Joint Source and Relay Design for Multiuser MIMO Nonregenerative Relay Networks With Direct Links

Haibin Wan and Wen Chen, *Senior Member, IEEE*

Abstract—In this paper, we investigate joint source precoding matrices and relay processing matrix design for multiuser multiple-input–multiple-output (MIMO) nonregenerative relay networks in the presence of the direct source–destination links. We consider both capacity and mean-square-error (MSE) criteria subject to the distributed power constraints, which are nonconvex and apparently have no simple solutions. Therefore, we propose an optimal source precoding matrix structure based on the point-to-point MIMO channel technique and a new relay processing matrix structure under the modified power constraint at relay node, based on which a nested iterative algorithm of jointly optimizing sources precoding and relay processing is established. We show that the capacity-based optimal source precoding matrices share the same structure with the MSE-based matrices and so does the optimal relay processing matrix. Simulation results demonstrate that the proposed algorithm outperforms the existing results.

Index Terms—Direct link, multiuser multiple-input multiple-output (MU-MIMO), nonregenerative relay, precoding matrix.

I. INTRODUCTION

Recently, multiple-input–multiple-output (MIMO) relay networks have attracted considerable interest from both academic and industrial communities. It has been verified that wireless relay can increase the coverage and capacity of wireless networks [1]. Meanwhile, MIMO techniques can provide significant improvement for the spectral efficiency and link reliability in scattered environments because of their multiplexing and diversity gains [2]. A MIMO relay network, combining the relaying and MIMO techniques, can make use of both advantages to increase the data rate in the network edge and extend the network coverage. It is a promising technique for next-generation wireless communications.

The capacity of MIMO relay network has extensively been investigated in the literature [3]–[7]. Recent works on MIMO nonregenerative relay have focused on how the source precoding and relay processing matrices are designed. For a single-user MIMO relay network, an optimal relay processing matrix that maximizes the end-to-end mutual information is independently designed in [8] and [9], and the optimal structures of jointly designed source precoding matrix and relay processing matrix are derived in [10]. In [11] and [12], the relay processing matrix for minimizing the mean square error (MSE) at the destination is developed. A unified framework for jointly optimizing

the source precoding and the relay processing matrices is established in [13]. For multiuser single-antenna relay networks, the optimal relay processing is designed to maximize the system capacity [14]–[16]. In [17], the optimal source precoding matrices and a relay processing matrix are developed in the downlink and uplink scenarios of an MU-MIMO relay network without considering source–destination (S–D) links. Only a few works consider the direct S–D links. In [18] and [19], the optimal relay processing matrix is designed based on the MSE criterion with and without the optimal source precoding matrix in the presence of direct links, respectively. However, for a relay network with direct S–D links, jointly optimizing the source precoding and the relay processing matrices based on capacity or MSE is much difficult, particularly for an MU-MIMO relay network.

In this paper, we consider an MU-MIMO nonregenerative relay network where each node is equipped with multiple antennas. We take the effect of the S–D link into the joint optimization of the source precoding matrices and the relay processing matrix, which is more complicated than the relatively simple case without considering S–D links [17]. To our best knowledge, there is no such work in the literature on the joint optimization of source precoding and relay processing for MU-MIMO nonregenerative relay networks with direct S–D links. Two major contributions of this paper over the conventional works are listed as follows.

- We first introduce a general strategy to the joint design of source precoding matrices and relay processing matrix by transforming the network into a set of parallel scalar subsystems as a point-to-point MIMO channel under a relay-modified power constraint, and we show that the capacity-based source precoding matrices and relay processing matrix, respectively, share the same structures with the MSE-based matrices.
- A nested iterative algorithm is presented to solve the joint optimization of source precoding and relay processing based on capacity and MSE, respectively. Simulation results show that the proposed algorithm outperforms the existing methods.

The rest of this paper is organized as follows. Section II illustrates the system model. Section III presents the optimal structures of source precoding and relay processing, and a nested iterative algorithm to solve the joint optimization of sources precoding and relay processing. Section IV is devoted to the simulation results. Finally, Section V concludes this paper.

Notations: Lowercase, boldface lowercase, and boldface uppercase letters denote the scalar, vector, and matrix, respectively. $\mathsf{E}(\cdot)$, $\mathrm{tr}(\cdot)$, $(\cdot)^{-1}$, $(\cdot)^{\dagger}$, $|\cdot|$, and $\|\cdot\|_F$ denote the expectation, trace, inverse, conjugate transpose, determinant, and Frobenius norm of a matrix, respectively. $\mathbf{I}_N$ stands for the identity matrix of order N. $\mathrm{diag}(a_1, \ldots, a_N)$ is a diagonal matrix with the ith diagonal entry a_i. log is of base 2. $\mathcal{C}^{M \times N}$ represents the set of $M \times N$ matrices over a complex field, and $\sim \mathcal{CN}(x, y)$ means satisfying a circularly symmetric complex Gaussian distribution with mean x and covariance y. $[x]^+$ denotes $\max\{0, x\}$.

H. Wan is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China, and also with the School of Physics Science and Technology, Guangxi University, Nanning 530004, China (e-mail: dahai_good@sjtu.edu.cn).

W. Chen is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China, and also with the State Key Laboratory for Integrated Service Networks, Xidian University, Xi'an 710071, China (e-mail: wenchen@sjtu.edu.cn).

II. SYSTEM MODEL

We consider a multiple-access MIMO relay network with two source nodes (SNs), one relay node (RN) and one destination node (DN), as illustrated in Fig. 1, where the channel matrices have been shown. The numbers of antennas equipped at the SNs, RN, and DN are N_s, N_r and N_d, respectively. For simplicity, we assume that there are only two SNs and both SNs have the same number of antennas. However, it is easy to be generalized to the scenario of multiple SNs with different numbers of antennas at each SN. In this paper, we
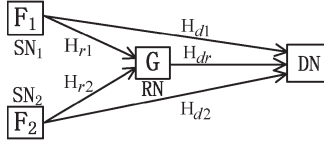
Fig. 1.   Multiple-access relay network with two source nodes, one relay node, and one destination node.

consider a nonregenerative half-duplex relaying strategy applied at the RN to process the received signals. Thus, the transmission will take place in two phases. Suppose that perfect synchronization has been established between SN_1 and SN_2 prior to transmission and both SN_1 and SN_2 simultaneously transmit their independent messages to the RN and DN during the first phase. Then, the RN processes the received signals and forwards them to the DN during the second phase.

Let $\mathbf{H}_{ri} \in \mathcal{C}^{N_r \times N_s}$, $\mathbf{H}_{di} \in \mathcal{C}^{N_d \times N_s}$, and $\mathbf{H}_{dr} \in \mathcal{C}^{N_d \times N_r}$ denote the channel matrices of the ith SN to the RN, to the DN, and the RN to the DN, respectively. Each entry of the channel matrices is assumed to be a complex Gaussian variable with zero mean and variance σ_h^2. Furthermore, all the channels involved are assumed to be quasistatic independent and identically distributed (i.i.d.) Rayleigh fading combining with large-scale fading over a common narrowband. Let $\mathbf{F}_1 \in \mathcal{C}^{N_s \times N_s}$ and $\mathbf{F}_2 \in \mathcal{C}^{N_s \times N_s}$ denote the precoding matrices for SN_1 and SN_2, respectively, which satisfy the power constraint $\mathsf{E}[\mathbf{s}_i^\dagger \mathbf{F}_i^\dagger \mathbf{F}_i \mathbf{s}_i] = \mathrm{tr}(\mathbf{F}_i^\dagger \mathbf{F}_i) \leq P_i$. Let $\mathbf{G} \in \mathcal{C}^{N_r \times N_r}$ denote the relay processing matrix. Suppose that $\mathbf{n}_r \in \mathcal{C}^{N_r \times 1}$ and $\mathbf{n}_i \in \mathcal{C}^{N_d \times 1}$ are the noise vectors at the RN and DN, respectively, and all noise are i.i.d. additive white Gaussian noise with zero mean and unit variance. Then, the baseband signal vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ received at the DN during the two consecutive phases can be expressed as follows:

$$
\underbrace{\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} \mathbf{H}_{d1} \\ \mathbf{H}_{dr}\mathbf{G}\mathbf{H}_{r1} \end{bmatrix}}_{\mathbf{H}_1} \mathbf{F}_1 \mathbf{s}_1 + \underbrace{\begin{bmatrix} \mathbf{H}_{d2} \\ \mathbf{H}_{dr}\mathbf{G}\mathbf{H}_{r2} \end{bmatrix}}_{\mathbf{H}_2} \mathbf{F}_2 \mathbf{s}_2
$$
$$
+ \underbrace{\begin{bmatrix} \mathbf{I}_{N_d} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{dr}\mathbf{G} & \mathbf{I}_{N_d} \end{bmatrix}}_{\mathbf{H}_3} \underbrace{\begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_r \\ \mathbf{n}_2 \end{bmatrix}}_{\mathbf{N}} \quad (1)
$$

where $\mathbf{s}_i \in \mathcal{C}^{N_s \times 1}$ is assumed to be a zero-mean circularly symmetric complex Gaussian signal vector that was transmitted by the ith SN and satisfies $\mathsf{E}(\mathbf{s}_i \mathbf{s}_i^\dagger) = \mathbf{I}_{N_s}$. Let $\mathbf{Y}$, $\mathbf{H}_i (i=1,2,3)$, and $\mathbf{N}$, as shown in (1), denote the effective receive signal, effective channels, and effective noise, respectively. Then, $\mathbf{H}_3\mathsf{E}[\mathbf{N}\mathbf{N}^\dagger]\mathbf{H}_3^\dagger = \mathbf{H}_3\mathbf{H}_3^\dagger = \mathrm{diag}(\mathbf{I}_{N_d}, \mathbf{R})$, where $\mathbf{R} = \mathbf{I}_{N_d} + \mathbf{H}_{dr}\mathbf{G}\mathbf{G}^\dagger\mathbf{H}_{dr}^\dagger$ is the covariance matrix of the effective noise at the DN during the second phase.

III. OPTIMAL COORDINATES OF THE JOINT SOURCE AND RELAY DESIGN

In this section, the capacity and MSE for the minimum-mean-square-error (MMSE) detector with successive interference cancelation (SIC) at the DN are analyzed. Then, we will exploit the optimal structures of source precoding and relay processing based on capacity and MSE, respectively. Then, a new algorithm for jointly optimizing the sources precoding matrices and the relay processing matrix is proposed to maximize the capacity or minimize MSE of the entire network.

A. Decoding Scheme

Conventional receivers such as the matched filter (MF), zero forcing (ZF), and MMSE decoder have been well studied in the previous works. The MF receiver has bad performance in the region with a

high signal-to-noise ratio (SNR), whereas the ZF produces a noise enhancement effect in the low-SNR region. The MMSE detector with SIC has significant advantage over MF and ZF, which is information lossless and optimal [20]. Therefore, we consider the MMSE-SIC receiver at the DN and first decode the signal from SN_2 without loss of generality. With the predetermined decoding order, the interference from SN_2 to SN_1 is virtually absent. To exploit the optimal structures of the matrices at the SNs, we first set up the RN with a fixed processing matrix $\mathbf{G}$ without considering the power control. With the predetermined decoding order, the MMSE receive filter for SN_i ($i = 1, 2$) is given as [21], [22]

$$
\mathbf{A}_i^{\mathrm{MMSE}} = \mathbf{F}_i^\dagger \mathbf{H}_i^\dagger \left(\mathbf{H}_i \mathbf{F}_i \mathbf{F}_i^\dagger \mathbf{H}_i^\dagger + \mathbf{R}_{Z_i} \right)^{-1} \quad (2)
$$

where $\mathbf{R}_{Z_1} \triangleq \mathbf{H}_3\mathbf{H}_3^\dagger$, and $\mathbf{R}_{Z_2} \triangleq \mathbf{H}_3\mathbf{H}_3^\dagger + \mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^\dagger\mathbf{H}_1^\dagger$. Then, the MSE matrix for SN_i can be expressed as

$$
\mathbf{E}_i = \mathsf{E}\left[\left(\mathbf{A}_i^{\mathrm{MMSE}}\mathbf{Y}_i - \mathbf{s}_i \right) \left(\mathbf{A}_i^{\mathrm{MMSE}}\mathbf{Y}_i - \mathbf{s}_i \right)^\dagger \right]
$$
$$
= \left(\mathbf{I}_{N_s} + \mathbf{F}_i^\dagger\mathbf{H}_i^\dagger\mathbf{R}_{Z_i}^{-1}\mathbf{H}_i\mathbf{F}_i \right)^{-1} \quad (3)
$$

where $\mathbf{Y}_1 = \mathbf{Y} - \mathbf{H}_2\mathbf{F}_2\mathbf{s}_2$, and $\mathbf{Y}_2 = \mathbf{Y}$. Hence, the capacity for SN_i is given as [20]

$$
C_i = \log \left| \mathbf{I}_{N_s} + \mathbf{F}_i^\dagger\mathbf{H}_i^\dagger\mathbf{R}_{Z_i}^{-1}\mathbf{H}_i\mathbf{F}_i \right| = \log \left| \mathbf{E}_i^{-1} \right|. \quad (4)
$$

B. Optimal Precoding Matrices at SNs

In this section, we will introduce two lemmas, which will be used to exploit the optimal source precoding matrices and relay processing matrix, respectively.

Lemma 1: For a matrix $\mathbf{A}$, if matrix $\mathbf{B}$ is a positive definite matrix and $\mathbf{C} = \mathbf{A}\mathbf{B}^{-1}\mathbf{A}^\dagger$, then $\mathbf{C}$ is an Hermitian and positive semidefinite matrix (HPSDM).

Proof: Because $\mathbf{B}$ is a positive definite matrix, $\mathbf{B}^{-1}$ is also a positive definite matrix. For any nonzero column vector $\boldsymbol{x}$, let $\boldsymbol{y} = \mathbf{A}^\dagger\boldsymbol{x}$. Then, we have $\boldsymbol{x}^\dagger\mathbf{C}\boldsymbol{x} = \boldsymbol{x}^\dagger\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^\dagger\boldsymbol{x} = \boldsymbol{y}^\dagger\mathbf{B}^{-1}\boldsymbol{y} \geq 0$, which implies that $\mathbf{C}$ is an HPSDM. ∎

Lemma 2: If $\mathbf{A}$ and $\mathbf{B}$ are positive semidefinite matrices, then, $0 \leq \mathrm{tr}(\mathbf{A}\mathbf{B}) \leq \mathrm{tr}(\mathbf{A})\mathrm{tr}(\mathbf{B})$, and there is an $\alpha \in [0, 1]$ such that $\mathrm{tr}(\mathbf{A}\mathbf{B}) = \alpha\mathrm{tr}(\mathbf{A})\mathrm{tr}(\mathbf{B})$.

Proof: See [23]. ∎

Because $\mathbf{R}_{Z_i} (i = 1, 2)$ is a positive definite matrix [24], according to Lemma 1, $\mathbf{H}_{si} = \mathbf{H}_i^\dagger\mathbf{R}_{Z_i}^{-1}\mathbf{H}_i$ is an HPSDM, which can be decomposed as

$$
\mathbf{H}_{si} = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{U}_i^\dagger \quad (5)
$$

with a unitary matrix $\mathbf{U}_i$ and nonnegative diagonal matrices $\mathbf{\Lambda}_i$, in which diagonal entries are in descending order. One of the main results of this paper is described as follows.

Proposition 1: For a given matrix[1] $\mathbf{G}$ and predetermined decoding order, the precoding matrix for SN_i with the canonical form

$$
\mathbf{F}_i = \mathbf{U}_i\mathbf{\Sigma}_i \quad (i = 1, 2) \quad (6)
$$

is optimal with the water-filling power allocation policy (Policy A) based on capacity or with the inverse water-filling power allocation

---

[1]The relay power constraint problem will directly be dealt with by an iterative algorithm later.

policy (Policy B) based on MSE, where

$$\mathbf{\Sigma}_i^2 = \left[\mu - \mathbf{\Lambda}_i^{-1}\right]^+ \qquad \text{(Policy - A)} \qquad (7a)$$

$$\mathbf{\Sigma}_i^2 = \left[\mu\mathbf{\Lambda}_i^{-1/2} - \mathbf{\Lambda}_i^{-1}\right]^+ \qquad \text{(Policy - B)} \qquad (7b)$$

$$\text{s.t} : \text{tr}\left(\mathbf{\Sigma}_i^2\right) = P_i. \qquad (7c)$$

Proof: Substituting $\mathbf{F}_1$ in (6) into (4) and (3), we, respectively, have

$$C_1 = \log\left|\mathbf{I}_{N_s} + \mathbf{\Sigma}_1^2\mathbf{\Lambda}_1\right|$$

$$\text{tr}(\mathbf{E}_1) = \text{tr}\left\{\left(\mathbf{I}_{N_s} + \mathbf{\Sigma}_1^2\mathbf{\Lambda}_1\right)^{-1}\right\}.$$

According to the Karush–Kuhn–Tucker (KKT) conditions [25], Policies A and B can make the capacity C_1 maximized and the MSE $\text{tr}(\mathbf{E}_1)$ minimized, respectively, under the power control P_1 at SN_1. This condition implies that $\mathbf{F}_1$ is optimal. After deciding on $\mathbf{F}_1$ and substituting $\mathbf{F}_1$ into $\mathbf{R}_{Z_2}$, we can prove that $\mathbf{F}_2$ is optimal. ∎

C. Nearly Optimal Processing Matrix at the Relay

In this section, we first exploit the structure of a relay processing matrix based on capacity for given $\mathbf{F}_1$ and $\mathbf{F}_2$. Then, we show that the same structure matrix at the RN can make the MSE of the entire network achieve the minimum with a different power allocation policy. The capacity of the entire network is [20]

$$C = \log\left|\mathbf{H}_1\mathbf{\Pi}_1\mathbf{H}_1^\dagger + \mathbf{H}_2\mathbf{\Pi}_2\mathbf{H}_2^\dagger + \mathbf{H}_3\mathbf{H}_3^\dagger\right| - \log\left|\mathbf{H}_3\mathbf{H}_3^\dagger\right|$$

where $\mathbf{\Pi}_i = \mathbf{F}_i\mathbf{F}_i^\dagger$. According to the determinant expansion formula of the block matrix [26], (8) can be rewritten as

$$C = \log|\mathbf{T}| + \log\left|\mathbf{H}_{dr}\mathbf{G}\mathbf{K}\mathbf{G}^\dagger\mathbf{H}_{dr}^\dagger + \mathbf{R}\right| - \log|\mathbf{R}| \qquad (8)$$

where

$$\mathbf{T} = \mathbf{I}_{N_d} + \sum_{i=1}^{2}\mathbf{H}_{di}\mathbf{\Pi}_i\mathbf{H}_{di}^\dagger \qquad (9a)$$

$$\mathbf{K} = \sum_{i=1}^{2}\mathbf{H}_{ri}\mathbf{\Pi}_i\mathbf{H}_{ri}^\dagger - \widetilde{\mathbf{K}} \qquad (9b)$$

$$\widetilde{\mathbf{K}} = \left(\sum_{i=1}^{2}\mathbf{H}_{ri}\mathbf{\Pi}_i\mathbf{H}_{di}^\dagger\right)\mathbf{T}^{-1}\left(\sum_{i=1}^{2}\mathbf{H}_{di}\mathbf{\Pi}_i\mathbf{H}_{ri}^\dagger\right). \qquad (9c)$$

Let $\Delta = \log|\mathbf{T}|$, which is independent of $\mathbf{G}$. Then, for given $\mathbf{F}_1$ and $\mathbf{F}_2$, the problem on the maximum capacity of the network can be formulated as

$$\arg\max_{\mathbf{G}} \tilde{C} = \log\left|\mathbf{H}_{dr}\mathbf{G}\mathbf{K}\mathbf{G}^\dagger\mathbf{H}_{dr}^\dagger + \mathbf{R}\right| - \log|\mathbf{R}| \qquad (10a)$$

$$\text{s.t. } \text{tr}\left\{\mathbf{G}\left(\mathbf{I}_{N_r} + \sum_{i=1}^{2}\mathbf{H}_{ri}\mathbf{\Pi}_i\mathbf{H}_{ri}^\dagger\right)\mathbf{G}^\dagger\right\} \le P_r. \qquad (10b)$$

To solve this problem and find a nearly optimal processing matrix $\mathbf{G}$, due to $\mathbf{K} = \mathbf{K}^\dagger$, we first decompose $\mathbf{K}$ based on eigenvalue

decomposition and then decompose $\mathbf{H}_{dr}$ based on singular value decomposition, i.e.,

$$\mathbf{K} = \mathbf{U}_K\mathbf{\Lambda}_K\mathbf{U}_K^\dagger$$

$$\mathbf{H}_{dr} = \mathbf{U}_H\mathbf{\Theta}\mathbf{V}_H^\dagger$$

where $\mathbf{U}_K$, $\mathbf{U}_H$, and $\mathbf{V}_H$ are unitary matrices, $\mathbf{\Lambda}_K = \text{diag}(\lambda_1, \cdots, \lambda_{N_r})$ is an $N_r \times N_r$ diagonal matrix, and $\mathbf{\Theta} = \text{diag}(\theta_1, \cdots, \theta_r)$ is an $N_r \times N_r$ diagonal matrix, in which diagonal entries are in descending order.

Based on (10a), it is easy to verify that the optimal left canonical of $\mathbf{G}$ is still given by $\mathbf{V}_H$ [8]. However, it is intractable to find the optimal right canonical for the processing matrix $\mathbf{G}$, because there is no matrix that can achieve the diagonalization of both the capacity cost function (10a) and the power constraint (10b). Nonetheless, we can modify the power constraint (10b) to another expression to find a matrix with the desired property. Because $\mathbf{K}$ is a deterministic matrix for the fixed sources precoding matrices, (10b) can be rewritten as

$$\text{tr}\left\{\mathbf{G}\left(\mathbf{I}_{N_r} + \mathbf{K}\right)\mathbf{G}^\dagger\right\} + \text{tr}\{\widetilde{\mathbf{K}}\mathbf{G}^\dagger\mathbf{G}\}$$

$$= \text{tr}\left\{\mathbf{G}\left(\mathbf{I}_{N_r} + \sum_{i=1}^{2}\mathbf{H}_{ri}\mathbf{\Pi}_i\mathbf{H}_{ri}^\dagger\right)\mathbf{G}^\dagger\right\} \le P_r.$$

Because $\mathbf{T}$ is a positive definite matrix, according to Lemma 1, $\widetilde{\mathbf{K}}$ in (9c) is also a positive semidefinite matrix. According to Lemma 1, the new power constraint at the RN can be expressed as

$$\text{tr}\left\{\mathbf{G}\left(\mathbf{I}_{N_r} + \mathbf{K}\right)\mathbf{G}^\dagger\right\} + \alpha\text{tr}\left\{\widetilde{\mathbf{K}}\right\}\text{tr}\{\mathbf{G}^\dagger\mathbf{G}\}$$

$$\approx \text{tr}\left\{\mathbf{G}(\mathbf{I}_{N_r} + \sum_{i=1}^{2}\mathbf{H}_{ri}\mathbf{F}_i\mathbf{F}_i^\dagger\mathbf{H}_{ri}^\dagger)\mathbf{G}^\dagger\right\} \le P_r \qquad (11)$$

where the exact value α can be found by an iterative method. Thus, applying the results in [8], [17], the processing matrix $\mathbf{G}$ with the following structure can achieve the desired diagonalization for both capacity cost function (10a) and the new power constraint (11) and will be optimal [8]:

$$\mathbf{G} = \mathbf{V}_H\mathbf{\Xi}\mathbf{U}_K^\dagger \qquad (12)$$

where $\mathbf{\Xi}^2 = \text{diag}(\xi_1, \cdots, \xi_{N_r})$ can be solved by an optimization method [8].

Let $\kappa = \text{tr}\{\widetilde{\mathbf{K}}\}$. Substituting $\mathbf{G}$ into (10a) and using the new power constraint (11) to replace (10b), the problem (10) to find ξ_i becomes

$$\arg\max_{\xi_1, \cdots, \xi_{N_r}} \tilde{C}(\xi_i) = \sum_{i=1}^{N_r}\log\frac{\theta_i^2\xi_i\lambda_i + \theta_i^2\xi_i + 1}{\theta_i^2\xi_i + 1} \qquad (13a)$$

$$\text{s.t. } \sum_{i=1}^{N_r}(\lambda_i + \alpha\kappa + 1)\xi_i \le P_r \text{ and } \xi_i \ge 0 \, \forall i. \qquad (13b)$$

Then, this optimization problem with respect to ξ_i is similar to a problem solved in [8], [17]. Then, we have

$$\xi_i = \frac{1}{2\theta_i^2(\lambda_i + 1)}\left[\sqrt{\lambda_i^2 + \frac{4\lambda_i\theta_i^2(\lambda_i + 1)\mu}{\lambda_i + 1 + \alpha\kappa}} - \lambda_i - 2\right]^+ \qquad (14)$$

$$\sum_{i=1}^{N_r}(\lambda_i + 1 + \alpha\kappa)\xi_i \le P_r \qquad (15)$$

where μ in (14) is decided by (15).

Next, we will show that the same structure matrix $\mathbf{G}$ can also make the MSE of the entire network achieve the minimum with a different

power allocation matrix $\mathbf{\Xi}$ for given $\mathbf{F}_1$ and $\mathbf{F}_2$. Due to the total MSE, it can be expressed as

$$
\begin{aligned}
J(\mathbf{G}) &= \mathrm{tr}(\mathbf{E}_1) + \mathrm{tr}(\mathbf{E}_2) \\
&\overset{a}{\leq} \mathrm{tr}(\tilde{\mathbf{E}}_1) + \mathrm{tr}(\mathbf{E}_2) \\
&= \mathrm{tr}\left\{ \left(\mathbf{I}_{2N_d} + \mathbf{F}^\dagger \mathbf{H}^\dagger \mathbf{R}_{Z_1}^{-1} \mathbf{H} \mathbf{F}\right)^{-1} \right\} \\
&\overset{b}{=} \mathrm{tr}\left(\mathbf{I}_{2N_d}\right) - \mathrm{tr}\left\{ \left(\mathbf{R}_{Z_1} + \mathbf{H}\mathbf{F}\mathbf{F}^\dagger \mathbf{H}^\dagger\right)^{-1} \mathbf{H}\mathbf{F}\mathbf{F}^\dagger \mathbf{H}^\dagger \right\} \\
&= \mathrm{tr}\left\{ \left(\mathbf{R}_{Z_1} + \mathbf{H}\mathbf{F}\mathbf{F}^\dagger \mathbf{H}^\dagger\right)^{-1} \mathbf{R}_{Z_1} \right\} \\
&\overset{c}{=} \beta\,\mathrm{tr}\left\{ \left(\mathbf{H}_{dr}\mathbf{G}\mathbf{K}\mathbf{G}^\dagger \mathbf{H}_{dr}^\dagger + \mathbf{R}\right)^{-1} \right\} \mathrm{tr}\left\{ \left(\mathbf{I}_{N_d} + \mathbf{R}\right) \right\} \\
&\triangleq \beta \tilde{J}(\mathbf{G})
\end{aligned} \tag{16}
$$

where $\mathbf{F} = \mathrm{diag}(\mathbf{F}_1, \mathbf{F}_2)$, $\mathbf{H} = [\mathbf{H}_1 \mathbf{H}_2]$, and β is a scalar factor. In (16), (a) comes from the fact that noise is enhanced using $\tilde{\mathbf{R}}_{Z_1} = \mathbf{H}_3\mathbf{H}_3^\dagger + \mathbf{H}_2\mathbf{\Pi}_2\mathbf{H}_2^\dagger$ to replace $\mathbf{R}_{Z_1}$ in calculating $\mathrm{tr}(\tilde{\mathbf{E}}_1)$, ($b$) follows from the Woodbury identity and $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, and ($c$) follows from Lemma 2 and the Schur complement to inverse a block matrix [26]. Based on (16), minimizing $J(\mathbf{G})$ is equivalent to minimizing $\tilde{J}(\mathbf{G})$. Then, for given $\mathbf{F}_1$ and $\mathbf{F}_2$, the optimal $\mathbf{G}$ to minimize MSE is

$$
\arg\min_{\mathbf{G}} \tilde{J}(\mathbf{G}) \tag{17a}
$$
$$
\text{s.t. : } (11). \tag{17b}
$$

Based on the aforementioned analysis, the structure of $\mathbf{G}$ in (12) can also achieve the diagonalization of (17) but has a new power allocation matrix $\mathbf{\Xi}$ that is different from the capacity-based matrix. Then, substituting $\mathbf{G}$ in (12) into (17) to find the new $\mathbf{\Xi}$, (17) becomes

$$
\arg\min_{\xi_1,\dots,\xi_{N_r}} \tilde{J}(\xi_i) \tag{18a}
$$
$$
\text{s.t. : } (13b) \tag{18b}
$$

where

$$
\tilde{J}(\xi_i) = \left(\sum_{i=1}^{N_r} \left(\theta_i^2 \lambda_i \xi_i + \theta_i^2 \xi_i + 1\right)^{-1} \right) \left(\sum_{i=1}^{N_r} \left(\theta_i^2 \xi_i + 2\right) \right).
$$

This problem can be solved by numerical optimization methods [25].

D. Iterative Algorithm

In the aforementioned discussion, with predetermined decoding order and fixed $\mathbf{G}$, $\mathbf{F}_1$ and $\mathbf{F}_2$ can be optimized. For $\mathbf{F}_1$ and $\mathbf{F}_2$, $\mathbf{G}$ can be optimized. Therefore, we propose an iterative algorithm to jointly optimize $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{G}$ based on capacity. Note that the MSE-based algorithm can also be easily obtained. The convergence analysis of the proposed iterative algorithm is intractable. However, it can yield much better performance than the existing methods, which will be demonstrated by the simulation results in the next section.

In summary, we outline the nested iterative algorithm as follows.

Algorithm 1: A nested iterative algorithm.
- **Initialization: G.**
- **Repeat:** Update $k := k + 1$;
  —Compute $\mathbf{F}_1^{(k)}$ based on $\mathbf{G}^{(k)}$;
  —Compute $\mathbf{F}_2^{(k)}$ based on $\mathbf{G}^{(k)}$ and $\mathbf{F}_1^{(k)}$;
  —Compute $\mathbf{G}^{(k+1)} = \mathbf{V}_H \mathbf{\Xi} \mathbf{U}_K$ based on $\mathbf{F}_1^{(k)}$ and $\mathbf{F}_2^{(k)}$ by the following inner repeat to find $\mathbf{\Xi}$;
    ○ **Initial:** α;
    ○ **Inner Repeat:** Update $n := n + 1$;
      —Compute $\mathbf{\Xi}^{(n)}$ based on $\alpha^{(n)}$;
      —Compute $\alpha^{(n+1)}$ based on $\mathbf{\Xi}^{(n)}$;
    ○ **Inner Until:** Convergence.
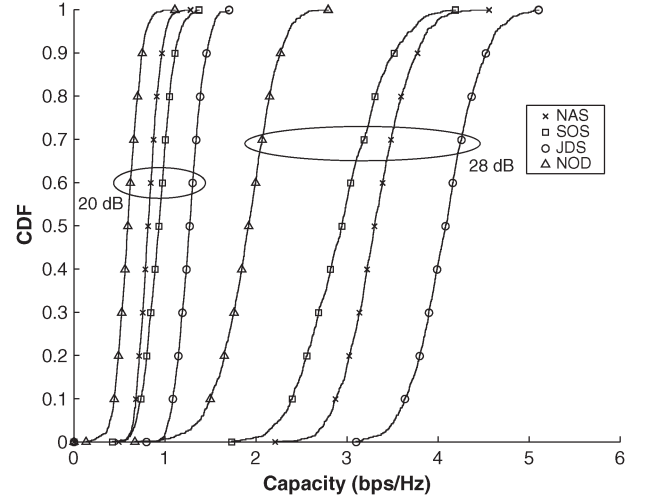- **Until:** The termination criterion is satisfied.

Fig. 2. CDF of the capacity for different power constraints, $P_1 = P_2 = P_r = 20$ db, and $P_1 = P_2 = P_r = 28$ db, $N_s = N_r = N_d = 4$, $\ell_{sd} = 10$, $\ell_{sr} = \ell_{rd} = 5$.

IV. SIMULATION RESULTS

In this section, simulation results are carried out to verify the performance superiority of the proposed joint source–relay design scheme for an MU-MIMO relay network with direct links. We first compare the proposed scheme with three other schemes in terms of the ergodic capacity and the cumulative distribution function (cdf) of the instantaneous capacity of the MIMO relaying networks and then compare the sum MSE of the networks. The alternative schemes listed as follows.

1) *Naive scheme (NAS)*. The source covariances are fixed to be scaled by the identity matrices $(P_1/N_S)\mathbf{I}$ and $(P_2/N_S)\mathbf{I}$ at SN_1 and SN_2, respectively, and the relay processing matrix is $\mathbf{G} = \eta\mathbf{I}$, where $\eta = \sqrt{P_r / (\mathrm{tr}(\mathbf{I} + \sum_{i=1}^2 \mathbf{H}_{ri}\mathbf{F}_i\mathbf{F}_i^\dagger \mathbf{H}_{ri}^\dagger))}$ is a power control factor. The S–D links contribution is included.

2) *Suboptimal scheme (SOS)*. This scheme is proposed in [17] for an MU-MIMO relay network without considering S–D links in the design. However, the S–D links contribution of capacity is included in the simulation for fair comparison. Note that this scheme is optimal for the scenario without considering the S–D links.

3) *No-direct-links scheme (NOD)*. This scheme is similar to SOS, but without S–D links contribution.

Note that both SOS and NOD have different power control polices to accommodate the capacity and MSE criteria. In the simulations, we consider a linear 2-D symmetric network geometry, as depicted in Fig. 1, where both SNs are deployed at the same position, and the distance between SNs (or RN) and the DN is set to be ℓ_{sd} (or ℓ_{rd}), and $\ell_{sd} = \ell_{sr} + \ell_{rd}$. The channel gains are modeled as the combination of large-scale fading (related to distance) and small-scale fading (Rayleigh fading), and all channel matrices have i.i.d. $\mathcal{CN}(0, 1/\ell^\tau)$ entries, where ℓ is the distance between two nodes, and $\tau = 3$ is the path-loss exponent.

Figs. 2–4 are based on the capacity criterion. Fig. 2 shows the cdf of instantaneous capacity for different power constraints when all node positions are fixed. Fig. 3 shows the capacity of the network versus the power constraints when all node positions are fixed. These two figures show that capacity offered by the proposed relaying scheme is better than both the SOS and NOD schemes at all SNR regimes, particularly at the high-SNR regime. NAS surpasses both the SOS and NOD schemes at the high-SNR regime, which demonstrates that the
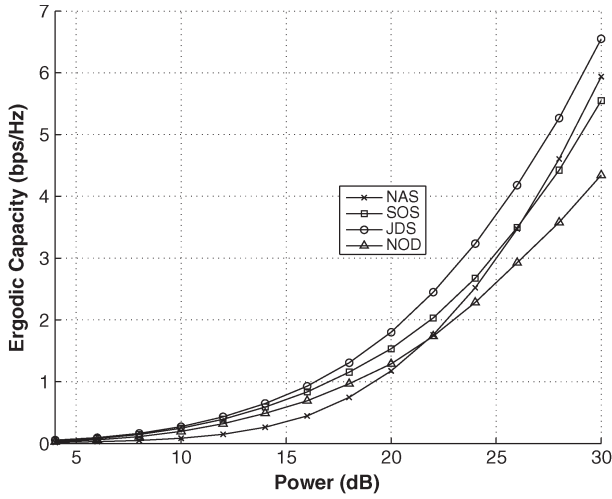
Fig. 3. Capacity versus power constraints $P_i (i = 1, 2, r)$ (in decibels), $P_1 = P_2 = P_r$, and $N_s = N_r = N_d = 4$, $\ell_{sd} = 10$, $\ell_{sr} = \ell_{rd} = 5$.
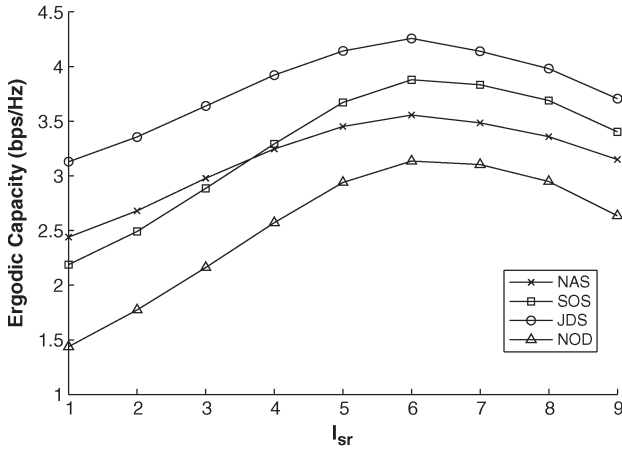

Fig. 5. Sum MSE versus the power constraints $P_i (i = 1, 2, r)$ (in decibels), $P_1 = P_2 = P_r$, and $N_s = N_r = N_d = 4$, $\ell_{sd} = 10$, $\ell_{sr} = \ell_{rd} = 5$.

Fig. 4. Capacity versus the distance between the source and the relay (ℓ_{sr}), $\ell_{sd} = 10$, $\ell_{rd} = \ell_{sd} - \ell_{sr}$, and $P_1 = P_2 = P_r = 26$ db, $N_s = N_r = N_d = 4$.

Fig. 6. Sum MSE versus the distance between the source and the relay (ℓ_{sr}), $\ell_{sd} = 10$, $\ell_{rd} = \ell_{sd} - \ell_{sr}$, and $P_1 = P_2 = P_r = 26$ db, $N_s = N_r = N_d = 4$.

direct S–D link should not be ignored in the design. Fig. 4 shows the capacity of the network versus the distance (ℓ_{sr}) between SNs and RN for fixed ℓ_{sd}. It is clear that the capacity offered by the proposed scheme is better than the SOS, NAS, and NOD schemes. The NOD scheme has the worst performance at any relay position at the moderate- and high-SNR regimes.

Figs. 5 and 6 are based on the MSE criterion, and similar conclusions can be drawn.

V. CONCLUSION

In this paper, we have proposed an optimal structure of the source precoding matrices and a relay processing matrix for an MU-MIMO nonregenerative relay network with direct S–D links based on capacity and MSE, respectively. We show that the capacity-based optimal source precoding matrices share the same structures with the MSE-based matrices, and so does the relay processing matrix. A nested iterative algorithm that jointly optimizes the source precoding and relay processing has been proposed. Simulation results show that the proposed algorithm provides better performance than the existing methods.
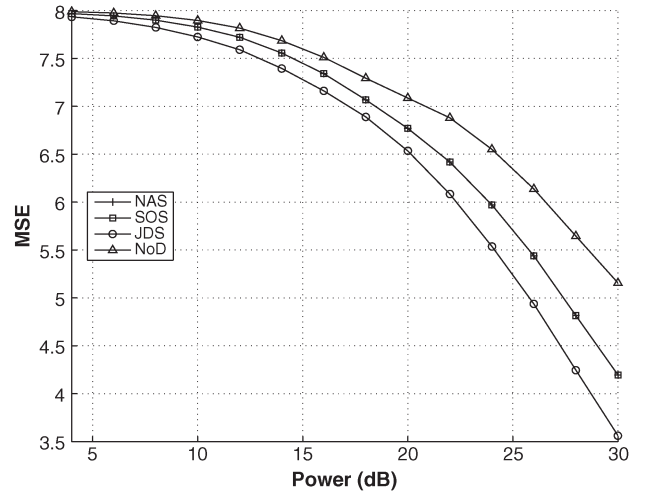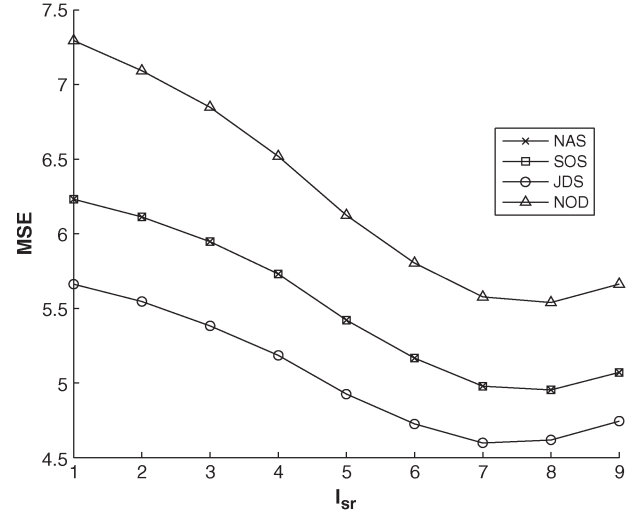
REFERENCES

[1] R. Pabst, B. H. Walke, D. C. Schultz, S. Mukherjee, H. Viswanathan, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 80–89, Sep. 2004.

[2] L. Zheng and D. N. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 8, no. 1, pp. 2–25, Aug. 2002.

[3] B. Wang, J. Zhang, and A. Hst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[4] Z. Wang, W. Chen, F. Gao, and J. Li, "Capacity performance of relay beamformings for MIMO multirelay networks with imperfect $\mathcal{R} - \mathcal{D}$ CSI at relays," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2608–2619, Jul. 2011.

[5] C. K. Lo, S. Vishwanath, and J. R. W. Heath, "Rate bounds for MIMO relay channels using precoding," in *Proc. IEEE GLOBECOM*, St.Louis, MO, Nov. 2005, vol. 3, pp. 1172–1176.

[6] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Inf. Theory*, vol. 5, no. 6, pp. 1433–1444, Jun. 2006.

[7] H. Wan, W. Chen, and J. Ji, "Efficient linear transmission strategy for MIMO relaying broadcast channels with direct links," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 14–17, Feb. 2012.

[8] X. Tang and Y. Hua, "Optimal design of nonregenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007.

[9] O. M. Munoz Medina, J. Vidal, and A. Agustıacute;n, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal. Process.*, vol. 55, no. 6, pp. 2593–2604, Jun. 2007.

[10] Z. Fang, Y. Hua, and J. C. Koshy, "Joint source and relay optimization for a nonregenerative MIMO relay," in *Proc. IEEE Workshop Sens. Array Multichannel Process.*, Waltham, MA, 2006, pp. 239–243.

[11] W. Guan and H. Luo, "Joint MMSE transceiver design in nonregenerative MIMO relays systems," *IEEE Commun. Lett.*, vol. 12, no. 7, pp. 517–519, Jul. 2008.

[12] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal. Process.*, vol. 56, no. 10, pp. 5062–5073, Oct. 2008.

[13] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal. Process.*, vol. 57, no. 12, pp. 4837–4851, Dec. 2009.

[14] L. Weng and R. D. Murch, "Multiuser MIMO relay system with self-interference cancellation," in *Proc. IEEE WCNC*, 2007, pp. 958–962.

[15] C.-B. Chae, T. Tang, R. W. Heath, and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. Signal. Process.*, vol. 56, no. 2, pp. 727–738, Feb. 2008.

[16] K. S. Gomadam and S. A. Jafar, "Duality of MIMO multiple access channel and broadcast channel with amplify-and-forward relays," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 211–217, Jan. 2010.

[17] Y. Yu and Y. Hua, "Power allocation for a MIMO relay system with multiple-antenna users," *IEEE Trans. Signal. Process.*, vol. 58, no. 5, pp. 2823–2835, May 2010.

[18] Y. Rong and F. Gao, "Optimal beamforming for nonregenerative MIMO relays with direct link," *IEEE Commun. Lett.*, vol. 13, no. 12, pp. 926–928, Dec. 2009.

[19] Y. Rong, "Optimal joint source and relay beamforming for MIMO relays with direct link," *IEEE Commun. Lett.*, vol. 14, no. 5, pp. 390–392, May 2010.

[20] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[21] S. M. Kay, *Fundamental of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[22] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[23] E. H. Lieb and W. Thirring, *Studies in Mathematical Physics, Essays in Honor of Valentine Bartmann*. Princeton, NJ: Princeton Univ. Press, 1976.

[24] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 1996.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[26] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

Comments on "Performance Analysis of MRC Diversity for Cognitive Radio Systems"

Wei Xu, Jianhua Zhang, and Ping Zhang

Abstract—The exact asymptotic symbol error rate (SER) for cognitive ratio systems with maximum ratio combining (MRC) considered by Li is derived, which circumvents the precondition requirement in the derivation made by Li. Monte Carlo simulations validate the theoretical analysis.

Index Terms—Cognitive radio, diversity order, maximum ratio combining (MRC), symbol error rate (SER).

I. INTRODUCTION

In [1], the ergodic capacity and average symbol error rate (SER) are asymptotically investigated to show the capacity scaling law and the achievable diversity order for the cognitive ratio system with maximum ratio combining (MRC). However, the asymptotic average SER analysis in [1, Sec. IV] is only evaluated under the condition $P \ll Q$, where P is the maximum transmit power of the cognitive user and Q is the interference power constraint at the primary user. However, if this condition does not hold,[1] the accuracy of the asymptotic SER in [1] will decrease. In this commentary, it is pointed out that the requirement $P \ll Q$ is not necessary for the derivation of exact asymptotic SER. Moreover, it is shown that the asymptotic SER in [1] is just a special case of the general results, which was derived in this commentary. The numerical and Monte Carlo simulations are also presented to verify the theoretical observations.

II. EXACT ASYMPTOTIC SYMBOL ERROR RATE ANALYSIS

According to [1, eq. (3)], the received signal-to-noise ratio (SNR) at the cognitive receiver can be formulated as

$$\gamma_{\mathrm{MRC}} = p\overline{\gamma} \sum_{i=1}^{K} g_{c_i} \qquad (1)$$

where $p = \min\{P, Q/g_p\}$ is the transmit power of the cognitive transmitter, $\overline{\gamma}$ is the inverse of the power of additive white Gaussian noise, K is the number of the cognitive receive antennas, and g_{c_i}, $i = 1, 2, \ldots K$ is the channel power from the cognitive transmitter to the ith cognitive receive antenna [1].

[1]It should be noted that $P \ll Q$ may not be always satisfied for the practical cognitive radio system, in which the interference tolerance level, i.e., Q, is always finite [2].