times n_k$ unitary matrix. For example, if $d = m - q = 4$ and $\Sigma = \text{diag}(a, a, b, b)$ ($a > b > 0$), the optimal $Q^*$ is given by

$$Q^* = \begin{bmatrix} Q^*_1 & 0 \\ 0 & Q^*_2 \end{bmatrix}$$

where $Q^*_1$ and $Q^*_2$ can be any $2 \times 2$ unitary matrices. Specifically, if all $\rho_i$'s are distinct, $Q^*$ must be a diagonal matrix with entries of 1 or $e^{j\theta}$; if all $\rho_i$'s are identical, $Q^*$ can be an arbitrary unitary matrix.

B. Nonzero Forcing Beamforming

In the following, we focus on the NFB problem (P2) with $\xi > 0$ and solve it by examining its Lagrangian dual. Although problem (P2) is nonconvex, we provide a sufficient condition for the strong duality to hold, and therefore general convex optimization algorithms can be applied to solve the problem.

1) Dual Problem: As compared to ZFB, NFB relaxes the interference constraint at the PR by allowing a nonzero $\xi$, therefore it is possible for the secondary system to use fewer secondary transmit antennas or transmit more secondary data streams. Moreover, it provides more opportunities for the secondary system to access the channel.
The Lagrangian of (P2) is defined as
\[
L(V, y) \triangleq \text{tr}(\Sigma V^{H}M^{-1}V) + y(\text{tr}(\Sigma V^{H}M_{x}V) - \xi)
\]
\[
= \text{tr}(\Sigma V^{H}(M^{-1} + yM_{x})V) - y\xi
\]
(29)
where \( y \geq 0 \) is the Lagrange multiplier. Define the dual objective \( g(\lambda) \) as an unconstrained minimization of the Lagrangian
\[
g(y) = \min_{V \in \text{St}(m,d)} L(V, y).
\]
(30)

Let \( \lambda_{i}, i = 1, \ldots, m \), denote the \( i \)-th eigenvalue of \( M^{-1} + yM_{x} \) with \( \lambda_{1} \leq \cdots \leq \lambda_{m} \). Given a \( y \), we can derive from Lemma 3 that \( L(V, y) \) is minimized when \( V \) is constructed as
\[
V_{y} = [v_{1}, \ldots, v_{d}]
\]
(31)
where \( v_{i}, i = 1, \ldots, d \), denotes the eigenvector of \( M^{-1} + yM_{x} \) corresponding to \( \lambda_{i} \). Therefore, matrix \( V_{y} \), as well as the corresponding sum transmit power \( \text{tr}(\Sigma V_{y}^{H}M^{-1}V_{y}) \) and interference power \( \text{tr}(\Sigma V_{y}^{H}M_{x}V_{y}) \), are functions of \( y \). The Lagrange dual problem associated with (P2) is given by
\[
(P5) : \quad \max g(y)
\]
\[
\quad \text{s.t. } y \geq 0.
\]
(32a)
(32b)

Note that the original problem (P2) is nonconvex because the domain of \( V \) is nonconvex. Therefore, we cannot directly tell if the strong duality holds or not. The following theorem provides a sufficient condition for strong duality between the primal problem (P2) and the dual problem (P5).

**Theorem 3:** The strong duality between (P2) and (P5) holds if the feasibility test is passed, i.e., \( \xi \geq \xi_{0} \).

**Remarks:** The feasible domain is nonempty if and only if \( \xi \geq \xi_{0} \). As a result, as long as the ST can access the channel, we can always solve the dual problem to obtain the optimal beamforming solution, which is of zero duality gap.
2) **Proof of Theorem 3**: To the best of our knowledge, although there are a number of constraint qualifications under which strong duality holds even if the primal problem is nonconvex [17], [21], [22], none of them can be directly applied to decide if the strong duality holds in our case.

Define $\epsilon(y) \triangleq y(\text{tr}(\Sigma V_y^H M_x V_y) - \xi)$. Note that for an arbitrary $y$, $\epsilon(y)$ represents the difference between $g(y)$ and the corresponding primal objective function value $\text{tr}(\Sigma V_y^H M^{-1} V_y)$. It is known that the dual function $g(y)$ yields a lower bound of (P2) for all $y \geq 0$. If there exists a $y$ such that $\epsilon(y) = 0$, then the result of the dual problem (P5) becomes $\text{tr}(\Sigma V_y^H M^{-1} V_y)$, which is exactly equal to the corresponding primal objective function value, and serves as the infimum of (P2). That is, $y$ and $V_y$ are the optimal solution of the dual problem (P5) and the primal problem (P2), respectively, and therefore zero duality gap holds. Otherwise, if no $y$ and the corresponding $V_y$ make $\epsilon(y) = 0$, we cannot find the optimal $V$ by solving the dual problem (P5). As a result, $g(y)$ becomes a loose lower bound of (P2) for all $y \geq 0$, which implies that the zero duality gap does not hold.

We will need the following lemma to prove Theorem 3.

**Lemma 4**: The transmit power $\text{tr}(\Sigma V_y^H M^{-1} V_y)$ is a monotonically increasing function of $y$ and the interference power $\text{tr}(\Sigma V_y^H M_x V_y)$ is a monotonically decreasing function of $y$. 

**Proof**: Please refer to Appendix D. 

Lemma 4 suggests that the minimum interference power is achieved at $y = \infty$. If the secondary transmission is feasible, i.e., $\xi \geq \xi_0$, it is easy to show that either of the following two conditions is satisfied:

\[
\begin{align*}
\text{tr}(\Sigma V_y^H M_x V_y) &= \xi, \exists y > 0 \\
\text{tr}(\Sigma V_y^H M_x V_y) &\leq \xi, y = 0.
\end{align*}
\]

Therefore, $\epsilon(y) = 0$ is always satisfied, i.e., the strong duality holds. That is to say, passing the feasibility test is the sufficient condition for zero duality gap in our problem. Furthermore, the $V_y$ corresponding to the minimum feasible $y$ of the dual problem is the optimal solution that minimizes the sum transmit power. General convex optimization algorithms can be applied to find the optimal $y$ and consequently the optimal transmit beamforming matrix.
Fig. 2. Optimal beamforming versus feasible beamforming. $d = 1$.

V. NUMERICAL RESULTS

Numerical results are provided in this section to illustrate the performance of the proposed beamforming solutions under various channel conditions and system requirements. We assume that each entry of channel matrices $G_x$, $H_x$, and $H$ is an i.i.d. random variable, distributed as $CN(0, 1)$. The primary transmit power $P_p$ is assumed to be 1, unless otherwise specified. Throughout this section, we set $p = q = 2$ and $m = n = 5$, and $d$ is selected to satisfy $m - q \geq d$ such that secondary access is always possible. Monte Carlo simulations with 2000 randomly generated channel-groups ($H, H_x, G_x$) are implemented, and the average minimum secondary transmit power is plotted versus the secondary per data stream SNR constraint or the primary interference power constraint in Figs. 2-8. In the case of multiple secondary data streams (i.e., $d > 1$), equal SNR requirements are considered in all the figures except Fig. 8.

Firstly, we consider the case of single secondary data stream in Figs. 2 and 3. Fig. 2 shows the minimum required average secondary transmit power based on the derived optimal beamforming matrix for ZFB. As a comparison, the required secondary transmit power averaged over all
feasible beamforming matrices is also shown in Fig. 2. Here, the feasible beamforming matrix refers to any beamforming matrix that satisfies the interference constraint in (3b) and the SNR requirement in (3c). From (5), it is clear that the required secondary transmit power linearly increases with the secondary SNR requirement, as shown in Fig. 2. In Fig. 2 we observe that the optimal beamforming matrix brings significant power saving over the feasible beamforming matrix and that this power saving increases with the primary transmit power.

In Fig. 3 we compare the SDP relaxation approach in Section III and the Stiefel manifold transformation approach in Section IV for both ZFB and NFB cases. Fig. 3 verifies that the SDP relaxation and the Stiefel manifold transformation approach achieve the same optimal performance, as expected.

Next, we consider the case of multiple secondary data streams. For simplicity, only the optimal beamforming solutions are presented in the rest of this section. The impact of interference power constraint $\xi$ and secondary per data stream SNR requirement $\rho$ on the minimum required secondary transmit power is shown in Figs. 4 and 5 where $d = 2$ secondary data streams are
considered. From Fig. 4 it is observed that for a fixed $\rho$, the required transmit power decreases significantly when $\xi$ is slightly increased over the zero value (the beamforming is changed from ZFB to NFB) while the required power decreases slowly when the positive $\xi$ further increases. This is not unexpected since ZFB significantly restricts the available beamforming dimensions as compared to NFB. In this case, there are only three dimensions available for ZFB and the other two dimensions are not permitted to use due to the nature of ZFB, while NFB can use all the five dimensions for beamforming. On the other hand, for both ZFB and NFB (i.e., zero and nonzero $\xi$ values), a lower secondary SNR requirement always brings a significant reduction in the required secondary transmit power. The lower bound of the secondary transmit power when $\xi \to \infty$ (it can be easily obtained by dropping constraint (3b) in (P1)) is also plotted in Fig. 4 where we observe that $\rho$ is more dominant than $\xi$ in determining the required secondary power. The similar finding can be observed in Fig. 5 as well as the relatively small difference among the nonzero $\xi$ values. In particular, we observe that at small $\rho$ values, there is almost no difference between the nonzero $\xi$ values in Fig. 5. This is because when the secondary SNR requirement
is small, the required secondary power is small and thus the interference caused to the primary receiver is also small and always below the underlying primary interference constraint $\xi$. When $\rho$ increases, the required secondary transmit power needs to increase and therefore the caused interference may no longer be lower than $\xi$. For a more strict interference constraint (i.e., a lower $\xi$), the available secondary beamforming directions are subject to more restrictions and therefore more secondary transmit power is needed to satisfy the primary interference constraint as well as the secondary per data stream SNR constraint at the same time.

The minimum required secondary power per data stream is investigated in Fig. 6 where the SNR requirement of each secondary data stream is assumed to be equal to each other, denoted as $\rho$. Together with Fig. 3 it is seen that when $d$ is increased from 1 to 3, more power needs to be allocated to each data stream on average. When $d$ is small, ST can choose channels with good conditions to transmit data streams and therefore the required power per data stream can be relatively small. As $d$ increases, while satisfying the primary interference power constraint as a higher priority, ST may have to use channels with poor conditions and hence more power
is required to transmit data streams with the same secondary SNR requirement per data stream. Based on Figs. 3 and 6, the total minimum required transmit power with different number of secondary data streams is shown in Fig. 7.

Fig. 8 illustrates the impact of distinct and identical secondary SNR requirements for both ZFB and NFB (zero and nonzero \( \xi \) values) cases with \( d = 2 \). Here, the SNR requirements have been chosen such that the achievable sum-rate is the same for both distinct and identical SNR requirements. Fig. 8 shows that the case of distinct SNR requirements costs less power than the identical SNR requirements case. The reason is given as follows. MIMO provides parallel channels for the transmission of multiple data streams but the underlying conditions of different channels may not be the same. From the secondary system point of view, with distinct SNR requirements, the secondary data stream with a higher SNR requirement can be allocated to a channel with better conditions to save the transmit power. This is why the case of distinct SNR requirements outperforms the case of identical SNR requirements in terms of lower secondary transmit power.
Fig. 7. Impact of the number of secondary data streams.

VI. CONCLUSION

Cognitive beamforming for multiple secondary data streams is considered in this paper. The optimal beamforming strategy is designed to minimize the secondary transmit power consumption, subject to the individual secondary per data stream SNR constraints and primary interference power constraint. By exploiting the Stiefel manifold, we derived the closed form solution for zero forcing beamforming; while for nonzero forcing beamforming, we prove that the strong duality holds for the nonconvex primal problem and thus the optimal solution can be easily obtained by solving the dual problem. As for the future work, we are interested in exploring the impact of imperfect channel state information at secondary users and designing robust cognitive beamforming solutions.
Letting $T = [t_1, \cdots, t_d]$ and denoting $r_i$ ($i = 1, \ldots, d$) as the receive beamforming vector for the $i$th data stream, the SNR for the $i$th data stream is given by

$$
\hat{\rho}_i = \frac{r_i H t_i H^H r_i^H}{r_i (\sum_{k \neq i} H t_k H^H + W) r_i^H} = \frac{r_i M_1 r_i^H}{r_i M_2 r_i^H}.
$$

Eqn. (33) is a generalized Rayleigh quotient and the $r_i$ that maximizes $\hat{\rho}_i$ is the dominant eigenvector of $(M_2^{-1/2})^H M_1 M_2^{-1/2}$. It thus follows that the maximum $\hat{\rho}_i$ is equal to the $i$-th eigenvalue of $T^H H^H W^{-1} H T$.

APPENDIX B

For any $T$ that satisfies (2), we define $P \triangleq M^{1/2} T \in \mathbb{C}^{m \times d}$ and let $p_i$ denote the $i$th column of $P$. Since $\Sigma$ is a diagonal matrix, it follows from (2) that all $p_i$’s are orthogonal to each other and the norm of $p_i$ is equal to $\Sigma_{ii}^{1/2}$. We normalize each $p_i$ to form a matrix $V = [p_1/\|p_1\|_2, \ldots, p_d/\|p_d\|_2] = P \Sigma^{-1/2} \in St(m, d)$. This directly yields $T = M^{-1/2} P = M^{-1/2} V \Sigma^{1/2}$. 

Fig. 8. Distinct versus identical secondary SNR requirements.
APPENDIX C

It is shown in [19] that for any \( w \times w \) Hermitian matrices \( A \) and \( B \),
\[
\sum_{i=1}^{w} \lambda_i(A)\lambda_{w-i+1}(B) \leq \text{tr}(AB)
\]  
(34)

where \( \lambda_i(X) \) denotes the \( i \)-th eigenvalue of matrix \( X \in \mathbb{C}^{w \times w} \) and \( \lambda_1(X) \geq \cdots \geq \lambda_w(X) \).

Let \( \Delta' \triangleq \text{diag}(\Delta, 0) \in \mathbb{C}^{v \times v} \) and \( \Theta' \triangleq [\Theta \Theta_{\perp}] \in \mathbb{C}^{v \times v} \) where \( \text{rank}(\Theta_{\perp}) = v - u \) and \( \Theta_{\perp}^H\Theta = 0 \). Note that \( \lambda_{u+1}(\Delta') = \cdots = \lambda_v(\Delta') = 0 \) and the eigenvalues of \( \Theta'^H\Omega\Theta' \) are equal to those of \( \Omega \). By using (34), we get
\[
\sum_{i=1}^{u} \delta_{i}\omega_{v-i+1} = \sum_{i=1}^{v} \lambda_i(\Delta')\lambda_{v-i+1}(\Omega)
\leq \text{tr}(\Delta'\Theta'^H\Omega\Theta')
\leq \text{tr}(\Delta\Theta'^H\Omega\Theta).
\]
The equality holds when \( \Theta'^H\Omega\Theta' = \text{diag}(\omega_v, \cdots, \omega_1) \), i.e., \( \Theta' = [\varphi_v, \cdots, \varphi_1] \). It thus follows that \( \Theta = [\varphi_v, \cdots, \varphi_{v-u+1}] \). This completes the proof.

APPENDIX D

Let \( \Sigma' = \text{diag}(\Sigma, 0) \in \mathbb{R}^{m \times m} \), and \( V' = [V_y V_{y\perp}] \) where \( V_{y\perp} \in \mathbb{C}^{m \times (m-d)} \) and its \( i \)-th \((i = 1, \ldots, m-d)\) column is the eigenvector of \( M^{-1} + yM_x \) corresponding to \( \lambda_{d+i} \). The dual problem (P5) can be equivalently rewritten as
\[
\max_{y \geq 0} \min_{V' \in St(m,m)} \text{tr}(\Sigma'V'^H(M^{-1} + yM_x)V' - y\xi).
\]  
(35)

Let \( \pi : \mathbb{C}^{m \times m} \rightarrow St(m,m) \) be a projection operator mapping an \( m \times m \) complex matrix to the closest point on the Stiefel manifold [16]. Suppose that \( Z \) is the derivative of the \( V' \) with respect to \( y \). It means that given a sufficiently small increment \( \Delta y \), the \( V' \) related to \( M^{-1} + (y + \Delta y)M_x \), denoted by \( V'' \), can be approximated as
\[
V'' \approx \pi(V' + \Delta yZ) = V' + \Delta yZ + O(\Delta y^2)
\]
where \( Z \) can be written as \( Z = V'C \) and \( C \in \mathbb{C}^{m \times m} \) is a skew-Hermitian matrix \( (C^H + C = 0) \) to be determined [16]. Let us define \( \Lambda \triangleq \text{diag}(\lambda_1, \cdots, \lambda_m) \) and \( F \triangleq V'^H M_x V' \). Since the columns of \( V'' \) are the eigenvectors of \( M^{-1} + (y + \Delta y)M_x \), the expression \( V''^H (M^{-1} + (y + \Delta y)M_x)V'' \) is a
diagonal matrix consisting of the eigenvalues of $M^{-1} + (y + \Delta y) M_x$. $V''H (M^{-1} + (y + \Delta y) M_x) V''$ can be decomposed as

$$V''H (M^{-1} + (y + \Delta y) M_x) V'' 
\approx \pi (V' + \Delta y Z)^H (M^{-1} + (y + \Delta y) M_x) \pi (V' + \Delta y Z)$$

$$= V'H (M^{-1} + y M_x) V' + \Delta y (Z'H (M^{-1} + y M_x) V'$$

$$+ V'H (M^{-1} + y M_x) Z + V''H M_x V') + O(\Delta y^2)$$

$$= \Lambda + \Delta y (-C \Lambda + \Lambda C + F) + O(\Delta y^2) \quad (36)$$

which indicates that matrix $(-C \Lambda + \Lambda C + F)$ has to be diagonal. By computing the expressions of the entries of $-C \Lambda + \Lambda C + F$ and setting the off-diagonal entries to be zero, it turns out that

$$(\lambda_i - \lambda_j) C_{ij} = -F_{ij}, \forall i \neq j. \quad (37)$$

Firstly, we consider the impact of positive $\Delta y$ on the interference power. Let $C^{(d)}$ and $C^{(od)}$ denote the matrices that only contain the diagonal and off-diagonal entries of $C$, respectively. With (37), the interference power difference at a sufficiently small increment $\Delta y > 0$ can be shown to be

$$\text{tr}(\Sigma' \pi (V' + \Delta y Z)^H M_x \pi (V''H + \Delta y Z)) - \text{tr}(\Sigma' V''H M_x V')$$

$$\approx 2\Delta y \text{Re}\{\text{tr}(\Sigma' V'H M_x Z)\}$$

$$= 2\Delta y \text{Re}\{\text{tr}(\Sigma' F(C^{(d)} + C^{(od)})\})$$

$$\stackrel{(a)}{=} 2\Delta y \text{Re}\{\text{tr}(\Sigma' FC^{(od)})\}$$

$$= 2\Delta y \sum_{i=1}^{d} \rho_i \sum_{j \neq i} F_{ij} C^{(od)}_{ij}$$

$$\stackrel{(b)}{=} 2\Delta y \sum_{i=1}^{d} \rho_i \sum_{j \neq i} (\lambda_i - \lambda_j) |C^{(od)}_{ij}|^2$$

$$= 2\Delta y \left( \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} (\rho_i - \rho_j) (\lambda_i - \lambda_j) |C^{(od)}_{ij}|^2 + \sum_{i=1}^{d} \sum_{j=d+1}^{m} \rho_i (\lambda_i - \lambda_j) |C^{(od)}_{ij}|^2 \right)$$

$$\leq 0. \quad (38)$$

In (38), (a) holds because $\text{Re}\{\text{tr}(\Sigma' V'H M_x V'C^{(d)})\} = \text{Re}\{\text{tr}((C^{(d)} \Sigma')(V''H M_x V'))\} = 0$, where $C^{(d)} \Sigma'$ is skew-Hermitian and $V''H M_x V'$ is Hermitian; (b) is obtained by substituting (37); (c)
is because $\rho_i \geq \rho_j$ and $\lambda_i \leq \lambda_j$ for $i < j$, and the equality holds if and only if $\rho_i = \rho_j$, $\lambda_i = \lambda_j$, $\forall i \neq j$. Considering the random nature of wireless channels, it is nearly impossible for the equality under (c) to hold. Therefore, we conclude that the interference power in the form of $\text{tr}(\Sigma'V'H\Sigma_{x}V')$ is a monotonically decreasing function of $y$.

Next, we consider how the value of $y$ affects the required secondary transmit power. Eqn. (29) can be rewritten as

$$L(V', y) = (1 + y)(\text{tr}(\Sigma'V'H(\alpha\Sigma^{-1} + (1 - \alpha)\Sigma_{x})V')) - y\xi$$

(39)

where $\alpha = 1/(1 + y) \in [0, \infty)$. Note that the selected $V'$ in the dual problem is determined by the weighted sum of $\Sigma^{-1}$ and $\Sigma_{x}$. A larger $y$ increases the weight of $\Sigma_{x}$ and at the same time decreases the weight of $\Sigma^{-1}$. Let $\Delta\alpha$ be the change of $\alpha$ as $y$ becomes $y + \Delta y$ and $E \triangleq V'H\Sigma^{-1}V'$. Similar to (36), it can be derived that

$$V''H((\alpha + \Delta\alpha)\Sigma^{-1} + (1 - \alpha - \Delta\alpha)\Sigma_{x})V''$$

$$= \alpha\Lambda - \frac{\Delta\alpha}{\alpha}[( - \Lambda C + \Lambda C) - \alpha(E - F)] + O(\Delta\alpha^2)$$

$$= \alpha\Lambda - \Delta\alpha[\alpha(-\Lambda C + \Lambda C + F) + (1 - \alpha)(-\Lambda C + \Lambda C - \alpha E/(1 - \alpha))] + O(\Delta\alpha^2)$$

(40)

where matrix $(-\Lambda C + \Lambda C - \alpha E/(1 - \alpha))$ should be diagonal. It thus follows that

$$(\lambda_i - \lambda_j)C_{ij} = \alpha E_{ij}/(1 - \alpha) = E_{ij}/y.$$ 

(41)

With (41), the difference of the required transmit power at $\Delta y > 0$ can be shown to be

$$\text{tr}(\Sigma'\pi(V' + tZ)H\Sigma^{-1}\pi(V'H + tZ)) - \text{tr}(\Sigma'\Sigma'H\Sigma^{-1}V')$$

$$\approx -\frac{2\Delta y}{y} \left( \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} (\rho_i - \rho_j) (\lambda_i - \lambda_j)|C_{ij}^{(od)}|^2 + \sum_{i=1}^{d} \sum_{j=d+1}^{m} \rho_i (\lambda_i - \lambda_j)|C_{ij}^{(od)}|^2 \right)$$

$$\geq 0.$$ 

(42)

With a similar argument used for (38), we conclude that the required secondary transmit power in terms of $\text{tr}(\Sigma'\Sigma'H\Sigma^{-1}V')$ is a monotonically increasing function of $y$. The monotonicity of functions $\text{tr}(\Sigma'\Sigma'H\Sigma_{x}V')$ and $\text{tr}(\Sigma'\Sigma'H\Sigma^{-1}V')$ will not change if $\Sigma'$ and $V'$ are replaced by $\Sigma$ and $V_{y}$, respectively. This completes the proof.
REFERENCES

[1] FCC, “Spectrum policy task force report,” *ET Docket*, no. 02-155, Nov. 2002.
[2] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.
[3] W. D. Lu, Y. Gong, *et al.*, “Cooperative OFDM relaying for opportunistic spectrum sharing: protocol design and resource allocation,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2126-2135, June 2012.
[4] J. M. Peha, “Sharing spectrum through spectrum policy reform and cognitive radio,” *Proc. IEEE*, vol. 97, no. 4, pp. 708-719, 2009.
[5] R. Tandra, S. M. Mishra, and A. Sahai, “What is a spectrum hole and what does it take to recognize one?,” *Proc. IEEE*, vol. 97, no. 5, pp. 824-848, 2009.
[6] M. H. Islam, Y.-C. Liang, and A.T. Hoang, “Joint power control and beamforming for cognitive radio networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2415-2419, July 2008.
[7] R. Zhang and Y.-C. Liang, “Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks,” *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88-102, Feb. 2008.
[8] S. Burer and R. D. Monteiro, “A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization,” *Mathematical Programming*, vol. 95, no. 2, pp. 329-357, Feb. 2003.
[9] Y. J. Zhang and A. M.-C. So, “Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 1683-1693, Feb. 2011.
[10] Z. Q. Luo and W. Yu, “An introduction to convex optimization for communications and signal processing,” *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1426-1438, Aug. 2006.
[11] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, “Transmit beamforming for physical-layer multicasting,” *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239-2251, June 2006.
[12] M. Schubert and H. Boche, “Solution of the multiuser downlink beamforming problem with individual SINR constraints,” *IEEE Trans. Vehicular Technology*, vol. 53, no. 1, pp. 18-28, Jan. 2004.
[13] A. M. Khachan, A. J. Tenenbaum, and R. S. Adve, “Linear processing for the downlink in multiuser MIMO systems with multiple data streams,” in *Proc. IEEE Int. Conf. Commun.*, June 2006, vol. 9, pp. 4113-4118.
[14] Y.-H. Yang, S.-C. Lin, and H.-J. Su, “Multiuser MIMO downlink beamforming design based on group maximum SINR filtering,” *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1746-1758, Apr. 2011.
[15] H. Yu and V. K. N. Lau, “Rank-constrained Schur-convex optimization with multiple trace/log-det constraints,” *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 304-314, Jan. 2011.
[16] J. H. Manton, “Optimization algorithms exploiting unitary constraints,” *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635-650, Mar. 2002.
[17] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge, U.K.: Cambridge Univ. Press, 2003.
[18] E. Polak, *Optimization Algorithms and Consistent Approximations*, New York: Springer-Verlag, 1997.
[19] J. B. Lasserre, “A trace inequality for matrix product,” *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1500-1501, Aug. 1995.
[20] Y. Huang, A. De Maio, and S. Zhang, “Semidefinite programming, matrix decomposition, and radar code design,” in *Convex Optimization in Signal Processing and Communications*, D. P. Palomar and Y. C. Eldar, Eds., pp. 192-228. Cambridge Univ. Press, New York, 2010.
[21] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, 2003.
[22] A. E. Ozdaglar and D. P. Bertsekas, “The relation between pseudonormality and quasiregularity in constrained optimization,” [Online]. Available: [http://web.mit.edu/dimitrib/www/Quasiregularity.pdf](http://web.mit.edu/dimitrib/www/Quasiregularity.pdf)