IET Communications

Research Article

Resource optimisation using bandwidth-power product for multiple-input multiple-output orthogonal frequency-division multiplexing access system in cognitive radio networks

ISSN 1751-8628
Received on 18th November 2014
Revised on 28th April 2015
Accepted on 31st May 2015
doi: 10.1049/iet-com.2014.1108
www.ietdl.org

Bo Chen ✉, *Minjian Zhao, Lei Zhang, Ming Lei*

Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, People's Republic of China
✉ *E-mail: 3070902061chenbo@zju.edu.cn*

Abstract: This study investigates resource allocation problems for a point-to-point multi-carrier multiple-input multiple-output cognitive radio network. Different from conventional resource optimisation problems, a joint power and bandwidth resource optimisation framework using the novel optimisation metric, namely bandwidth-power product, is developed. Besides, rate requirement of the secondary system is satisfied, and the interferences introduced to the primary users (PUs) are below threshold of tolerance. The optimal source precoding matrix is designed and two methods, namely the project-channel singular value decomposition (SVD) and direct-channel SVD methods, are applied to satisfy interference power constraints for PUs. Then the unified power and channel allocation problem is derived and found to be a mixed-integer programming problem. Hence, a sub-optimal and tractable algorithm with low complexity is proposed. The innovative idea is to determine the channel resource budget by selecting the best channels, where the criterion for evaluating the quality of channel is detailed discussed. Then the power optimisation subproblem and channel allocation subproblem can be performed independently using the Lagrange-duality theory and Gauss–Newton method, respectively. The simulation results show significant improvement in spectral efficiency by using this framework compared to classical power optimisation framework using the waterfilling scheme.

1 Introduction

Radio spectrum is one of the most scarce and valuable resources for wireless communication, and the greatly increasing demand for wireless applications results in a usage crisis of current spectrum. In contrast to scarcity, the majority of licensed spectrum is underutilised at given location and time [1]. Cognitive radio (CR) is a promising technique to improve spectrum efficiency by allowing unlicensed CR nodes to share spectrum with licensed primary radio (PR) nodes under two kinds of spectrum sharing mechanisms, i.e. underlay spectrum sharing and overlay spectrum sharing. As to underlay spectrum sharing, the CR can transmit simultaneously with the PR over the same spectrum under the constraint that interferences introduced to the PR must be limited by predefined interference temperature. Whereas in overlay spectrum sharing, the CR can only make use of the idle spectrum that has not been utilised by the PR [2].

As is well known, orthogonal frequency-division multiplexing access (OFDMA) is proposed and regarded as a promising air interface technology for CR networks (CRNs) [3]. Resource allocation for OFDMA wireless systems has attracted great attention. The resource optimisation problem was divided into two categories in [4]: rate-adaptive (RA) and margin-adaptive (MA). For conventional OFDMA wireless systems, power optimisation, and channel allocation were extensively investigated in [5]. Taking the unique nature of CR into consideration, several works were done to evaluate their performance in CRNs by adding the primary user (PU) interference limits as an additional constraint [6–9].

Majority of prior studies on the resource optimisation problem for CRNs focus on frequency and/or time domain, assuming that both the primary and secondary transceivers own single antenna. Multiple-input multiple-output (MIMO) has attracted great attention during the past decade. MIMO, taking advantage of multiple antennas equipped by transceivers, extends resource allocation to space domain. The superiority of MIMO was fully validated in [10–12]. Opportunistic spectrum sharing in CRNs exploiting multi-antennas was fundamentally investigated in [13]. Furthermore, in [14], the author designed the optimal MIMO transceiver via majorisation theory. As research goes deep, by properly employing the MIMO structure, the secondary transmission, which is treated as interference by the PR should lie in the subspace orthogonal to the space spanned by the primary signal. Therefore, interference alignment (IA) has become a hot topic [15, 16]. However, IA fails to work in some cases due to the limited system resource [17], then interference cancellation (IC) is applied. By introducing cooperation at the transmitter and receiver, IA and cancellation (IAC) was presented in MIMO CRNs [18, 19]. Note that all these works tried to minimise the transmit power while the channel resource was totally ignored.

In addition to the total power and the system throughout metrics used in the resource optimisation problem, the space–bandwidth product (SBP) has been proven to be a valid metric in CRNs. SBP can be used to grasp the efficiency of spectrum utilisation, which is driven by the fact that wireless communication consumes space [20]. In [21], optimum spectrum sharing was proposed with the SBP metric in multi-hop CRNs. However, the spectral footprint was controlled only through the spectral bandwidth while the transmit power spectral density was fixed. A joint bandwidth and power optimisation problem in CRNs was introduced and fundamentally solved in [22]. Bandwidth–power product (BPP) has been put forward as a typical implementation of SBP and demonstrated to improve performance effectively in CRNs. However, to the best of our knowledge, few research works have been done to regulate optimum spectrum sharing using the BPP metric for MIMO-OFDMA CRNs.

Different from the aforementioned works using conventional performance metrics, e.g. using total power or system throughout as the objective function, we propose the BPP metric for resource optimisation in MIMO-OFDMA CRNs. We extend the work in [22] to the MIMO area while the BPP rather than the transmit power is optimised compared to [15–19]. The contributions of this paper are summarised as follows:

- Apply the novel optimisation metric (BPP) to the joint power and channel resource optimisation framework and formulate the corresponding optimisation problem mathematically.
- Design the optimal structure of source precoding matrix and propose the project-channel singular value decomposition (PC-SVD) and direct-channel singular value decomposition (DC-SVD) methods to deal with the interference power constraints at the PUs. Besides, the advantages and disadvantages of each method are detailed analysed.
- Derive the simplified power and channel allocation problem and propose a sub-optimal and tractable algorithm with low complexity. Besides, a simple criterion for evaluating the quality of the channel is introduced.
- Provide the persuasive simulation results, whose meaning is twofold. Firstly, the proposed sub-optimal algorithm achieves the acceptable performance with significant reduction of computation. Secondly, compared to the conventional resource optimisation in a MIMO CR area that employs the water-filling scheme, the efficiency of spectrum utilisation is well improved by our scheme.

The rest of this paper is organised as follows. The system model is briefly described in Section 2. In Section 3, we formulate the resource allocation problem. Then, detail analysis is made to fundamentally solve the corresponding resource optimisation problem in Section 4. In Section 5, we provide simulation results and demonstrate the performance of the proposed solution. Finally, we conclude this paper in Section 6.

2 System model

In this paper, the CRN uses the OFDMA transmission scheme. It comprises of a single pair of secondary transmitter–receiver and K primary receivers as shown in Fig. 1. The underlay spectrum sharing mechanism is applied, namely the secondary transmitter can share the same spectrum simultaneously with all the primary receivers under the constraints that interferences introduced to the primary receivers are below threshold of tolerance. Besides, we assume multiple antennas for both the secondary transceiver and all the primary receivers. Here, we denote N as the total number of channels available for allocation and B as the bandwidth of each channel. $M_{\mathrm{t,s}}$, $M_{\mathrm{r,s}}$, and $M_{k,\mathrm{p}}$ stand for the number of antennas at the secondary transmitter, secondary receiver and the kth primary receiver, respectively.

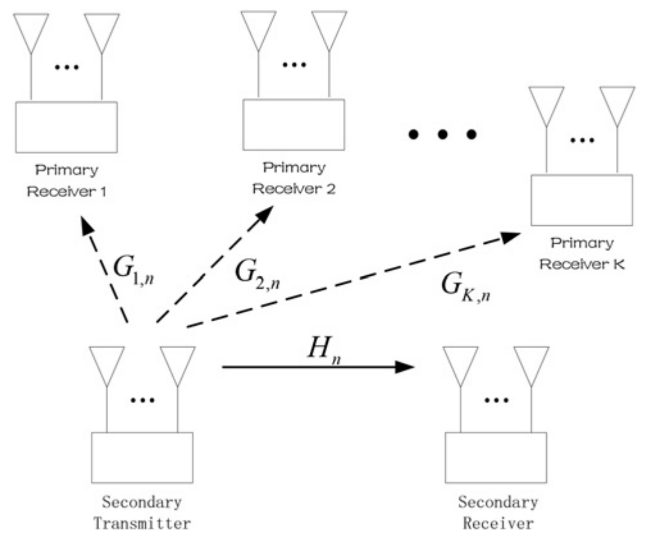

Fig. 1 *CR network: secondary user shares same transmit spectrum with* K *primary users*

We assume that the secondary transmitter precisely knows channel information from the secondary transmitter to the secondary and primary receivers. Based on accurate channel knowledge, the transmitter can optimally balance between lowering interferences at primary receivers and maximising its own benefit by adapting its transmit resources, such as power and spatial spectrum. In practice, the secondary receiver estimates channels from the secondary transmitter to it and then feeds channel-state information (CSI) back to the transmitter. While primary receivers have no duty to help the transmitter sense channels between them. By exploiting the channel reciprocity between the secondary transmitter and primary receivers, the secondary transmitter can directly obtain the CSI by periodically sensing transmitted signals from the primary receivers provided that the primary transmission employs time-division-duplex. In a deep fading condition, it may be difficult for the secondary transmitter to perfectly sense instantaneous channels. For such cases, the results shown in this paper provide upper-bounds of performance.

In [20], the author showed us the fact that wireless communication needs to consume space. Therefore, the novel bandwidth-power product metric, which is clearly defined as the multiplication of occupied bandwidth and sending power (a typical measurement for coverage space), can be used to embody the efficiency of spectrum utilisation. Classical resource allocation algorithms using the waterfilling scheme attempt to utilise as much bandwidth as possible so that transmit power can be minimised. This greedy allocation of channels leads to a big waste of spectral opportunities in CRNs using the overlay spectrum sharing mechanism. On the contrary, the spectral footprint optimisation method, aiming to minimise spectral footprint (the BPP metric), reduces the number of channels occupied by increasing sending power. This motivates the deeper research below.

3 Problem formulation

In this section, we evolve mathematical formulation of the resource optimisation framework for a MIMO-OFDMA CRN, which aims to minimise the BPP metric. The $M_{\mathrm{t,s}} \times 1$ signal vector transmitted on the nth channel by the secondary transmitter is given by

$$\boldsymbol{s}_n = \boldsymbol{U}_n \boldsymbol{b}_n, \tag{1}$$

where $\boldsymbol{b}_n$ is the $M_{b,n} \times 1$ source symbol vector, and $\boldsymbol{U}_n$ is the $M_{\mathrm{t,s}} \times M_{b,n}$ source precoding matrix. We assume that $E\left[\boldsymbol{b}_n \boldsymbol{b}_n^{\mathrm{H}}\right] = \boldsymbol{I}_{M_{b,n}}$, where $E[\cdot]$ denotes statistical expectation, $(\cdot)^{\mathrm{H}}$ stands for Hermitian transpose, and $\boldsymbol{I}_n$ is an $n \times n$ identity matrix. Also we define the transmit covariance matrix as $\boldsymbol{S}_n = E\left[\boldsymbol{s}_n \boldsymbol{s}_n^{\mathrm{H}}\right]$. The consumed transmit power on the nth channel by the secondary transmitter can be written as

$$P_n = \mathrm{tr}\left(\boldsymbol{S}_n\right) = \mathrm{tr}\left(\boldsymbol{U}_n \boldsymbol{U}_n^{\mathrm{H}}\right), \tag{2}$$

where $\mathrm{tr}(\boldsymbol{X})$ denotes the trace of matrix $\boldsymbol{X}$.

The $M_{\mathrm{r,s}} \times 1$ signal vector received on the nth channel at the secondary receiver is expressed as

$$\boldsymbol{y}_n = \boldsymbol{H}_n \boldsymbol{s}_n + \boldsymbol{v}_n, \tag{3}$$

where $\boldsymbol{H}_n$ is the $M_{\mathrm{r,s}} \times M_{\mathrm{t,s}}$ MIMO fading channel matrix between the secondary transmitter and secondary receiver, and $\boldsymbol{v}_n$ is the $M_{\mathrm{r,s}} \times 1$ independent and identically distributed additive white Gaussian noise vector. We assume that the noises are complex circularly symmetric with zero mean and a variance of ρ.

By employing the linear minimum-mean-square-error (MMSE) method, the estimated signal vector at the secondary receiver on the nth channel is

$$\hat{\boldsymbol{b}}_n = \boldsymbol{W}_n^{\mathrm{H}} \boldsymbol{y}_n, \tag{4}$$

where $\boldsymbol{W}_n$ is the $M_{b,n} \times M_{\mathrm{r,s}}$ weight matrix of the linear MMSE receiver. According to [23], $\boldsymbol{W}_n$ can be derived as

$$\boldsymbol{W}_n = \left(\boldsymbol{H}_n \boldsymbol{U}_n \boldsymbol{U}_n^{\mathrm{H}} \boldsymbol{H}_n^{\mathrm{H}} + \boldsymbol{C}_{v_n}\right)^{-1} \boldsymbol{H}_n \boldsymbol{U}_n, \tag{5}$$

where $(\cdot)^{-1}$ denotes the matrix inversion, and $\boldsymbol{C}_{v_n}$ is the noise covariance matrix, which can be clearly defined as $\boldsymbol{C}_{v_n} = E\left[\boldsymbol{v}_n \boldsymbol{v}_n^{\mathrm{H}}\right] = \rho \boldsymbol{I}_{M_{r,s}}$. In [24], mathematical deductions were conducted to show that the diagonal elements of the mean-square-error (MSE) matrix can be used to measure system performance and accurately formulate important parameters, such as signal-to-interference-noise ratio (SINR) and mutual information. The estimated MSE matrix of the linear MMSE receiver is given by

$$\begin{aligned} \boldsymbol{E}_n &= E\left[\left(\hat{\boldsymbol{b}}_n - \boldsymbol{b}_n\right)\left(\hat{\boldsymbol{b}}_n - \boldsymbol{b}_n\right)^{\mathrm{H}}\right] \\ &= \left(\rho^{-1} \boldsymbol{U}_n^{\mathrm{H}} \boldsymbol{H}_n^{\mathrm{H}} \boldsymbol{H}_n \boldsymbol{U}_n + \boldsymbol{I}_{M_{\mathrm{t,s}}}\right)^{-1}. \end{aligned} \tag{6}$$

The SINR of the jth data stream on the nth channel is expressed as

$$\mathrm{SINR}_{j,n} = 1/\left[\boldsymbol{E}_n\right]_{j,j} - 1, \tag{7}$$

where $[\boldsymbol{X}]_{m,n}$ denotes the (m, n)th element of matrix $\boldsymbol{X}$.

The total interference constraint on the nth channel introduced to each primary receiver over all receive antennas is

$$\mathrm{tr}\left(\boldsymbol{G}_{k,n} \boldsymbol{S}_n \boldsymbol{G}_{k,n}^{\mathrm{H}}\right) \le \Gamma_k, \quad k = 1, \ldots, K, \tag{8}$$

where $\boldsymbol{G}_{k,n}$ is the $M_{k,\mathrm{p}} \times \mathrm{M}_{\mathrm{t,s}}$ MIMO fading channel matrix between the secondary transmitter and the kth primary receiver, and Γ_k stands for the total power interference constraint for the kth primary receiver over all receive antennas. In [13], the author proved the fact that the capacity loss of the kth primary receiver because of the interference from the secondary transmitter is limited if the transmit covariance matrix $\boldsymbol{S}_n$ satisfies (8). Then the quality of service of primary transmission can be exactly guaranteed through dynamically adapting parameter Γ_k.

Now we are ready to develop the mathematical optimisation problem described as follows: (***P1***):

$$\min\left\{F_{\mathrm{B}} F_{\mathrm{P}}\right\}: F_{\mathrm{B}} = \sum_{n=1}^{N} \boldsymbol{x}_n B,\ F_{\mathrm{P}} = \sum_{n=1}^{N} \boldsymbol{x}_n \mathrm{tr}\left(\boldsymbol{U}_n \boldsymbol{U}_n^{\mathrm{H}}\right), \tag{9}$$

subject to

$$\sum_{n=1}^{N} \mathrm{tr}\left(\boldsymbol{G}_{k,n} \boldsymbol{S}_n \boldsymbol{G}_{k,n}^{\mathrm{H}}\right) \le \Gamma_k, \quad k = 1, 2, \ldots, K, \tag{10}$$

$$\sum_{n=1}^{N} \sum_{j=1}^{J_n} \boldsymbol{x}_n R_{j,n} \ge \phi, \tag{11}$$

$$\forall \boldsymbol{x}_n \in \{0, 1\}, \quad n = 1, \ldots, N, \tag{12}$$

where $\boldsymbol{x}_n$ is a binary variable. It is equal to zero if the nth channel is not used for the secondary transmission, and is one otherwise. We define the channel resource budget $\mathcal{C}$ for the secondary transmission as $\mathcal{C} = \left\{n | \boldsymbol{x}_n = 1\right\}$. Besides, ϕ is the minimum rate requirement for the secondary transmission and J_n is the total number of data streams that the nth channel can bear. $R_{j,n}$ is the rate of the jth data stream on the nth channel, and written as

$$R_{j,n} = B \log\left(1 + \mathrm{SINR}_{j,n}\right). \tag{13}$$

Equation (10) is equivalent to the interference power constraints at the primary receivers. Here, the total interferences over all receive antennas and channels for each primary receiver are limited. Besides, the constraint in (11) corresponds to satisfying the minimum rate requirement for the secondary transmission.

4 Performance analysis

In this section, detailed analysis is made to fundamentally solve the problem ***P1***. The complete process consists of three steps. We begin by calculating the optimal structure of the source precoding matrix for each channel while the interference power constraints (10) are not considered. Then, the PC-SVD and DC-SVD methods are introduced to properly solve the interference constraints. Besides, advantages and disadvantages of each method are discussed in detail. Finally, the simplified power and channel resource optimisation problem is derived and found to be a mixed-integer programming problem. Hence, we propose a sub-optimal and tractable algorithm with low complexity. The innovative idea is to relax the channel selection so that the power optimisation subproblem and channel allocation subproblem can be independently performed using the Lagrange-duality theory [25] and Gauss–Newton method [26], respectively. Note that as to determining the channel resource set $\mathcal{C}$ in the power optimisation process, we design a simple criterion to measure the quality of each channel and sort all channels in the decreasing order, then $\mathcal{C}$ can be easily obtained by choosing the best N^* channels if N^* channels are assumed to be occupied.

4.1 Optimal source precoding matrix

In this subsection, we are ready to design the optimal structure of the source precoding matrix for each channel by neglecting interference power constraints (10). The corresponding subproblem can be expressed as (***P2***)

$$\min_{\boldsymbol{U}_n} \mathrm{tr}\left(\boldsymbol{U}_n \boldsymbol{U}_n^{\mathrm{H}}\right), \tag{14}$$

$$\text{s.t.} \sum_{j=1}^{J_n} R_{j,n} \ge \phi. \tag{15}$$

Before putting forwarding the critical theorem on the solution of the subproblem, two important definitions in majorisation theory need to be introduced according to [27].

Definition 1: [18, 1.A.1]: Consider any two $N \times 1$ vectors $\boldsymbol{x}, \boldsymbol{y}$, sorted in the decreasing order, which means $\boldsymbol{x}_{[1]} \ge \boldsymbol{x}_{[2]} \ge \cdots \boldsymbol{x}_{[N]}$, $\boldsymbol{y}_{[1]} \ge \boldsymbol{y}_{[2]} \ge \cdots \boldsymbol{y}_{[N]}$. Vector $\boldsymbol{y}$ majorises vector $\boldsymbol{x}$, represented as $\boldsymbol{x} \prec \boldsymbol{y}$, if $\sum_{i=1}^{n} \boldsymbol{x}_{[i]} \le \sum_{i=1}^{n} \boldsymbol{y}_{[i]}$, $n = 1, 2, \ldots, N-1$, and $\sum_{i=1}^{N} \boldsymbol{x}_{[i]} = \sum_{i=1}^{N} \boldsymbol{y}_{[i]}$.

Definition 2: [18, 3.A.1]: A real value function f is called Schur-concave if $f(\boldsymbol{x}) \ge f(\boldsymbol{y})$ for $\boldsymbol{x} \prec \boldsymbol{y}$, and called Schur-convex if $f(\boldsymbol{x}) \le f(\boldsymbol{y})$ for $\boldsymbol{x} \prec \boldsymbol{y}$.

Equation (15) can be identically expressed as

$$q\left(\boldsymbol{d}\left(\boldsymbol{E}_n\right)\right) \ge \phi, \tag{16}$$

where $\boldsymbol{d}(\boldsymbol{X})$ denotes a column vector including all the main diagonal entry of matrix $\boldsymbol{X}$, and $q(\boldsymbol{x}) = \sum_{j=1}^{J_n} -B \log\left(\boldsymbol{x}_j\right)$.

Denote the singular value decomposition (SVD) of $\boldsymbol{H}_n$ as

$$\boldsymbol{H}_n = \Omega_{H_n}\boldsymbol{\Lambda}_{H_n}^{1/2}\left(\boldsymbol{V}_{H_n}\right)^{\mathrm{H}}, \tag{17}$$

where the dimensions of Ω_{H_n}, $\boldsymbol{\Lambda}_{H_n}$, and $\boldsymbol{V}_{H_n}$ are $M_{\mathrm{r,s}} \times M_{\mathrm{r,s}}$, $M_{\mathrm{r,s}} \times M_{\mathrm{t,s}}$, and $M_{\mathrm{t,s}} \times M_{\mathrm{t,s}}$ respectively. Also, we assume the diagonal elements of $\boldsymbol{\Lambda}_{H_n}$ are sorted in the decreasing order.

Theorem 1: We assume $J_n = M_{b,n} = \mathrm{rank}(\boldsymbol{H}_n)$. As to the problem ***P2***, the optimal source precoding matrix $\boldsymbol{U}_n$ is given by

$$\boldsymbol{U}_n = \boldsymbol{V}_{H_n,1}\boldsymbol{\Lambda}_{U_n}^{1/2}, \tag{18}$$

where the $M_{\mathrm{t,s}} \times M_{b,n}$ matrix $\boldsymbol{V}_{H_n,1}$ contains the rightmost $M_{b,n}$ columns from $\boldsymbol{V}_{H_n}$. The $M_{b,n} \times M_{b,n}$ matrix $\boldsymbol{\Lambda}_{U_n}$ can be expressed as $\boldsymbol{\Lambda}_{U_n} = \mathrm{diag}\left(\boldsymbol{\sigma}_{U_n}\right)$, where $\boldsymbol{\sigma}_{U_n} = \left[\sigma_{1,n}, \ldots \sigma_{M_{b,n},n}\right]^{\mathrm{T}}$ is the $M_{b,n} \times 1$ vector whose elements need to be optimised and diag($\mathbf{x}$) stands for a diagonal matrix whose main diagonal elements vector is $\boldsymbol{x}$.

Proof: See Appendix 1. □

The assumption we make in Theorem 1, namely $J_n = M_{b,n} = \mathrm{rank}$ ($\boldsymbol{H}_n$), maximises the number of independent data streams on each channel so that no transmit power is wasted. Note that the optimal source precoding matrix diagonalises the source–destination channel matrix and converts the MIMO channel into multiple parallel single-input single-output (SISO) spatial-channels.

4.2 Interference power constraints analysis

In this subsection, we will take (10) into consideration for the optimisation problem. From Theorem 1, we can infer that solutions on interference power constraints depend on the SVD of $\boldsymbol{H}_n$. Besides, the projected-channel and direct-channel methods were introduced to properly solve this problem in [13]. Therefore, we apply the PC-SVD and DC-SVD algorithms to deal with (10) in this paper. Also the advantages and disadvantages of each algorithm are explicitly analysed.

4.2.1 Projected-channel SVD: In the PC-SVD, using the zero-forcing technique, $\boldsymbol{U}_n$ is properly designed so that the interferences at the primary receivers introduced from the secondary transmission are completely avoided. Note that the PC-SVD is effective if and only if the number of transmit antennas is larger than that of total PUs' receive antennas, as the reason will be discussed later. We denote the $M_{\mathrm{p}} \times \mathrm{M}_{\mathrm{t,s}}$ matrix $\boldsymbol{G}_n$ as the nth channel from the secondary transmitter to all the primary receivers, where $M_{\mathrm{p}} = \sum_{k=1}^{K} M_{k,\mathrm{p}}$. Alternatively, $\boldsymbol{G}_n = \left[\boldsymbol{G}_{1,n}^{\mathrm{T}}, \ldots, \boldsymbol{G}_{K,n}^{\mathrm{T}}\right]^{\mathrm{T}}$, where $\boldsymbol{X}^{\mathrm{T}}$ stands for the transpose of $\boldsymbol{X}$. The SVD of $\boldsymbol{G}_n^{\mathrm{H}}$ is expressed as

$$\boldsymbol{G}_n^{\mathrm{H}} = \boldsymbol{\Omega}_{G_n^{\mathrm{H}}}\boldsymbol{\Lambda}_{G_n^{\mathrm{H}}}^{1/2}\boldsymbol{V}_{G_n^{\mathrm{H}}}^{\mathrm{H}}, \tag{19}$$

where the dimensions of $\boldsymbol{\Omega}_{G_n^{\mathrm{H}}}$, $\boldsymbol{\Lambda}_{G_n^{\mathrm{H}}}$, and $\boldsymbol{V}_{G_n^{\mathrm{H}}}$ are $M_{\mathrm{t,s}} \times M_{\mathrm{t,s}}$, $M_{\mathrm{t,s}} \times M_{\mathrm{p}}$, and $M_{\mathrm{p}} \times M_{\mathrm{p}}$, respectively. As to the PC-SVD, the secondary channel $\boldsymbol{H}_n$ is first projected into the null space of $\boldsymbol{G}_n^{\mathrm{H}}$, which is

$$\boldsymbol{H}_{\perp,n} = \boldsymbol{H}_n\left(\boldsymbol{I}_{M_{\mathrm{t,s}}} - \boldsymbol{\Omega}_{G_n^{\mathrm{H}},1}\left(\boldsymbol{\Omega}_{G_n^{\mathrm{H}},1}\right)^{\mathrm{H}}\right), \tag{20}$$

where the $M_{\mathrm{t,s}} \times \mathrm{rank}\left(\boldsymbol{G}_n^{\mathrm{H}}\right)$ matrix $\boldsymbol{\Omega}_{G_n^{\mathrm{H}},1}$ contains the rightmost columns corresponding to non-zero singular values. According to (20), $\boldsymbol{H}_{\perp,n}$ becomes a zero matrix if $M_{\mathrm{t,s}} \le M_{\mathrm{p}}$. Therefore, the PC-SVD can only be meaningful when $M_{\mathrm{t,s}} > M_{\mathrm{p}}$. Define the SVD of $\boldsymbol{H}_{\perp,n}$ as

$$\boldsymbol{H}_{\perp,n} = \boldsymbol{\Omega}_{H_{\perp,n}}\boldsymbol{\Lambda}_{H_{\perp,n}}^{1/2}\left(\boldsymbol{V}_{H_{\perp,n}}\right)^{\mathrm{H}}, \tag{21}$$

where the diagonal elements of $\boldsymbol{\Lambda}_{H_{\perp,n}}$ are sorted in the decreasing order. By applying Theorem 1 to (20), we can obtain

$$\boldsymbol{U}_n = \boldsymbol{V}_{H_{\perp,n},1}\boldsymbol{\Lambda}_{U_n}^{1/2}, \tag{22}$$

where the $M_{\mathrm{t,s}} \times M_{b,n}$ matrix $\boldsymbol{V}_{H_{\perp,n},1}$ contains rightmost $M_{b,n}$ columns from $\boldsymbol{V}_{H_{\perp,n}}$ and $M_{b,n} = \mathrm{rank}(\boldsymbol{H}_{\perp,\mathrm{n}}) = \min(\mathrm{M}_{\mathrm{r,s}}, M_{\mathrm{t,s}} - M_{\mathrm{p}})$.

Theorem 2: The interferences introduced to primary receivers can be completely eliminated by designing $\boldsymbol{U}_n$ according to (22).

Proof: See Appendix 2. □

4.2.2 Direct-channel SVD: In the DC-SVD, (18) is applied to design the $\boldsymbol{U}_n$, while the interference constraints are satisfied by adjusting the power allocation $\boldsymbol{\Lambda}_{U_n}$. Then $M_{b,\ n} = \min(\mathrm{M}_{\mathrm{t,s}}, M_{\mathrm{r,s}})$. We denote $\boldsymbol{G}_{k,n}^{\mathrm{H}} = \left[\boldsymbol{g}_{1,n}^{k}, \ldots, \boldsymbol{g}_{M_{k,p},n}^{k}\right]$, $k = 1, \ldots, K$, $n = 1, \ldots, N$, and $\boldsymbol{V}_{H_n,1} = [\boldsymbol{v}_{1,n}, \ldots, \boldsymbol{v}_{M_{b,n},n}]$, $n = 1, \ldots, N$. Note that $\boldsymbol{g}_{1,n}^{k}, \ldots, \boldsymbol{g}_{M_{k,p},n}^{k}$, $k = 1, \ldots, K$, $n = 1, \ldots, N$, and $\boldsymbol{v}_{1,n}, \ldots, \boldsymbol{v}_{M_{b,n}}$, $n = 1, \ldots, N$ are all $M_{\mathrm{t,s}} \times 1$ vectors. Besides, we define $I_{j,m,n}^{k}$ as the interference introduced from the jth data stream transmission to the mth receive antenna of the kth primary receiver on the nth channel. Then, it is not difficult to find that

$$I_{j,m,n}^{k} = \alpha_{j,m,n}^{k}\sigma_{j,n}, \tag{23}$$

where $\alpha_{j,m,n}^{k} = \left(\left(\boldsymbol{g}_{m,n}^{k}\right)^{\mathrm{T}}\boldsymbol{v}_{j,n}\right)^2$, $j = 1, \ldots, M_{b,n}$, $k = 1, \ldots, K$, $m = 1, \ldots, M_{k,\mathrm{p}}$, and $n = 1, \ldots, N$. According to (23), constraints (10) can be equivalently expressed as

$$\sum_{j=1}^{M_{b,n}}\sum_{m=1}^{M_{k,\mathrm{p}}}\sum_{n=1}^{N} I_{j,m,n}^{k} = \sum_{j=1}^{M_{b,n}}\sum_{n=1}^{N}\bar{\alpha}_{j,n}^{k}\sigma_{j,n} \le \Gamma_k, \quad k = 1, 2\ldots, K, \tag{24}$$

where $\bar{\alpha}_{j,n}^{k} = \sum_{m=1}^{M_{k,\mathrm{p}}}\alpha_{j,m,n}^{k}$.

In order to make deduction much more tractable, this paper places more stringent requirements on interference constraints. Here, the interference introduced from each data stream transmission on the nth channel to each primary receive antenna is guaranteed to be below threshold of tolerance. Under such restriction, the power allocated in a data stream, which is transmitted on an excellent secondary channel, may be greatly reduced because of severe interference to a single primary receive antenna. Hence the results obtained in this paper provide a lower-bound for the DC-SVD algorithm.

In the similar way, we denote $\boldsymbol{G}_n^{\mathrm{H}} = \left[\hat{\boldsymbol{g}}_{1,n}, \ldots, \hat{\boldsymbol{g}}_{M_{\mathrm{p}},n}\right]$, where $\hat{\boldsymbol{g}}_{1,n}, \ldots, \hat{\boldsymbol{g}}_{M_{\mathrm{p}},n}$ are all $M_{\mathrm{t,s}} \times 1$ vectors. Besides, the interference introduced from the jth data stream transmission to the kth primary receive antenna on the nth channel is defined as $\hat{I}_{j,k,n}$. Then, the interference constraints can be derived as

$$\hat{I}_{j,k,n} = \hat{\alpha}_{j,k,n}\sigma_{j,n} \le \Gamma, \tag{25}$$

where $\hat{\alpha}_{j,k,n} = \left(\hat{\boldsymbol{g}}_{k,n}^{\mathrm{T}}\boldsymbol{v}_{j,n}\right)^2$, $j = 1, \ldots, M_{b,n}$, $k = 1, \ldots, M_{\mathrm{p}}$, and $n = 1, \ldots, N$. Also Γ stands for the corresponding interference threshold and is set equal for all the interference constraints. Furthermore, (25) can be equivalently written as

$$\sigma_{j,n} \le \Gamma/\tilde{\alpha}_{j,n}, \quad \forall j \in \{1, \ldots, M_{b,n}\}, \quad \forall n \in \{1, \ldots, N\}, \quad (26)$$

where $\tilde{\alpha}_{j,n} = \max_k \{\hat{\alpha}_{j,k,n}\}$, $k \in \{1, \ldots, M_p\}$.

4.2.3 Performance comparison: As discussed above, both the PC-SVD and DC-SVD algorithms rely on the SVD of the secondary channel matrix $\boldsymbol{H}_n$ and decompose the secondary MIMO channel matrix into multiple SISO spatial-channels. The main difference between them lies in that, the PC-SVD performs the channel decomposition which is based on the SVD of $\boldsymbol{H}_{\perp,n}$ (the projection of $\boldsymbol{H}_n$ into the null space of $\boldsymbol{G}_n$), while for the DC-SVD, the channel decomposition directly depends on the SVD of $\boldsymbol{H}_n$. Then, the DC-SVD accordingly maintains the maximum number of data streams each channel can hold and restricts the transmit power of each data stream to satisfy the interference constraints. Meanwhile the PC-SVD reduces the total number of data streams in order to make the transmit space lie in the null space of primary channels and removes the interferences totally.

Furthermore, as to the computational complexity, the PC-SVD has a larger complexity for obtaining the optimal precoding matrix because of the extra channel projection (20), while so does the DC-SVD for determining the optimal power allocation due to the interference constraints (24). Besides, according to [13], the PC-SVD is optimal if the interference threshold Γ is exactly zero. On the other hand, while the Γ is sufficiently large so that (26) is inactive, then the DC-SVD is optimal indeed. Finally, under the good condition, which means the channel allocation rather than the power optimisation dominates the optimisation problem, it is conjectured that the DC-SVD outperforms the PC-SVD, while on the contrary, the PC-SVD provides a better performance than the DC-SVD under the poor condition. Simulation results discussed in Section 5 validate this conjecture.

4.3 Power optimisation and channel allocation

In this subsection, a unified framework considering both power optimisation and channel allocation is developed. To start with, several important definitions are listed below in order to make the deduction simpler. We denote $\boldsymbol{\Lambda}_{H_n,1}$ and $\boldsymbol{\Lambda}_{H_{\perp,n},1}$ as the diagonal matrices, whose diagonal elements contain all the non-zero diagonal elements from $\boldsymbol{\Lambda}_{H_n}$ and $\boldsymbol{\Lambda}_{H_{\perp,n}}$, respectively. Correspondingly, they can be expressed as

$$\Lambda_{H_n,1} = \mathbf{diag}\left(\boldsymbol{\lambda}^1_{H_n}\right), \quad \boldsymbol{\Lambda}_{H_{\perp,n},1} = \mathbf{diag}\left(\boldsymbol{\lambda}^1_{H_{\perp,n}}\right), \quad (27)$$

where $\boldsymbol{\lambda}^1_{H_n} = \left[\lambda_{H_n,1}, \ldots, \lambda_{H_n,M_{b,n}}\right]^{\mathrm{T}}$, and $\boldsymbol{\lambda}^1_{H_{\perp,n}} = \left[\lambda_{H_{\perp,n},1}, \ldots, \lambda_{H_{\perp,n},M_{b,n}}\right]^{\mathrm{T}}$. Both the $\boldsymbol{\lambda}^1_{H_n}$ and $\boldsymbol{\lambda}^1_{H_{\perp,n}}$ are $M_{b,n} \times 1$ vectors, while their elements are sorted in the decreasing order. Besides, we denote $\bar{\boldsymbol{\lambda}}_n$ as

$$\bar{\boldsymbol{\lambda}}_n = \begin{cases} \boldsymbol{\lambda}^1_{H_{\perp,n}}, \mathrm{PC-SVD} \\ \boldsymbol{\lambda}^1_{H_n}, \mathrm{DC-SVD} \end{cases}. \quad (28)$$

Note that $M_{b,n}$ can be easily obtained as

$$M_{b,n} = \begin{cases} \min\left(M_{\mathrm{t,s}} - M_{\mathrm{p}}, M_{r,\mathrm{s}}\right), \mathrm{PC-SVD} \\ \min\left(M_{\mathrm{t,s}}, M_{\mathrm{r,s}}\right), \mathrm{DC-SVD} \end{cases}. \quad (29)$$

Furthermore, we define $\bar{\sigma}_{j,n}$ as

$$\bar{\sigma}_{j,n} = \begin{cases} \infty, \mathrm{PC-SVD} \\ \Gamma/\tilde{\alpha}_{j,n}, \mathrm{DC-SVD} \end{cases}, \quad (30)$$

where ∞ means the threshold is sufficiently large so that $\sigma_{j,n}$ is not restricted by $\bar{\sigma}_{j,n}$ any more. Note that, as to the DC-SVD, we use (26) instead of (24) to satisfy the interference constraints, thus a lower-bound for the DC-SVD is provided in this paper.

Next, we are ready to deal with the joint power optimisation and channel allocation problem. Based on the analysis results and definitions discussed above, the problem ***P1*** can be equivalently written as (***P5***):

$$\min\{F_{\mathrm{B}} F_{\mathrm{P}}\}: F_{\mathrm{B}} = \sum_{n=1}^{N} x_n B, \quad F_{\mathrm{P}} = \sum_{n=1}^{N} \sum_{j=1}^{M_{b,n}} x_n \sigma_{j,n}, \quad (31)$$

subject to

$$\sigma_{j,n} \le \bar{\sigma}_{j,n}, \quad j = 1, \ldots, M_{b,n}, \quad n = 1, \ldots, N, \quad (32)$$

$$\sum_{n=1}^{N} \sum_{j=1}^{M_{b,n}} x_n B \log\left(1 + \frac{\sigma_{j,n}\bar{\lambda}_{j,n}}{\rho}\right) \ge \phi, \quad (33)$$

$$\forall x_n \in \{0, 1\}, \quad n = 1, \ldots, N, \quad (34)$$

where $\bar{\lambda}_{j,n}$ denotes the jth element in $\bar{\boldsymbol{\lambda}}_n$. It is not difficult to see that (34) makes the problem ***P5*** a mixed integer programming, and an exhaustive search over all feasible regions is required to achieve the optimal solution. Hence, a sub-optimal and tractable algorithm with low-complexity is developed in this subsection. Here, the problem ***P5*** is further decomposed into two parts: the power optimisation subproblem and channel allocation subproblem. The innovative idea is to relax the channel selection so that each subproblem can be performed independently. As to the power optimisation subproblem, for the given channel resource budget $\mathcal{C}$, a simple and iterative algorithm using the Lagrangian–dual method [25] is proposed. While in the channel allocation subproblem, in order to minimise the BPP, the optimal channel resource set $\mathcal{C}$ is iteratively obtained using the Gauss–Newton method [26]. Note that the critical point for determining the set $\mathcal{C}$ is to sort all the N channels in a decreasing order according to the quality of each channel. Designing the quality evaluation criterion is a challenging issue and will be discussed later. Once channels are ordered, determining the set $\mathcal{C}$ is refined to select the best N^* channels if N^* channels are occupied.

4.3.1 Power-optimisation: For the given channel resource budget $\mathcal{C}$, namely the best N^* channels are occupied for the secondary transmission, the power optimisation problem can be written as (***P6***):

$$\min\{F^*_{\mathrm{B}} F_{\mathrm{P}}\}: F^*_{\mathrm{B}} = |\mathcal{C}|B, \quad F_{\mathrm{P}} = \sum_{n\in\mathcal{C}} \sum_{j=1}^{M_{b,n}} \sigma_{j,n}, \quad (35)$$

subject to

$$\sigma_{j,n} \le \bar{\sigma}_{j,n}, \quad j = 1, \ldots, M_{b,n}, \quad n \in \mathcal{C}, \quad (36)$$

$$\sum_{n\in\mathcal{C}} \sum_{j=1}^{M_{b,n}} B \log\left(1 + \frac{\sigma_{j,n}\bar{\lambda}_{j,n}}{\rho}\right) \ge \phi, \quad (37)$$

where $|\mathcal{X}|$ denotes the cardinality of set $\mathcal{X}$. Here, according to the assumption, we can obtain $|\mathcal{C}| = N^*$. Then, based on the Lagrangian–dual theory [25], the corresponding Lagrangian function is derived by taking (37) into consideration as

$$L = F^*_{\mathrm{B}} F_{\mathrm{P}} + \nu\left(\phi - \sum_{n\in\mathcal{C}} \sum_{j=1}^{M_{b,n}} B \log\left(1 + \frac{\sigma_{j,n}\bar{\lambda}_{j,n}}{\rho}\right)\right), \quad (38)$$

where ν is a Lagrange multiplier corresponding to constraint (37) and $\nu \geq 0$. Based on the Karush–Kuhn–Tucker (KKT) condition, the optimal solution of $\sigma_{j,n}$ can be derived by minimising the function L respect to $\sigma_{j,n}$ as

$$\sigma_{j,n} = \left[\nu - \frac{1}{\bar{\lambda}_{j,n}/\rho}\right]^{+}, \tag{39}$$

where $[\boldsymbol{x}]^{+} = \max\{0, \boldsymbol{x}\}$. In order to satisfy the constraint (36), (39) is adjusted as

$$\sigma_{j,n} = \min\left\{\bar{\sigma}_{j,n}, \left[\nu - \frac{1}{\bar{\lambda}_{j,n}/\rho}\right]^{+}\right\}, \tag{40}$$

where $j = 1, \ldots, M_{b,n}$ and $n \in \mathcal{C}$. Furthermore, the Lagrange multiplier ν can be obtained when the equality of (37) holds. It is not difficult to see that in (40), a smaller ν will result in a lower system throughout [the left part of (37)], and vice versa. Therefore, the optimal ν can be efficiently obtained by the bisection method [25] through comparing the optimal system throughout for a given ν with predefined throughout threshold ϕ. In conclusion, the following Algorithm 1 (see Fig. 2) can be properly used to solve the power optimisation problem.

As to determining the set $\mathcal{C}$, designing the criterion for evaluating the quality of channels is the key step and still left to be discussed. Here, we consider a special case, where $\bar{\sigma}_{j,n}$, $j = 1, \ldots, M_{b,n}$, $n \in \mathcal{C}$ are all sufficiently large so that (36) is inactive, in other words, the PC-SVD method is applied or channels between the secondary transmitter and primary receivers are all in severe fading. Besides, we denote the $S \times 1$ vector $\tilde{\boldsymbol{\lambda}}$ as $\tilde{\boldsymbol{\lambda}} = [\tilde{\lambda}_1, \ldots, \tilde{\lambda}_S]^{\mathrm{T}}$, where $S = \sum_{n \in \mathcal{C}} M_{b,n}$. The $\tilde{\boldsymbol{\lambda}}$ can be directly obtained by sorting all the elements of the gain set $\left\{\bar{\lambda}_{j,n}, j = 1, \ldots, M_{b,n}, n \in \mathcal{C}\right\}$ in the decreasing order. Define Q_n as

$$Q_n = \frac{1}{\prod_{\tilde{\lambda}_j \in \boldsymbol{C}_n} \tilde{\lambda}_j}, \tag{41}$$

where $\boldsymbol{C}_n = \left\{\bar{\lambda}_{j,k}, j = 1, \ldots, M_{b,n}, k = n \in \mathcal{C}\right\}$.

Remark 1: Q_n is indeed an outstanding and simple criterion for evaluating the quality of channels, and the smaller Q_n is, the better quality of the channel is. This conjecture is proved in Appendix 3. The numerical results discussed in Section 5.1 further validate this remark.

4.3.2 Channel allocation:

Once the power optimisation process is accomplished for a given channel resource set $\mathcal{C}$, we continue to determine the optimal $\mathcal{C}$ in order to minimise the BPP. As discussed above, based on the assumption that all channels are sorted in the decreasing order, the channel allocation problem is refined to select the best N channels, where N need to be optimised as

$$\min_{N}\left\{NB \times \left(\sum_{n=1}^{N}\sum_{j=1}^{M_{b,n}} \min\left\{\bar{\sigma}_{j,n}, \left[\nu^{*} - \frac{1}{\bar{\lambda}_{j,n}/\rho}\right]^{+}\right\}\right)\right\}, \tag{42}$$

where ν^* is the optimal solution of ν according to Fig. 2. Problem (42) can be iteratively solved using the Gauss–Newton method [26].

Although the closed-form expression of optimal N is hard to obtain, it can be tractable under the special case: (i) $\bar{\sigma}_{j,n}, j = 1, \ldots, M_{b,n}, n \in \mathcal{C}$ are sufficiently large so that (36) is inactive; (ii) the variations of the channel gain-to-interference-and-noise ratio over all the candidate subspaces are very small and all the secondary channel matrices are full-rank, namely $\tilde{\lambda}_j \simeq \tilde{\lambda}, j = 1, \ldots, S$ and $M_{b,n} = \bar{M}, n = 1, \ldots, N$; and (iii) the rate requirement ϕ is large enough to make all the subspaces be occupied for the secondary transmission for a given $\mathcal{C}$, i.e. $S^* = S = N\bar{M}$. Note that this case is more likely to happen in indoor communication where the coherence bandwidth is usually larger compared with outdoor communication [28]. Then, (42) can be derived as

$$\min_{N}\left\{NB \times \left(\sum_{s=1}^{N\bar{M}}\left(\frac{\rho 2^{\frac{\phi}{(BN\bar{M})}}}{\tilde{\lambda}} - \frac{1}{\tilde{\lambda}/\rho}\right)\right)\right\}. \tag{43}$$

Based on the KKT condition, we can obtain $N_{\mathrm{opt}} = \dfrac{0.435\phi}{B\bar{M}}$.

4.3.3 Complexity analysis:

As mentioned above, the sub-optimal algorithm proposed in this paper decomposes the unified resource optimisation problem into two independent parts, namely power optimisation subproblem and channel allocation subproblem. The bisection method is applied to deal with the power optimisation subproblem (Fig. 2), while the channel allocation subproblem is optimised using the Gauss–Newton method. Note that both the bisection method and Gauss–Newton

Algorithm 1

Input:

$\mathcal{C}, \rho, B, \phi$;

$\{\bar{\lambda}_{j,n}\}, \{\bar{\sigma}_{j,n}\}, j = 1, \ldots M_{b,n}, n \in \mathcal{C}$;

δ_ν: a small positive constant that controls the accuracy of algorithm;

$\hat{\nu}$: a positive constant that sets the upper bound for ν;

Output:

$\{\sigma_{j,n}\}, j = 1, \ldots M_{b,n}, n \in \mathcal{C}$;

```
Initialize: v_min = 0, ν_max = ν̂;
while ((ν_max − v_min) > δ_ν) do
    Set ν = (ν_max + v_min)/2;
    Obtain {σ_j,n} according to (40);
    If Σ_{n∈C} Σ_{j=1}^{M_b,n} B log(1 + σ_j,n λ̄_j,n / ρ) ≥ φ, set ν_max = ν; otherwise, ν_min = ν;
end while
return {σ_j,n};
```

Fig. 2 *Algorithm 1 Power optimisation algorithm* method have low complexity for obtaining the optimal solutions, we can conclude that the corresponding algorithm indeed is low-complexity. As to the worst case, the channel resource budget $\mathcal{C}$ takes N possibilities into consideration, and for each case, $N_v \sum_{n=1}^{N} M_{b,n}$ calculations are needed to obtain the optimal solution for the power allocation subproblem, where N_v is the number of iterations for the bisection method. Therefore, the sub-optimal algorithm proposed takes at most $NN_v \sum_{n=1}^{N} M_{b,n}$ operations. Note that $\sum_{n=1}^{N} M_{b,n}$ in the DC-SVD is larger than that in the PC-SVD, therefore the DC-SVD has a larger complexity for determining the optimal resource allocation, while so does the PC-SVD for obtaining $\bar{\boldsymbol{\lambda}}_n$, as discussed above.

4.4 Summary of the analysis

In conclusion, the original optimisation problem ***P1*** has been completely solved. The total solution consists of three procedures. Firstly, the optimal precoding matrix for each channel has been designed. Secondly, two methods, namely the PC-SVD and the DC-SVD, have been proposed to properly solve the interference power constraints. Thirdly, the unified power optimisation and channel allocation problem has been derived and a sub-optimal and tractable algorithm with low-complexity has been introduced. To be clear, the following Algorithm 2 (see Fig. 3) gives a summary of the solution in detail, where $N^{(t)}$ stands for the number of the channels used for the secondary transmission in the tth iteration.

5 Numerical results

In this section, several numerical results are provided to validate the performance of the proposed scheme. All channel gains $\boldsymbol{H} = \{\boldsymbol{H}_1, \ldots, \boldsymbol{H}_N\}$ and $\boldsymbol{G} = \{\boldsymbol{G}_{1,1}, \ldots, \boldsymbol{G}_{1,N}, \ldots, \boldsymbol{G}_{K,1}, \ldots, \boldsymbol{G}_{K,N}\}$, are assumed to be independent of each other and follow circularly-symmetric-complex-Gaussian distribution as $CN(0, 1)$ for $\boldsymbol{H}$, and $CN(0, 0.1)$ for $\boldsymbol{G}$, respectively. Without loss of generality, we set $N = 64$, $B = 1$, and $M_{\rm t,s} = M_{\rm r,s} = 4$. One thousand independent Monte Carlo simulations are implemented to average the experimental results.

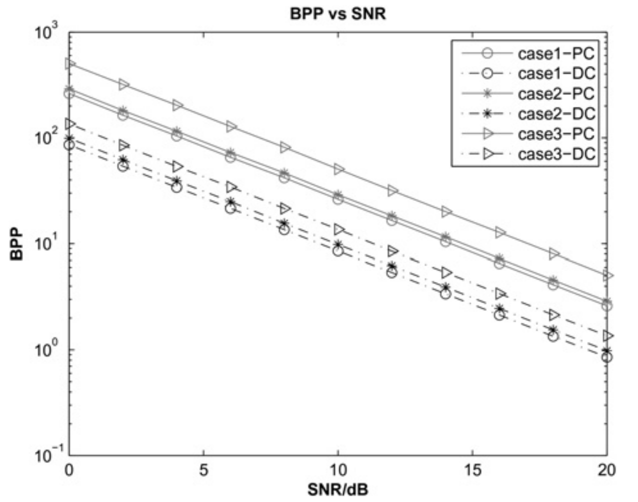

Fig. 4 *Example 1: Relationship between BPP and SNR for the proposed scheme under different evaluating criteria*

5.1 Performance comparison for different criteria of evaluating the channel quality

Fig. 4 compares performance for different criteria of evaluating the channel quality. We set $\Gamma = 0.1$, $\phi = 50$, $K = 3$, and $M_{k,\rm p} = 1$, $k = 1, \ldots, 3$. Also the noise variance can be obviously obtained by $\rho = 1/10^\wedge(\mathrm{SNR}/10)$. Three criteria to measure the quality of channel are considered: $Ca = \min_n Q_n = \min_n\left\{1/\left(\prod_{j=1}^{M_{b,n}} \bar{\lambda}_{j,n}\right)\right\}$, $Cb = \max_n\left\{\sum_{j=1}^{M_{b,n}} \bar{\lambda}_{j,n}\right\}$ and $Cc = \{\text{disorder/random}\}$. Then the circular, asteroid, and triangular shapes correspond to Ca, Cb, and Cc, respectively. Besides, the solid and dash-dotted lines represent the PC-SVD and the DC-SVD, respectively. According to Fig. 4, we can see that our proposed measuring criterion Ca outperforms both Cb and Cc. Furthermore, for the given parameters, the DC-SVD is superior to the PC-SVD, which is because the PC-SVD wastes too much transmit space resource under a good circumstance (large Γ and small ϕ).

Algorithm 2

Input:

$N, B, M_{t,s}, M_{r,s}, K, \{M_{k,p}, k = 1, \ldots K\}, \phi, \Gamma, \rho$;

$\mathbf{H} = \{\mathbf{H}_1, \ldots \mathbf{H}_N\}$; $\mathbf{G} = \{\mathbf{G}_{1,1}, \ldots, \mathbf{G}_{1,N}, \ldots \mathbf{G}_{K,1}, \ldots, \mathbf{G}_{K,N}\}$;

Parameters of the algorithm: $\mathcal{C}$, $N^{(t)}$, $\{\bar{\lambda}_{j,n}\}, \{\bar{\sigma}_{j,n}\}, \{\sigma_{j,n}\}, j = 1, \ldots M_{b,n}, n = 1, \ldots N$;

Output:

BPP;

1: Obtain $\{\bar{\lambda}_{j,n}\}, \{\bar{\sigma}_{j,n}\}, j = 1, \ldots M_{b,n}, n = 1, \ldots N$;

As to the PC-SVD method, $\bar{\lambda}_{j,n}$ can be acquired using the equations (19), (20), (21) and (22), while $\bar{\sigma}_{j,n}$ can be obtained according to the equation (30);

As to the DC-SVD method, $\bar{\lambda}_{j,n}$ can be achieved through the equation (18), while $\bar{\sigma}_{j,n}$ can be gained by employing the equations (25) and (26);

2: Set $N^{(0)} = 1$;

3: **while** $N^{(t+1)} \neq N^{(t)}$ **do**

4: Obtain $\mathcal{C}$ by choosing the best $N^{(t)}$ channels, where the criterion for evaluating the quality of the channel refers to the equation (41);

5: Obtain $\{\sigma_{j,n}\}$ according to the *Algorithm 1*;

6: Calculate the current BPP by employing the equation (31);

7: Update $N^{(t+1)}$ using the Gauss-Newton method;

8: **end while**

9: **return** BPP;

Fig. 3 *Algorithm 2: Resource optimisation algorithm using the BPP for the MIMO-OFDMA CRN*

IET Commun., 2015, Vol. 9, Iss. 14, pp. 1710–1720

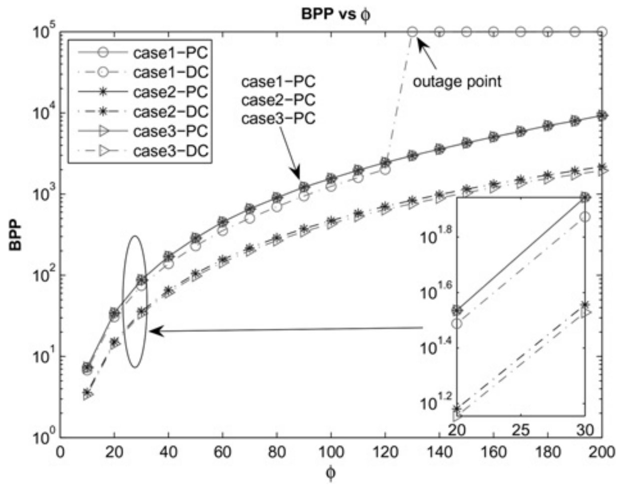

Fig. 5 *Example 2: Relationship between BPP and* ϕ *for the proposed scheme under different* Γ

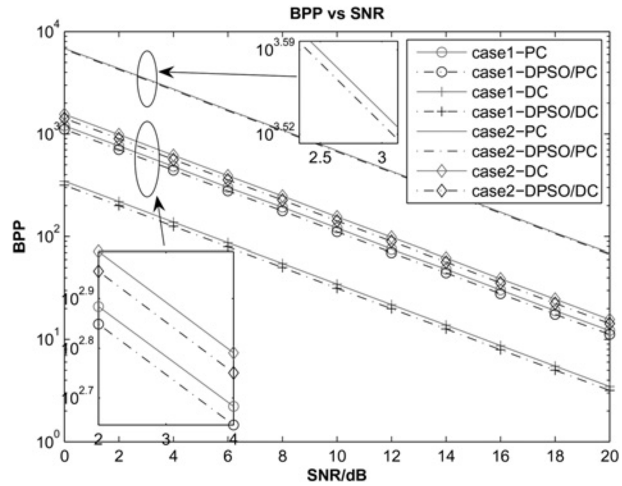

Fig. 7 *Example 4: Performance gap between the proposed sub-optimal resource optimisation algorithm and the optimal DPSO algorithm*

5.2 Performance evaluation for different ϕ

In this subsection, we are ready to study the relationship between BPP and ϕ as shown in Fig. 5 and relationship between optimum N and ϕ as shown in Fig. 6. We set $\rho = 1$, $K = 3$, and $M_{k,\mathrm{p}} = 1$, $k = 1, \ldots, 3$. Also the circular, asteroid, and triangular shapes stand for $\Gamma = 0.001$, $\Gamma = 0.01$, and $\Gamma = 0.1$, respectively. Besides, the solid and dash-dotted lines correspond to the PC-SVD and the DC-SVD, respectively.

In Fig. 5, we can observe that the PC-SVD converges to the same performance regardless of the Γ. This is because the PC-SVD confines the transmit space in the null space of primary channels and is independent of interference power constraints Γ. Furthermore, we can find that the PC-SVD is superior to the DC-SVD under the poor condition, namely stringent interference power constraints (small Γ) and high data rate requirement (large ϕ), and inferior to the DC-SVD under the good condition (large Γ and small ϕ). As analysed above, the PC-SVD sacrifices part of transmit space resource for unconstrained power optimisation in available transmit space. Then, under the good condition, the inferiority of the PC-SVD is due to wasting transmit space resource, while superiority of the PC-SVD results from optimising power resource without constraints under the poor condition. Particularly, as to $\Gamma = 0.001$, it is impossible to meet rate demand for the DC-SVD under given parameters when $\phi \geq 130$, which is referred to as outage event.

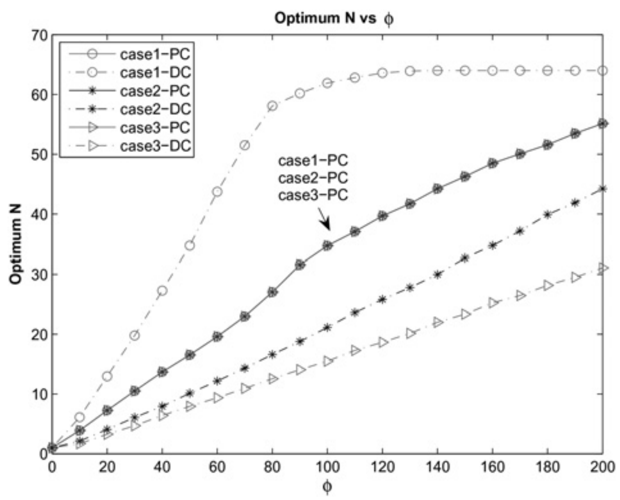

Fig. 6 *Example 3: Relationship between Optimum* N *and* ϕ *for proposed scheme*

In Fig. 6, we can find that as the condition becomes poor (small Γ and large ϕ), the optimum N increases. This is because under the poor condition, power optimisation is dominated while channel resource optimisation becomes critical under the good condition. Furthermore, the optimum N converges to the same regardless of Γ for the PC-SVD, as the reason has been discussed above.

5.3 Performance gap between the proposed resource optimisation scheme and the optimal one using the discrete particle swarm optimisation (DPSO) algorithm

As to the power optimisation and channel allocation problem discussed in Section 4.3, the optimal solution of ***P5*** can only be obtained through an exhaustive search over all feasible regions due to its mixed integer programming. Therefore, a sub-optimal algorithm is proposed to solve ***P5*** with significant reduction of computation in this paper. Then, it is necessary to make a comparison between the proposed sub-optimal algorithm and the optimal one. In this subsection, the DPSO [29] is applied to provide an approximate upper bound instead of the brutal search method to make the amount of computation acceptable. The DPSO employs a simple mechanism that mimics swarm behaviour in birds flocking to search for globally optimal solutions. Fig. 7 presents the performance gap between the proposed resource optimisation algorithm and the DPSO algorithm. Note that as to the DPSO, we consider both the PC-SVD and the DC-SVD to address the interference power constraints. We set $\Gamma = 0.1$, $\rho = 1/10$^(SNR/10), $K = 3$, and $M_{k,\mathrm{p}} = 1$, $k = 1, \ldots, 3$. To be clear, 'case 1' and 'case 2' in Fig. 7 denote $\phi = 90$ and $\phi = 180$, respectively. It is not difficult to observe that the performance of the proposed sub-optimal resource optimisation algorithm is very close to that of the DPSO for any given case. Furthermore, simulation processes show that the CPU time the DPSO spends is 50 times more than that of the proposed algorithm. Then, we can conclude that the proposed sub-optimal algorithm achieves the acceptable performance with great reduction of computation, which embodies the significance of this work. Finally, the consumed resource (BPP) increases as the condition becomes poor (small SNR and large ϕ).

5.4 Performance comparison between the proposed scheme and waterfilling scheme

In this subsection, we prepare to make a comparison between our proposed scheme and waterfilling scheme as shown in Fig. 8. Note

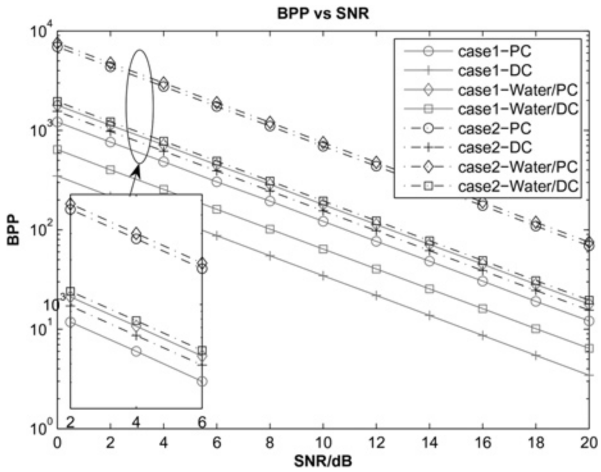

Fig. 8 *Example 5: Relationship between BPP and SNR for the proposed scheme and waterfilling scheme under different* ϕ

that as to waterfilling scheme, both the PC-SVD and the DC-SVD have been taken into account to deal with the interference power constraints. We set $\Gamma = 0.1$, $\rho = 1/10$^(SNR/10), $K = 3$, and $M_{k,\mathrm{p}} = 1$, $k = 1, \ldots, 3$. Also the solid and dash-dotted lines represent $\phi = 90$ and $\phi = 180$, respectively. It is easy to find that the proposed scheme provides a better performance than waterfilling scheme for any given case, which embodies the significance of this work. Furthermore, we see that as the condition becomes poor (small SNR and large ϕ), consumed resource (BPP) increases for both two schemes.

5.5 Performance evaluation for different number of total PUs' receive antennas

Fig. 9 compares the performance of PC-SVD with DC-SVD as the total number of primary receivers' antennas increases from 1 to 8. Without loss of generality, we assume single-antenna for each primary receiver and range K from 1 to 8. Also we set $\rho = 1$ and $\phi = 80$. The solid and dash-dotted lines correspond to $\Gamma = 0.001$ and $\Gamma = 0.1$, respectively. Observations from Fig. 9 are threefold. Firstly, as to the PC-SVD, an outage event happens when $K \geq 4$, validating that PC-SVD is practicable if and only if the number of transmit antennas is larger than that of total PUs' receive antennas. Emphasise again that the PC-SVD converges to the same performance regardless of Γ. Secondly, as mentioned above, the PC-SVD is superior to the DC-SVD under the poor condition and inferior to the DC-SVD under the good condition. Especially, note that as to the $\Gamma = 0.001$, outage event occurs for the DC-SVD when $K \geq 5$. Thirdly, for any given case, the proposed scheme provides a better performance than waterfilling scheme.

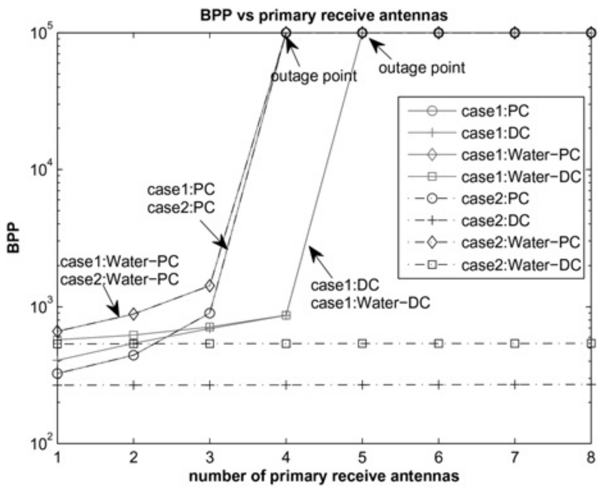

Fig. 9 *Example 6: Relationship between BPP and total PUs' receive antennas for the proposed scheme*

6 Conclusion

In this paper, we have developed a novel resource allocation optimisation framework for a MIMO-OFDMA CRN. We aim to minimise spectral footprint of the CRN using the BPP metric. The optimal source precoding matrix structure has been designed and mathematically proved. Meanwhile, two methods, namely the PC-SVD and DC-SVD, have been introduced to satisfy interference power constraints for the multiple primary receivers. In addition, a simple and a tractable algorithm with low complexity has been proposed to deal with the unified power and channel optimisation problem. Note that as to determining the channel resource budget, a simple quality evaluation criterion has been developed. Furthermore, the mean of numerical results presented in this work is twofold. Firstly, the proposed sub-optimal algorithm achieves the acceptable performance with significant reduction of computation. Secondly, the proposed framework greatly improves spectrum efficiency by striking an optimal balance between the consumed bandwidth and power. Further spectral opportunities are available compared to what can be obtained by waterfilling scheme in CRNs using overlay spectrum sharing mechanism. Hence, the framework proposed in this article offers an efficient way for resource allocation for the MIMO-OFDMA system in CRNs.

7 Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities, and the 5th Generation Mobile Communication Program in China (863 Project) under Grant number 2012AA01A507, 2014AA01A707.

8 References

1 Federal Communications Commission. 'Spectrum policy task force report'. FCC Report, ET Docket 02-135, November 2003
2 Haykin, S.: 'Cognitive radio: brain-empowered wireless communication', *IEEE J. Sel. Areas Commun.*, 2005, **23**, (2), pp. 201–220
3 Bansal, G., Hossian, M.J., Bhargava, V.K.: 'Optimal and suboptimal power allocation schemes for OFDM-based cognitive radio systems', *IEEE Trans. Wirel. Commun.*, 2008, **7**, (11), pp. 4710–4718
4 Sadr, S., Anpalagan, A., Raahemifar, K.: 'Radio resource allocation alogrithms for the downlink of multiuser OFDM communication system', *IEEE Comm. Survey Tutor.*, 2009, **11**, (3), pp. 92–106, Quarter
5 Wong, C.Y., Chen, R.S., Letaief, K.B., Murch, R.D.: 'Multiuser OFDM with adaptive subcarrier, bit and power allocation', *IEEE J. Select. Areas Commun.*, 1999, **17**, (10), pp. 1747–1758
6 Ngo, D., Tellambura, C., Nguyen, H.: 'Resource allocation for OFDMA-based cognitive radio multicast networks with primary user activity consideration', *IEEE Tran. Veh. Technol.*, 2010, **59**, (4), pp. 1668–1679
7 Wang, P., Zhao, M., Xiao, L., Zhou, S., Wang, J.: 'Power allocation in OFDM-based cognitive radio systems'. Proc. IEEE GLOBECOM, November 2007, pp. 4061–4065
8 Ma, Y., Kim, D.I., Wu, Z.: 'Optimization of OFDMA-based cellular cognitive radio networks', *IEEE Trans. Commun.*, 2010, **58**, (8), pp. 2265–2276
9 Hoang, A.T., Liang, Y.C.: 'Power control and channel allocation in cognitive radio networks with primary users' cooperation', *IEEE Trans. Mob. Comput.*, 2010, **9**, (3), pp. 348–360
10 Telatar, I.E.: 'Capacity of multi-antenna Gaussian channels', *Eur. Trans. Telecommun.*, 1999, **10**, (6), pp. 585–595
11 Tarokh, V., Seshadri, N., Calderbank, A.R.: 'Space-time block codes for high data rate wireless communication:perform criterion and code construction', *IEEE Trans. Inf. Theory*, 1998, **44**, (2), pp. 744–765
12 Farrokhi, F.R., Tassiulas, L., Liu, K.: 'Joint optimal power control and beamforming in wireless networks using antenna arrays', *IEEE Trans. Commun.*, 1998, **46**, pp. 1313–1323
13 Zhang, R., Liang, Y.C.: 'Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks', *IEEE J. Select. SP*, 2008, **2**, (1)

14 Palomar, D.P., Jiang, Y.: 'MIMO transceiver design via majorization theory', *Found. Trends Commun. Inf. Theory*, 2006, **3**, (4–5), pp. 331–551
15 Ayach, O.E., Peters, S.W., Heath, R.W.: 'The practical challenges of interference alignment', *IEEE Wirel. Commun. Mag.*, 2013, **20**, (1), pp. 35–42
16 Shin, W., Noh, W., Jang, K., *et al.*: 'Hierarchical interference alignment for downlink heterogeneous networks', *IEEE Trans. Wirel. Commun.*, 2012, **11**, (12), pp. 4549–4559
17 Razaviyayn, M., Lyubeznik, G., Luo, Z.: 'On the degrees of freedom achievable through interference alignment in a MIMO interference channel', *IEEE Trans. Signal Process.*, 2012, **60**, (2), pp. 812–821
18 Xin, Q., Kang, C.G.: 'On feasibility of interference alignment and cancellation for Gaussian interference multiple access channel'. IEEE Int. Conf. on Communications, 2012, pp. 2467–2471
19 Li, Z., Shen, B., Li, J.: 'Interference alignment and cancellation based concurrent transmission and scheduling scheme for multiuser CR-MIMO system', *IEEE Wirel. Commun. Mag.*, 2013, **10**, (8), pp. 36–43
20 Gupta, P., Kumar, P.: 'The capacity of wireless networks', *IEEE Trans. Inf. Theory*, 2000, **46**, (2), pp. 388–404
21 Hou, Y., Shi, Y., Sherali, H.: 'Optimal spectrum sharing for multi-hop software defined radio networks'. Proc. IEEE INFOCOM., May 2007, pp. 1–9
22 Tachwali, Y., Lo, B.F., Akyildiz, I.F., Agusti, R.: 'Multiuser resource allocation optimization using bandwidth-power product in cognitive radio networks', *IEEE J. Select. Areas Commun.*, 2013, **31**, (3), pp. 451–463
23 Kay, S.M.: 'Fundamentals of statistical signal process: estimation theory' (Prentice-Hall, Englewood Cliff, NJ, 1993)
24 Rong, Y., Tang, X.J., Hua, Y.B.: 'A unified framework for optimizing linear nonregenerative multicarrier mimo relay communication systems', *IEEE Trans. SP*, 2009, **57**, (12), pp. 4837–4851
25 Boyd, S., Vandenberghe, L.: 'Joint convex optimization' (Cambridge University Press, 2004)
26 Bertsekas, D.P.: 'Nonlinear programming' (Athena Scientific, 1999, 2nd edn.)
27 Marshall, A.W., Olkin, I.: 'Inequalities: theory of majorization and its application' (Academic Press, New York, 1979)
28 Hinostroza, V.: 'Frequency selectivity parameters for broadband signals'. ITST, June 2007, pp. 1–6
29 Kennedy, J., Eberhart, R.C.: 'A discrete binary version of the particle swarm algorithm', *IEEE Inter. Conf. Comput. Cybern. Simul.*, 1997, **5**, pp. 4104–4108, Oct. .SS
30 Palomar, D.P., Cioffi, J.M., Lagunas, M.A.: 'Joint tx-rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization', *IEEE Trans. Signal Process.*, 2003, **51**, pp. 2381–2401

9 Appendix

9.1 Appendix 1: Proof of Theorem 1

Proof of Theorem 1 consists of three stages. Firstly, the dual problem of ***P2*** can be written as (***P3***):

$$\max_{\boldsymbol{U}_n} q\left(\boldsymbol{d}\left(\boldsymbol{E}_n\right)\right), \tag{44}$$

$$\text{s.t. } \operatorname{tr}\left(\boldsymbol{U}_n\boldsymbol{U}_n^{\mathrm{H}}\right) \le P_n. \tag{45}$$

Based on the duality theory [25], ***P2*** and ***P3*** have the same optimal structure for $\boldsymbol{U}_n$. Furthermore, ***P3*** can easily be converted to ***P4*** stated below

$$\min_{\boldsymbol{U}_n} g\left(\boldsymbol{d}\left(\boldsymbol{E}_n\right)\right), \tag{46}$$

$$\text{s.t.}(45),$$

where $g(\boldsymbol{x}) = \sum_{j=1}^{J_n} B\log\left(x_j\right)$. In [30], the author proved that the function g in (46) is Schur-concave. Also it is not difficult to find that g is increasing with $\boldsymbol{d}[\boldsymbol{E}_n]$.

Secondly, based on [27], a few lemmas represented by the majorisation theory are introduced before proceeding to the proof of Theorem 1.

Lemma 1: [18, 9.B.1]: For a Hermitian matrix $\boldsymbol{X}$, $\boldsymbol{d}[\boldsymbol{X}]$ denotes the vector containing diagonal elements, while $\boldsymbol{\lambda}[\boldsymbol{X}]$ denotes the vector containing eigenvalues. There is $\boldsymbol{d}[\boldsymbol{X}] \prec \boldsymbol{\lambda}[\boldsymbol{X}]$.

Lemma 2: [18, 3.A.8]: If a real-valued function f is increasing and Schur-convex, then f satisfies $\boldsymbol{x} \prec \boldsymbol{y} \Rightarrow f(\boldsymbol{x}) \le f(\boldsymbol{y})$.

Lemma 3: [18, 9.H.1.h]: If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are both $N \times N$ positive semi-definite Hermitian matric with eigenvalues $\lambda_{X,i}$ and $\lambda_{Y,i}$, $i = 1, 2 \dots N$, sorted in the same order, then there is $\operatorname{tr}(\boldsymbol{XY}) \ge \sum_{i=1}^{N} \lambda_{X,i}\lambda_{Y,N+1-i}$.

Thirdly, we are ready to confirm that (18) is the optimal source precoding matrix for the problem ***P4***, which is equal to prove Theorem 1. Note that a similar proof was done in [24]. Let us define

$$\boldsymbol{X}_n = \boldsymbol{H}_n\boldsymbol{U}_n\boldsymbol{U}_n^{\mathrm{H}}\boldsymbol{H}_n^{\mathrm{H}} = \boldsymbol{\Omega}_{\boldsymbol{X}_n}\boldsymbol{\Lambda}_{\boldsymbol{X}_n}\left(\boldsymbol{\Omega}_{\boldsymbol{X}_n}\right)^{\mathrm{H}}, \tag{47}$$

where the $M_{b,n} \times M_{b,n}$ matrix $\boldsymbol{\Lambda}_{\boldsymbol{X}_n}$ is a diagonal matrix whose diagonal elements are the non-zero eigenvalues of $\boldsymbol{X}_n$ and sorted in the decreasing order, and $\boldsymbol{\Omega}_{\boldsymbol{X}_n}$ is a $M_{\mathrm{r,s}} \times M_{b,n}$ matrix containing associated eigenvectors. From (47), we can derive

$$\boldsymbol{H}_n\boldsymbol{U}_n = \boldsymbol{\Omega}_{\boldsymbol{X}_n}\left(\boldsymbol{\Lambda}_{\boldsymbol{X}_n}\right)^{1/2}\boldsymbol{Q}_n, \tag{48}$$

where the $M_{b,n} \times M_{b,n}$ matrix $\boldsymbol{Q}_n$ is an arbitrary unitary matrix satisfying $\boldsymbol{Q}_n\left(\boldsymbol{Q}_n\right)^{\mathrm{H}} = \boldsymbol{I}_{M_{b,n}}$. According to (6), we can obtain

$$\begin{aligned}
\boldsymbol{E}_n &\overset{A}{=} \boldsymbol{I}_{M_{b,n}} - \left(\boldsymbol{H}_n\boldsymbol{U}_n\right)^{\mathrm{H}}\left[\boldsymbol{H}_n\boldsymbol{U}_n\left(\boldsymbol{H}_n\boldsymbol{U}_n\right)^{\mathrm{H}} + \rho\right]^{-1}\left(\boldsymbol{H}_n\boldsymbol{U}_n\right) \\
&\overset{B}{=} \boldsymbol{I}_{M_{b,n}} - \boldsymbol{Q}_n^{\mathrm{H}}\left(\boldsymbol{\Lambda}_{\boldsymbol{X}_n}\right)^{1/2}\boldsymbol{\Omega}_{\boldsymbol{X}_n}^{\mathrm{H}}\left(\boldsymbol{\Omega}_{\boldsymbol{X}_n}\boldsymbol{\Lambda}_{\boldsymbol{X}_n}\boldsymbol{\Omega}_{\boldsymbol{X}_n}^{\mathrm{H}} + \rho\right)^{-1} \\
&\quad \times \boldsymbol{\Omega}_{\boldsymbol{X}_n}\left(\boldsymbol{\Lambda}_{\boldsymbol{X}_n}\right)^{1/2}\boldsymbol{Q}_n \\
&\overset{C}{=} \boldsymbol{I}_{M_{b,n}} - \boldsymbol{Q}_n^{\mathrm{H}}\boldsymbol{\Lambda}_{\boldsymbol{X}_n}\left(\boldsymbol{\Lambda}_{\boldsymbol{X}_n} + \rho\boldsymbol{I}_{M_{b,n}}\right)^{-1}\boldsymbol{Q}_n \\
&\overset{\Delta}{=} \boldsymbol{I}_{M_{b,n}} - \boldsymbol{M}_{E_n},
\end{aligned} \tag{49}$$

where $\boldsymbol{M}_{E_n} = \boldsymbol{Q}_n^{\mathrm{H}}\boldsymbol{\Lambda}_{\boldsymbol{X}_n}\left(\boldsymbol{\Lambda}_{\boldsymbol{X}_n} + \rho\boldsymbol{I}_{M_{b,n}}\right)^{-1}\boldsymbol{Q}_n$. Derivation A uses the lemma of matrix inversion: $(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} + \boldsymbol{C})^{-1}\boldsymbol{A} = \boldsymbol{C}^{-1}\boldsymbol{A}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{C}^{-1}\boldsymbol{A} + \boldsymbol{I})^{-1}$. By substituting (48) into the equation, expression $\boldsymbol{B}$ can be derived. Besides, derivation $\boldsymbol{C}$ exploits the lemma of matrix inversion: $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, and $\left(\boldsymbol{\Omega}_{\boldsymbol{X}_n}\right)^{-1} = \boldsymbol{\Omega}_{\boldsymbol{X}_n}^{\mathrm{H}}$.

Based on Lemma 1, we can have

$$\boldsymbol{d}\left[\boldsymbol{M}_{E_n}\right] \prec \boldsymbol{\lambda}\left[\boldsymbol{M}_{E_n}\right] = \boldsymbol{d}\left[\boldsymbol{\Lambda}_{\boldsymbol{X}_n}\left(\boldsymbol{\Lambda}_{\boldsymbol{X}_n} + \rho\boldsymbol{I}_{M_{b,n}}\right)^{-1}\right], \tag{50}$$

where $\boldsymbol{d}\left[\boldsymbol{M}_{E_n}\right]$ is majorised if $\boldsymbol{Q}_n$ is a diagonal matrix whose diagonal elements are unit norm. Then $\boldsymbol{Q}_n$ can be expressed as

$$\left|\left[\boldsymbol{Q}_n\right]_{i,i}\right| = 1, \quad \left|\left[\boldsymbol{Q}_n\right]_{i,j}\right| = 0, \quad i, j = 1, 2, \dots, M_{b,n}, \quad i \ne j. \tag{51}$$

For the convenience, we set $\boldsymbol{Q}_n = \boldsymbol{I}_{M_{b,n}}$.

As mentioned above, $g(\boldsymbol{d}(\boldsymbol{E}_n))$ is Schur-concave and increasing with respect to $\boldsymbol{d}(\boldsymbol{E}_n)$, then $g\left(\boldsymbol{d}\left(\boldsymbol{I}_{M_{b,n}} - \boldsymbol{M}_{E_n}\right)\right)$ is Schur-concave and decreasing with respect to $\boldsymbol{d}\left(\boldsymbol{M}_{E_n}\right)$. It is easy to see $-g\left(\boldsymbol{d}\left(\boldsymbol{I}_{M_{b,n}} - \boldsymbol{M}_{E_n}\right)\right)$ is Schur-convex and increasing with respect to $\boldsymbol{d}\left(\boldsymbol{M}_{E_n}\right)$. Applying Lemma 2 to (50), we obtain $-g\left(\boldsymbol{d}\left(\boldsymbol{I}_{M_{b,n}} - \boldsymbol{M}_{E_n}\right)\right) \le -g\left(\boldsymbol{d}\left(\boldsymbol{I}_{M_{b,n}} - \boldsymbol{\Lambda}_{X_n}\left(\boldsymbol{\Lambda}_{X_n} + \rho\boldsymbol{I}_{M_{b,n}}\right)^{-1}\right)\right)$, which can be equivalently expressed as $g\left(\boldsymbol{d}\left(\boldsymbol{I}_{M_{b,n}} - \boldsymbol{\Lambda}_{X_n}\left(\boldsymbol{\Lambda}_{X_n} + \rho\boldsymbol{I}_{M_{b,n}}\right)^{-1}\right)\right) \le g\left(\boldsymbol{d}\left(\boldsymbol{I}_{M_{b,n}} - \boldsymbol{M}_{E_n}\right)\right)$.

Substituting (17) into (48), we have

$$\begin{bmatrix} \mathbf{0}^{(M_{\rm r,s}-M_{b,n})\times(M_{\rm t,s}-M_{b,n})} & \mathbf{0}^{(M_{\rm r,s}-M_{b,n})\times M_{b,n}} \\ \mathbf{0}^{M_{b,n}\times(M_{\rm t,s}-M_{b,n})} & \left(\mathbf{\Lambda}_{H_n,1}^{M_{b,n}\times M_{b,n}}\right)^{1/2} \end{bmatrix} \boldsymbol{V}_{H_n}^{\rm H}\boldsymbol{U}_n = \mathbf{\Omega}_{H_n}^{\rm H}\mathbf{\Omega}_{X_n}\left(\mathbf{\Lambda}_{X_n}\right)^{1/2}, \quad (52)$$

where $\mathbf{\Lambda}_{H_n,1}^{M_{b,n}\times M_{b,n}}$ is a $M_{b,n}\times M_{b,n}$ diagonal matrix containing all non-zero singular values of $\boldsymbol{H}_n$.

Referring to [24], (52) can be finally derived as

$$\boldsymbol{V}_{H_n,1}^{\rm H}\boldsymbol{U}_n = \left(\mathbf{\Lambda}_{H_n,1}\right)^{-1/2}\mathbf{\Omega}_{H_n,1}^{\rm H}\mathbf{\Omega}_{X_n}\left(\mathbf{\Lambda}_{X_n}\right)^{1/2}, \quad (53)$$

where the matrix $\boldsymbol{V}_{H_n,1}$ and $\mathbf{\Omega}_{H_n,1}$ contain the rightmost columns corresponding to non-zero singular values from $\boldsymbol{V}_{H_n}$ and $\mathbf{\Omega}_{H_n}$, respectively. Then the transmit power at the nth channel is given by

$$\begin{aligned} \operatorname{tr}\left(\boldsymbol{U}_n\boldsymbol{U}_n^{\rm H}\right) &= \operatorname{tr}\left(\boldsymbol{V}_{H_n,1}^{\rm H}\boldsymbol{U}_n\left(\boldsymbol{V}_{H_n,1}^{\rm H}\boldsymbol{U}_n\right)^{\rm H}\right) \\ &= \operatorname{tr}\Big(\left(\mathbf{\Lambda}_{H_n,1}\right)^{-1/2}\mathbf{\Omega}_{H_n,1}^{\rm H}\mathbf{\Omega}_{X_n}\mathbf{\Lambda}_{X_n} \quad (54) \\ &\quad\times\mathbf{\Omega}_{X_n}^{\rm H}\mathbf{\Omega}_{H_n,1}\left(\mathbf{\Lambda}_{H_n,1}\right)^{-1/2}\Big). \end{aligned}$$

At last, (54) is minimised in order to optimise the power resource. From Lemma 3, it can be found that (54) is minimised if $\mathbf{\Omega}_{X_n} = \mathbf{\Omega}_{H_n,1}$. Substituting the result into (53), we can obtain $\boldsymbol{U}_n = \boldsymbol{V}_{H_n,1}\left(\mathbf{\Lambda}_{H_n,1}\right)^{-1/2}\left(\mathbf{\Lambda}_{X_n}\right)^{1/2} = \boldsymbol{V}_{H_n,1}\mathbf{\Lambda}_{U_n}^{1/2}$, where $\mathbf{\Lambda}_{U_n} = \left(\mathbf{\Lambda}_{H_n,1}\right)^{-1}\mathbf{\Lambda}_{X_n}$. The proof of Theorem 1 is completely finished.

9.2 Appendix 2: Proof of Theorem 2

Note that

$$\boldsymbol{H}_{\perp,n}\boldsymbol{G}_n^{\rm H} = \mathbf{0}. \quad (55)$$

Substituting (21) into (55), we can derive

$$\left(\boldsymbol{V}_{H_{\perp,n},1}\right)^{\rm H}\boldsymbol{G}_n^{\rm H} = \mathbf{0}. \quad (56)$$

Then, the interferences introduced to primary receivers can be obtained as

$$\operatorname{tr}\left(\boldsymbol{G}_n\boldsymbol{U}_n\boldsymbol{U}_n^{\rm H}\boldsymbol{G}_n^{\rm H}\right) = \operatorname{tr}\left(\boldsymbol{G}_n\boldsymbol{V}_{H_{\perp,n},1}\mathbf{\Lambda}_{U_n}\boldsymbol{V}_{H_{\perp,n},1}^{\rm H}\boldsymbol{G}_n^{\rm H}\right) = \mathbf{0}. \quad (57)$$

Theorem 2 is completely proved.

9.3 Appendix 3: Proof of Remark 1

Here, we define S^* as the total number of subspaces that are occupied for the secondary transmission, namely the power allocated in the $\{\tilde{\lambda}_m, m \le S^*\}$ is non-zero while that in the $\{\tilde{\lambda}_m, m > S^*\}$ becomes zero. Then, based on (37) and (39), the Lagrange multiplier ν can be solved as

$$\nu = \frac{\rho 2^{(\phi/BS^*)}}{\left(\prod_{j=1}^{S^*}\tilde{\lambda}_j\right)^{1/S^*}}, \quad (58)$$

where ν satisfies

$$\nu - \frac{1}{\tilde{\lambda}_{S^*}/\rho} \ge 0, \quad \text{and} \quad \nu - \frac{1}{\tilde{\lambda}_{S^*+1}/\rho} \le 0. \quad (59)$$

Note that S^* can be iteratively obtained by the bisection method. Then, P_n, the power allocated on the nth channel, can be derived as

$$\begin{aligned} P_n &= \sum_{j=1}^{M_{b,n}}\left[\frac{\rho 2^{(\phi/BS^*)}}{\left(\prod_{j=1}^{S^*}\tilde{\lambda}_j\right)^{1/S^*}} - \frac{1}{\bar{\lambda}_{j,n}/\rho}\right]^+ \\ &\overset{A}{=} \frac{\rho M_{b,n}2^{(\phi/BS^*)}}{\left(\prod_{\tilde{\lambda}_j\in\bar{\mathcal{C}}_n}\tilde{\lambda}_j\right)^{1/S^*}} \times \frac{1}{\left(\prod_{\tilde{\lambda}_j\in\mathcal{C}_n}\tilde{\lambda}_j\right)^{1/S^*}} - \sum_{j=1}^{M_{b,n}}\frac{\rho}{\bar{\lambda}_{j,n}} \\ &\le \frac{\rho M_{b,n}2^{(\phi/BS^*)}}{\left[\left(\prod_{\tilde{\lambda}_j\in\bar{\mathcal{C}}_n}\tilde{\lambda}_j\right)\left(\prod_{\tilde{\lambda}_j\in\mathcal{C}_n}\tilde{\lambda}_j\right)\right]^{1/S^*}} - \frac{\rho M_{b,n}}{\left(\prod_{\tilde{\lambda}_j\in\mathcal{C}_n}\tilde{\lambda}_j\right)^{1/M_{b,n}}} \quad (60) \\ &= \rho M_{b,n}(Q_n)^{1/S^*}\left[\frac{2^{\frac{\phi}{BS^*}}}{\left(\prod_{\tilde{\lambda}_j\in\bar{\mathcal{C}}_n}\tilde{\lambda}_j\right)^{1/S^*}} - Q_n^{(1/M_{b,n}-1/S^*)}\right], \end{aligned}$$

where $\bar{\mathcal{C}}_n = \{\bar{\lambda}_{j,k}, j = 1, \ldots, M_{b,n}, k \in \mathcal{C}\ \&\&\ k \ne n\}$. For the derivation A, we assume that the power allocated over all the subspaces on the n channel ($n \in \mathcal{C}$) is non-negative, namely $\nu - \left(1/\left[\bar{\lambda}_{j,n}/\rho\right]\right) \ge 0, j = 1, \ldots, M_{b,n}$. Note that designing the optimal quality criterion becomes impossible due to the difficulty of deriving the close-form of S^*. However, according to (60), P_n is greatly dominated by $(Q_n)^{1/S^*}$, and is increasing with Q_n. Then, we infer that Q_n is an outstanding and simple evaluation criterion and the smaller Q_n is, the better quality of the channel is. The proof is completed.