Adaptive Cooperative Transmission and Spectrum Sharing in MIMO-CCRN

Zhao Li, Xinyi Kong, Jiandong Li
State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an, China
zli19912@gmail.com, xy_kong@126.com, jdli@mail.xidian.edu.cn

***Abstract*—A two-phase spatiotemporal adaptive transmission and spectrum sharing mechanism is designed for multi-input multi-output cooperative cognitive radio network (MIMO-CCRN). By employing cognitive user(s) as relay(s), performance of end-to-end primary transmission is improved. On the other hand, secondary user (SU) can access to the spectrum for its own data transmission simultaneously with the primary, as a reward for assisting the primary's data delivery. With the proposed scheme, multiple SU pairs participate in the data relaying, and share the access authority following a time division multiplexing mode. By joint transmit precoding and receive filtering design, co-channel interference can be appropriately suppressed. Moreover, bottleneck effect in the end-to-end primary link is studied, and an adaptive time slot division scheme is employed as a solution to bottleneck elimination. Simulation results show that the primary spectral efficiency is improved with the assistance of SU(s). Time domain adaptation could effectively enhance the end-to-end performance. Meanwhile, spatiotemporal multiplexing of different types of transmission facilitates the spectral efficiency. As a result, a win-win situation for both the primary and secondary is achieved.**

Keywords—cooperative; cognitive radio; MIMO; interference; bottleneck

I. INTRODUCTION

With the rapid growth of wireless service, spectrum resource available for communication becomes more and more limited. On the other hand, there is serious waste in spectrum use [1]. Cognitive radio (CR) emerges as a promising technique to improve spectrum utilization [2]. After years of research, CR has been extended from opportunistic access to spectrum holes to several spectrum sharing paradigms including overlay, underlay, and interweave [3-4].

Recently, cooperative cognitive radio network (CCRN) [5-11] is proposed, in which SU obtains access authorities in frequency, time, etc. as a reward to carry out its own transmission, by assisting in relaying the primary information. In [6-8], cooperation schemes based on time slot division are proposed in which primary users (PUs) lease their spectrum to SUs for a fraction of time and in exchange, they get the cooperative transmission power from SUs. CR user chooses a power level to cooperate with the primary link to earn the access time which is related to its afforded power level. With this mechanism, a dedicated time portion should be assigned to the SUs. In [9], an FDMA based two-phase cooperation is established, in which a PU divides its spectrum into two orthogonal subbands, and broadcasts on the first subband in phase one. SU relays on the same subband in the second phase, and continuously transmits its own data on the second subband during two phases. With this scheme, PU only uses a fraction of its licensed bandwidth. Compared to the above works, [10-11] exploit new types of resource to avoid leasing dedicated portion of time or spectrum to SUs, facilitating both primary and secondary transmissions. In [10], a MIMO enabled two-phase cooperation is presented. By leveraging the degrees of freedom (DoF) provided by multi-antenna systems, SUs relay and carry out their own transmission simultaneously in the second phase. In [11], the DoF provided by orthogonal modulation is exploited in designing a two-phase cooperation scheme, by which SUs are able to relay the PU's traffic using the quadrature channel and transmit their own in in-phase modulation in the same time slot without interference.

MIMO technique can exploit spatial domain communication resource and has been accepted by LTE and LTE-A standards [12]. By incorporating MIMO with CCRN [10, 13-14], benefits such as interference suppression and spectral efficiency enhancement can be obtained. However, co-channel interference (CCI) is simplified unrealistically in [10], thus in practical use where CCI exists, the actual system performance will be deteriorated. In [13-14], algorithms that jointly determine beamforming vectors and power allocation for both the PU's and SU's data transmission are proposed in a simplified MIMO-CCRN model. However, in both papers only the secondary transmitter (ST) is employed as relay, yet the case in which secondary receiver (SR) operates as relay is neglected. Moreover, in the system models only ST is equipped with multiple antennas while the other devices are single antenna entities. Such schemes fail to take full advantage of the multiplexing capability of multi-antenna systems. In particular, since interference between PU and SU is not well studied in the above works, the methods cannot be extended to practical use.

In this paper, an adaptive cooperative transmission mechanism is proposed in MIMO-CCRN, both ST and SR can be recruited as relay. Multiple SU pairs equipped with multi-antennas are considered. Specifically, interference among concurrent data streams is investigated. By joint design of transmit precoding and receive filtering, the secondary traffic could achieve harmless coexistence with the primary in spatial domain, while in the time domain multiple SU transmitter-receiver pairs share spectrum in a TDMA mode.

978-1-4799-4912-0/14/$31.00 ©2014 IEEE

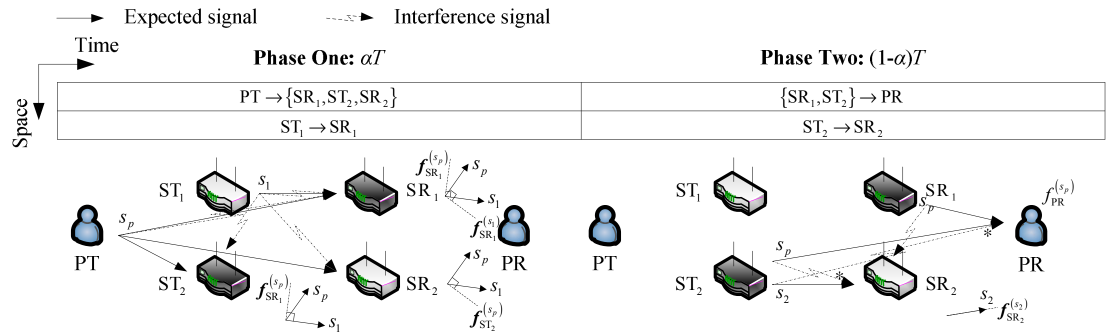

Fig. 1. System model and frame structure.

Throughout this paper, we use the following notation. The set of complex numbers is denoted as $\mathbb{C}$, while vectors and matrices are denoted by bold lower-case and upper-case letters, respectively. $\mathbb{E}(\cdot)$ denotes statistical expectation. The Hermitian (or conjugate transpose) and pseudo inverse of a vector or a matrix are denoted as $(\cdot)^H$ and $(\cdot)^\dagger$, respectively. The Euclidean norm of a vector and matrix is denoted as $\|\cdot\|$, whereas $|\cdot|$ denotes the absolute value of a scalar number. $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$ represents for the inner product of two vectors.

II. System Model

Consider a MIMO-CCRN consisting of one PU pair and N SU pairs. Since the primary transmitter (PT) may be far from its intended receiver (PR) or the link between PT and PR is blocked by buildings or undergoes deep fading, the direct link is not able to support the primary service. So PU is willing to employ SU(s) as relay(s). In this paper, we assume that the direct link can only support a reliable low rate control signal delivery, i.e., data transmission is via relay node(s) and no direct data link exists. In the network topology, a secondary network consisting of N SU pairs lies between the PU pair. Each SU can receive from PT and transmit to PR. Either ST or SR in each pair can be recruited as relay. The primary link could adopt a subset of SUs, denoted by $\mathcal{R}$, to participate in the cooperative transmission. The relay(s) can be selected based on certain criterions [15]. In this paper, we skip over the relay selection without loss of generality. Moreover, the overhead due to relay selection, control information exchange, etc. is negligible compared to the data transmission period.

As shown in Fig. 1, a time-division frame structure is used. The frame duration T is divided into two phases. The first phase is of duration αT ($0 < \alpha < 1$), whereas the second phase is $(1-\alpha)T$. For simplicity, we fixedly set $N = 2$ and $\mathcal{R} = \{\mathrm{SR}_1, \mathrm{ST}_2\}$. Both situations in which ST and SR are employed as relays are included. Two SU pairs share the communication opportunity based on TDMA. Note that the dashed line at the receiver indicates the receiving direction determined by the filtering vector. Specifically, in existing works about MIMO-CCRN [10, 13-14] only interferences marked as asterisk are taken into account. However, in practical the interference problem may become more complicated as illustrated in the figure.

The SUs are equipped with $M = 2$ antennas. We do not consider multiple antennas at PUs since our focus is on the processing at SUs. We use $\boldsymbol{H}_{s,d} \in \mathbb{C}^{M_d \times M_s}$ and $\boldsymbol{g}_{s,d} \in \mathbb{C}^{M_d \times M_s}$ to denote the channel matrix and vector from transmitter s to receiver d, whose element is modeled as independent and identically distributed zero-mean unit-variance complex Gaussian random variables, invariant during one frame, but generally varying over the frame, i.e., quasi-static block-fading model [10-11] is assumed. We assume that CSI is available at SUs and PR. CSI can be obtained [16] by channel training, estimation, feedback mechanisms, and etc. Compared to the cases where accurate CSI is not available, the results in this paper provide performance upper-bound. The primary data is denoted by s_p, whereas s_i ($i = 1,2$) represents for the secondary data from ST_i to SR_i. $\mathbb{E}(s_p s_p^H) = \mathbb{E}(s_i s_i^H) = 1$ holds. The transmit power at PT is P_T. As for relay node, power allocated for s_p and s_i are denoted by $P_s^{(s_p)}$ and $P_s^{(s_i)}$, respec-tively. Let $P_s^{(s_p)} = P_s^{(s_i)} = P_T/2$. The detailed communication procedure is described in detail as follows.

In phase one, PT broadcasts s_p to SUs in $\mathcal{R}$. ST_1 is granted for its own transmission since its partner SR_1 is employed as relay. Both ST_2 and SR_1 decode s_p. Although s_p is not the expected data for SR_2, SR_2 recovers s_p in order to implement interference cancellation (IC) in phase two. In the first phase, both ST_2 and SR_2 face the interference from ST_1. In phase two, SR_1 and ST_2 employs decode-and-forward (DF) to forward s_p to PR. ST_2 also sends its own data to SR_2. Since PR has only one antenna, its receiving direction cannot be adjusted to nullify the coming interference. Zero-forcing beamforming (ZFBF) is employed at ST_2 to pre-cancel the interference at PR. As for SR_2, there are three signal components, including one desired signal from ST_2 and two interferences from ST_2 and SR_1, respectively. By exploiting decoded data s_p in the last phase, IC is used to simplify the reception at SR_2, and then s_2 is obtained.

III. Cooperative Transmission and Spectrum Sharing in MIMO-CCRN

$$\tilde{\boldsymbol{r}}_{d}^{(s_x)} = \sqrt{P_{\rm PT}^{(s_p)}}\left[\boldsymbol{f}_{d}^{(s_x)}\right]^H \boldsymbol{g}_{{\rm PT},d} s_p + \sqrt{P_{{\rm ST}_1}^{(s_1)}}\left[\boldsymbol{f}_{d}^{(s_x)}\right]^H \boldsymbol{H}_{{\rm ST}_1,d}\boldsymbol{p}_{{\rm ST}_1}^{(s_1)} s_1 + \left[\boldsymbol{f}_{d}^{(s_x)}\right]^H \boldsymbol{n} \quad (2)$$

$$\boldsymbol{r}_{\rm PR} = \sqrt{P_{{\rm SR}_1}^{(s_p)}}\boldsymbol{g}_{{\rm SR}_1,{\rm PR}}\boldsymbol{p}_{{\rm SR}_1}^{(s_p)} s_p + \sqrt{P_{{\rm ST}_2}^{(s_p)}}\boldsymbol{g}_{{\rm ST}_2,{\rm PR}}\boldsymbol{p}_{{\rm ST}_2}^{(s_p)} s_p + \sqrt{P_{{\rm ST}_2}^{(s_2)}}\boldsymbol{g}_{{\rm ST}_2,{\rm PR}}\boldsymbol{p}_{{\rm ST}_2}^{(s_2)} s_2 + \boldsymbol{n} \quad (3)$$

$$\boldsymbol{r}_{{\rm SR}_2} = \sqrt{P_{{\rm SR}_1}^{(s_p)}}\boldsymbol{H}_{{\rm SR}_1,{\rm SR}_2}\boldsymbol{p}_{{\rm SR}_1}^{(s_p)} s_p + \sqrt{P_{{\rm ST}_2}^{(s_p)}}\boldsymbol{H}_{{\rm ST}_2,{\rm SR}_2}\boldsymbol{p}_{{\rm ST}_2}^{(s_p)} s_p + \sqrt{P_{{\rm ST}_2}^{(s_2)}}\boldsymbol{H}_{{\rm ST}_2,{\rm SR}_2}\boldsymbol{p}_{{\rm ST}_2}^{(s_2)} s_2 + \boldsymbol{n} \quad (4)$$

In this section, a two-phase spatiotemporal cooperation mechanism is proposed. The SUs assist the primary. As a reward, they are granted for accessing spectrum in a spatial multiplexing manner with the primary. As for the secondary, they share the medium in a TDMA mode.

A. Signal processing in phase one

In this phase, PT broadcasts s_p to $\rm SR_1$, $\rm ST_2$ and $\rm SR_2$. $\rm SR_1$ and $\rm ST_2$ decode s_p so as to forward to PR in the next phase. $\rm SR_2$ recovers s_p in phase one and exploits this information to implement IC in phase two. As a reward for $\rm SR_1$'s offering help to the primary, its partner $\rm ST_1$ is authorized for transmitting in this stage. So $\rm SR_1$ also needs to decode its own data s_1. As the receivers of s_p, $\rm ST_2$ and $\rm SR_2$ suffer from interference from $\rm ST_1$. The received signal at $\rm SR_1$, $\rm ST_2$ and $\rm SR_2$ can be expressed by a general formula:

$$\boldsymbol{r}_{d} = \sqrt{P_{\rm PT}^{(s_p)}}\boldsymbol{g}_{{\rm PT},d} s_p + \sqrt{P_{{\rm ST}_1}^{(s_1)}}\boldsymbol{H}_{{\rm ST}_1,d}\boldsymbol{p}_{{\rm ST}_1}^{(s_1)} s_1 + \boldsymbol{n} \quad (1)$$

where $\boldsymbol{H}_{s,d} \in \mathbb{C}^{2\times 2}$, $\boldsymbol{g}_{s,d} \in \mathbb{C}^{2\times 1}$, and $d \in \{{\rm SR}_1, {\rm ST}_2, {\rm SR}_2\}$. $\boldsymbol{p}_s^{(s_x)} \in \mathbb{C}^{2\times 1}$ denotes the precoder at transmitter $s \in \{{\rm PT}, {\rm ST}_1\}$ for data s_x where subscript $x \in \{1,2\}$. Transmit power $P_{\rm PT}^{(s_p)} = P_T$ and $P_{{\rm ST}_1}^{(s_1)} = P_T/2$. The first two terms on the right hand side (RHS) of (1) indicate the received signal at d from PT and $\rm ST_1$, respectively. $\boldsymbol{n}$ represents for the additive white Gaussian noise (AWGN) with zero-mean and variance σ_n^2. The receiver d employs filtering vector $\boldsymbol{f}_{d}^{(s_x)}$ to decode s_x. The estimated signal at d is given by (2).

In the following, $\boldsymbol{p}_s^{(s_x)}$ and $\boldsymbol{f}_{d}^{(s_x)}$ are to be determined. Substituting $s_x = s_p$ and $d = {\rm SR}_1$ into (2), the estimated signal s_p at $\rm SR_1$ is obtained. Since the transmission of secondary pair $({\rm ST}_1, {\rm SR}_1)$ is of lower priority with respect to the primary, $\rm ST_1$ adopts precoder $\boldsymbol{p}_{{\rm ST}_1}^{(s_1)}$ orthogonal to the equivalent matrix $\boldsymbol{H}_{{\rm ST}_1,{\rm SR}_1}^{eq} = [\boldsymbol{f}_{{\rm SR}_1}^{(s_p)}]^H \boldsymbol{H}_{{\rm ST}_1,{\rm SR}_1}$ so that the interference part, $[\boldsymbol{f}_{{\rm SR}_1}^{(s_p)}]^H \boldsymbol{H}_{{\rm ST}_1,{\rm SR}_1}\boldsymbol{p}_{{\rm ST}_1}^{(s_1)}$ becomes zero. Meanwhile, the filter for decoding s_p at $\rm SR_1$ should match with $\boldsymbol{g}_{{\rm PT},{\rm SR}_1}$ so as to achieve as good reception performance as possible, i.e., $[\boldsymbol{f}_{{\rm SR}_1}^{(s_p)}]^H = \boldsymbol{g}_{{\rm PT},{\rm SR}_1}^{\dagger} / \left\|\boldsymbol{g}_{{\rm PT},{\rm SR}_1}^{\dagger}\right\|$. To obtain $\boldsymbol{p}_{{\rm ST}_1}^{(s_1)}$, we apply singular value decomposition (SVD) to $\boldsymbol{H}_{{\rm ST}_1,{\rm SR}_1}^{eq}$ and get $\boldsymbol{H}_{{\rm ST}_1,{\rm SR}_1}^{eq} = \boldsymbol{U}_{{\rm ST}_1,{\rm SR}_1}^{eq}\boldsymbol{\Sigma}_{{\rm ST}_1,{\rm SR}_1}^{eq}(\boldsymbol{V}_{{\rm ST}_1,{\rm SR}_1}^{eq})^H$ where $\boldsymbol{V}_{{\rm ST}_1,{\rm SR}_1}^{eq} = [\boldsymbol{v}_{{\rm ST}_1,{\rm SR}_1}^{eq,(1)}\ \boldsymbol{v}_{{\rm ST}_1,{\rm SR}_1}^{eq,(2)}]$. Note that $\boldsymbol{H}_{{\rm ST}_1,{\rm SR}_1}^{eq} \in \mathbb{C}^{1\times 2}$, a solution $\boldsymbol{p}_{{\rm ST}_1}^{(s_1)} = \boldsymbol{v}_{{\rm ST}_1,{\rm SR}_1}^{eq,(2)}$ exists, satisfying $[\boldsymbol{f}_{{\rm SR}_1}^{(s_p)}]^H \boldsymbol{H}_{{\rm ST}_1,{\rm SR}_1}\boldsymbol{p}_{{\rm ST}_1}^{(s_1)} = 0$.

Besides decoding s_p, $\rm SR_1$ also needs to recover its own data s_1. We substitute $s_x = s_1$ and $d = {\rm SR}_1$ into (2), the estimated signal s_1 at $\rm SR_1$ is acquired. Similar to the above discussion, $\rm SR_1$ calculates $\boldsymbol{f}_{{\rm SR}_1}^{(s_1)}$ to meet $[\boldsymbol{f}_{{\rm SR}_1}^{(s_1)}]^H \boldsymbol{g}_{{\rm PT},{\rm SR}_1} = 0$, then s_1 can be extracted from $\boldsymbol{r}_{{\rm SR}_1}$. Applying SVD to $\boldsymbol{g}_{{\rm PT},{\rm SR}_1}$, we get $\boldsymbol{g}_{{\rm PT},{\rm SR}_1} = \boldsymbol{U}_{{\rm PT},{\rm SR}_1}\boldsymbol{\Sigma}_{{\rm PT},{\rm SR}_1}\boldsymbol{V}_{{\rm PT},{\rm SR}_1}^H$. Since $\boldsymbol{g}_{{\rm PT},{\rm SR}_1} \in \mathbb{C}^{2\times 1}$, $\boldsymbol{f}_{{\rm SR}_1}^{(s_1)} = \boldsymbol{u}_{{\rm PT},{\rm SR}_1}^{(2)}$ is obtained.

As for the reception of s_p at $\rm ST_2$ and $\rm SR_2$, s_1 is interference. Take $\boldsymbol{f}_{{\rm ST}_2}^{(s_p)}$ as an example, the estimated s_p at $\rm ST_2$ can be obtained by substituting $s_x = s_p$ and $d = {\rm ST}_2$ into (2). Applying SVD to $\boldsymbol{H}_{{\rm ST}_1,{\rm ST}_2}^{eq} = \boldsymbol{H}_{{\rm ST}_1,{\rm ST}_2}\boldsymbol{p}_{{\rm ST}_1}^{(s_1)}$, we have $\boldsymbol{H}_{{\rm ST}_1,{\rm ST}_2}^{eq} = \boldsymbol{U}_{{\rm ST}_1,{\rm ST}_2}^{eq}\boldsymbol{\Sigma}_{{\rm ST}_1,{\rm ST}_2}^{eq}(\boldsymbol{V}_{{\rm ST}_1,{\rm ST}_2}^{eq})^H$ where $\boldsymbol{U}_{{\rm ST}_1,{\rm ST}_2}^{eq} = [\boldsymbol{u}_{{\rm ST}_1,{\rm ST}_2}^{eq,(1)}\ \boldsymbol{u}_{{\rm ST}_1,{\rm ST}_2}^{eq,(2)}]$. Since $\boldsymbol{H}_{{\rm ST}_1,{\rm ST}_2}^{eq} \in \mathbb{C}^{2\times 1}$, $\boldsymbol{f}_{{\rm ST}_2}^{(s_p)} = \boldsymbol{u}_{{\rm ST}_1,{\rm ST}_2}^{eq,(2)}$ is obtained to satisfy $[\boldsymbol{f}_{{\rm ST}_2}^{(s_p)}]^H \boldsymbol{H}_{{\rm ST}_1,{\rm ST}_2}\boldsymbol{p}_{{\rm ST}_1}^{(s_1)} = 0$.

B. Signal processing in phase two

In this phase, the received signal at $\rm SR_2$ is composed of relaying traffic from $\rm SR_1$ and $\rm ST_2$ to PR, as well as the data transmission from $\rm ST_2$ to $\rm SR_2$. In the following, $\rm SR_2$ exploits s_p obtained in the last phase to cancel the interference from $\rm SR_1$ and $\rm ST_2$, so that the reception of s_2 is simplified. On the other hand, since PR has only one antenna, its receiving direction cannot be adjusted to eliminate the coming interference. In order to solve this problem, $\rm ST_2$ should properly design its precoder so that no interference is yielded at PR.

The received signal at PR and $\rm SR_2$ is given by (3) and (4), respectively. We fixedly set $P_{{\rm SR}_1}^{(s_p)} = P_{{\rm ST}_2}^{(s_p)} = P_{{\rm ST}_2}^{(s_2)} = P_T/2$. By applying SVD to $\boldsymbol{g}_{{\rm ST}_2,{\rm PR}}$, $\boldsymbol{g}_{{\rm ST}_2,{\rm PR}} = \boldsymbol{U}_{{\rm ST}_2,{\rm PR}}\boldsymbol{\Sigma}_{{\rm ST}_2,{\rm PR}}\boldsymbol{V}_{{\rm ST}_2,{\rm PR}}^H$ is obtained, where $\boldsymbol{V}_{{\rm ST}_2,{\rm PR}} = [\boldsymbol{v}_{{\rm ST}_2,{\rm PR}}^{(1)}\ \boldsymbol{v}_{{\rm ST}_2,{\rm PR}}^{(2)}]$. Let $\boldsymbol{p}_{{\rm ST}_2}^{(s_2)} = [\boldsymbol{v}_{{\rm ST}_2,{\rm PR}}^{(2)}]^H$ so that the sending of s_2 can be nullified at PR, i.e., $\boldsymbol{g}_{{\rm ST}_2,{\rm PR}}\boldsymbol{p}_{{\rm ST}_2}^{(s_2)}$ in (3) becomes zero.

Recall that PU has higher priority than SU, in order to obtain as large desired SNR at PR as possible, the precoders $\boldsymbol{p}_{{\rm SR}_1}^{(s_p)}$ and $\boldsymbol{p}_{{\rm ST}_2}^{(s_p)}$ should match with $\boldsymbol{g}_{{\rm SR}_1,{\rm PR}}$ and $\boldsymbol{g}_{{\rm ST}_2,{\rm PR}}$, respectively. Take $\boldsymbol{p}_{\mathrm{ST}_2}^{(s_p)}$ as an example, $\boldsymbol{p}_{\mathrm{ST}_2}^{(s_p)} = \boldsymbol{g}_{\mathrm{ST}_2,\mathrm{PR}}^{\dagger} \Big/ \left\| \boldsymbol{g}_{\mathrm{ST}_2,\mathrm{PR}}^{\dagger} \right\|$. The received signal at PR after pre-cancelling interference from ST_2 becomes:

$$\boldsymbol{r}_{\mathrm{PR}}^{(s_p)} = \sqrt{P_T/2}\left[\boldsymbol{g}_{\mathrm{SR}_1,\mathrm{PR}}\boldsymbol{p}_{\mathrm{SR}_1}^{(s_p)} + \boldsymbol{g}_{\mathrm{ST}_2,\mathrm{PR}}\boldsymbol{p}_{\mathrm{ST}_2}^{(s_p)}\right]s_p + \boldsymbol{n} \tag{5}$$

We assume that PR is aware of channel vectors $\boldsymbol{g}_{\mathrm{SR}_1,\mathrm{PR}}$ and $\boldsymbol{g}_{\mathrm{ST}_2,\mathrm{PR}}$, as well as precoders $\boldsymbol{p}_{\mathrm{SR}_1}^{(s_p)}$ and $\boldsymbol{p}_{\mathrm{ST}_2}^{(s_p)}$. Then s_p is readily obtained.

The processing at SR_2 includes two parts. First, since $\boldsymbol{H}_{\mathrm{ST}_2,\mathrm{SR}_2}$, $\boldsymbol{H}_{\mathrm{SR}_1,\mathrm{SR}_2}$, $\boldsymbol{p}_{\mathrm{SR}_1}^{(s_p)}$ and $\boldsymbol{p}_{\mathrm{ST}_2}^{(s_p)}$ are available at SR_2, the interference terms $\boldsymbol{H}_{\mathrm{ST}_2,\mathrm{SR}_2}\boldsymbol{p}_{\mathrm{ST}_2}^{(s_p)}s_p$ and $\boldsymbol{H}_{\mathrm{SR}_1,\mathrm{SR}_2}\boldsymbol{p}_{\mathrm{SR}_1}^{(s_p)}s_p$ in (4) can be eliminated by exploiting these information along with the decoded s_p. Then (6) is obtained as follows:

$$\boldsymbol{r}_{\mathrm{SR}_2}^{(s_2)} = \sqrt{P_{\mathrm{ST}_2}^{(s_2)}}\boldsymbol{H}_{\mathrm{ST}_2,\mathrm{SR}_2}\boldsymbol{p}_{\mathrm{ST}_2}^{(s_2)}s_2 + \boldsymbol{n} \tag{6}$$

Secondly, SR_2 computes $(\boldsymbol{H}_{\mathrm{ST}_2,\mathrm{SR}_2}^{eq})^{\dagger}$ where $\boldsymbol{H}_{\mathrm{ST}_2,\mathrm{SR}_2}^{eq} = \boldsymbol{H}_{\mathrm{ST}_2,\mathrm{SR}_2}\boldsymbol{p}_{\mathrm{ST}_2}^{(s_2)}$. Then $\boldsymbol{f}_{\mathrm{SR}_2}^{(s_2)}$ is obtained by normalizing $\left[(\boldsymbol{H}_{\mathrm{ST}_2,\mathrm{SR}_2}^{eq})^{\dagger}\right]^H$. s_2 can be recovered according to (7).

$$\tilde{\boldsymbol{r}}_{\mathrm{SR}_2}^{(s_2)} = \sqrt{P_{\mathrm{ST}_2}^{(s_2)}}\left[\boldsymbol{f}_{\mathrm{SR}_2}^{(s_2)}\right]^H \boldsymbol{H}_{\mathrm{ST}_2,\mathrm{SR}_2}\boldsymbol{p}_{\mathrm{ST}_2}^{(s_2)}s_2 + \left[\boldsymbol{f}_{\mathrm{SR}_2}^{(s_2)}\right]^H \boldsymbol{n} \tag{7}$$

C. *Achievable spectral efficiency analysis*

We now analyze the achievable spectral efficiency of different links. As for multi-antenna based communication, equal power allocation is adopted for its simplicity. Since there is no direct link from PT to PR, the end-to-end achievable rate is limited by the worse hop of the two [17]. Moreover, assume that relay node operates in half-duplex mode. The achievable spectral efficiency of the two-hop transmission via relay $r \in \mathcal{R}$ ($\mathcal{R} = \{\mathrm{SR}_1, \mathrm{ST}_2\}$) is given by (8) as follows:

$$C_{\mathrm{PT},r,\mathrm{PR}}^{(s_p)} = \min\left\{\alpha\log_2\left[1+\gamma_{\mathrm{PT},r}^{(s_p)}\right], (1-\alpha)\log_2\left[1+\gamma_{r,\mathrm{PR}}^{(s_p)}\right]\right\} \tag{8}$$

$\gamma_{\mathrm{PT},r}^{(s_p)} = P_T \left\| [\boldsymbol{f}_r^{(s_p)}]^H \boldsymbol{g}_{\mathrm{PT},r} \right\|^2 \Big/ \sigma_n^2$ and $\gamma_{r,\mathrm{PR}}^{(s_p)} = P_r^{(s_p)} \left\| \boldsymbol{g}_{r,\mathrm{PR}} \boldsymbol{p}_r^{(s_p)} \right\|^2 \Big/ \sigma_n^2$ denote the SNR at r and PR, respectively. For $r = \mathrm{SR}_1$, SR_1 forwards s_p in the second phase. $P_T/2$ is allocated to the relaying transmission. For $r = \mathrm{ST}_2$, ST_2 transmits to SR_2 and relays to PR simultaneously, we adopt $P_{\mathrm{ST}_2}^{(s_p)} = P_{\mathrm{ST}_2}^{(s_2)} = P_T/2$, i.e., two data streams are allocated the same power.

If two relays are employed, the end-to-end spectral efficiency $C_{\mathrm{PT},\mathcal{R},\mathrm{PR}}^{(s_p)}$ can be calculated as follows:

$$C_{\mathrm{PT},\mathcal{R},\mathrm{PR}}^{(s_p)} = \min\left[\alpha \min_{r\in\mathcal{R}} \log_2\left[1+\gamma_{\mathrm{PT},r}^{(s_p)}\right], (1-\alpha)\log_2\left[1+\gamma_{\mathcal{R},\mathrm{PR}}^{(s_p)}\right]\right] \tag{9}$$

where $\gamma_{\mathcal{R},\mathrm{PR}}^{(s_p)} = P_T\left(\varphi^{eq}\right)^2 \Big/ \left(2\sigma_n^2\right)$ and $\varphi^{eq} = \boldsymbol{g}_{\mathrm{SR}_1,\mathrm{PR}}\boldsymbol{p}_{\mathrm{SR}_1}^{(s_p)} + \boldsymbol{g}_{\mathrm{ST}_2,\mathrm{PR}}\boldsymbol{p}_{\mathrm{ST}_2}^{(s_p)}$. Note that if $\gamma_{\mathcal{R},\mathrm{PR}}^{(s_p)}$ is larger than $\min_{r\in\mathcal{R}} \gamma_{\mathrm{PT},r}^{(s_p)}$, the combination of relaying signals at PR is unnecessary, since the bottleneck exists in the first hop. In this case, we can simply select one relay r_{opt} through which the maximum end-to-end spectral efficiency is achieved, i.e., $C_{\mathrm{PT},r_{opt},\mathrm{PR}}^{(s_p)} = \max_{r\in\mathcal{R}} C_{\mathrm{PT},r,\mathrm{PR}}^{(s_p)}$.

As for the direct transmission from ST_i to SR_i ($i = 1,2$), the spectral efficiency is given by (10) and (11), respectively.

$$C_{\mathrm{ST}_1,\mathrm{SR}_1}^{(s_1)} = \alpha\log_2\left\{1+\gamma_{\mathrm{ST}_1,\mathrm{SR}_1}^{(s_1)}\right\} \tag{10}$$

$$C_{\mathrm{ST}_2,\mathrm{SR}_2}^{(s_2)} = (1-\alpha)\log_2\left\{1+\gamma_{\mathrm{ST}_2,\mathrm{SR}_2}^{(s_2)}\right\} \tag{11}$$

where $\gamma_{\mathrm{ST}_i,\mathrm{SR}_i}^{(s_i)} = \dfrac{P_T}{2\sigma_n^2}\left\|\left[\boldsymbol{f}_{\mathrm{SR}_i}^{(s_i)}\right]^H \boldsymbol{H}_{\mathrm{ST}_i,\mathrm{SR}_i}\boldsymbol{p}_{\mathrm{ST}_i}^{(s_i)}\right\|^2$.

D. *Time domain adaptation for primary spectral efficiency enhancement*

Based on the above analysis, adaptive time slot partition is employed to eliminate bottleneck effect for the end-to-end primary link. Take an arbitrary single-relay two-hop link, say from PT to r ($r \in \mathcal{R}$) and from r to PR as an example. Among the two hops, the achievable end-to-end spectral efficiency is restricted by the worse one. If the slot partition factor α can be adjusted in terms of the channel condition in each hop, i.e., to meet (12), the bottleneck can be eliminated.

$$\alpha\log_2\left[1+\gamma_{\mathrm{PT},r}^{(s_p)}\right] = (1-\alpha)\log_2\left[1+\gamma_{r,\mathrm{PR}}^{(s_p)}\right] \tag{12}$$

Recall that $P_{\mathrm{PT}}^{(s_p)} = P_T$ and $P_{\mathrm{SR}_1}^{(s_p)} = P_{\mathrm{ST}_2}^{(s_p)} = P_T/2$. α can be easily computed based on (13).

$$\alpha = \frac{\log_2\left[1+\gamma_{r,\mathrm{PR}}^{(s_p)}\right]}{\log_2\left[1+\gamma_{\mathrm{PT},r}^{(s_p)}\right] + \log_2\left[1+\gamma_{r,\mathrm{PR}}^{(s_p)}\right]} = \frac{C_{r,\mathrm{PR}}^{(s_p)}}{C_{\mathrm{PT},r}^{(s_p)} + C_{r,\mathrm{PR}}^{(s_p)}} \tag{13}$$

As for the multi-relay two-hop situation, bottleneck exists either in the worst sub-link in first hop or in the second hop. The bottleneck effect can be expressed as follows:

$$C_{\mathrm{PT},\mathcal{R},\mathrm{PR}}^{(s_p)} = \min\left\{\alpha \min_{r\in\mathcal{R}}\left[C_{\mathrm{PT},r}^{(s_p)}\right], (1-\alpha)C_{\mathcal{R},\mathrm{PR}}^{(s_p)}\right\} \tag{14}$$

where $C_{\mathcal{R},\mathrm{PR}}^{(s_p)} = \log_2[1+\gamma_{\mathcal{R},\mathrm{PR}}^{(s_p)}]$. To maximize $C_{\mathrm{PT},\mathcal{R},\mathrm{PR}}^{(s_p)}$, $\alpha \min_{r\in\mathcal{R}}[C_{\mathrm{PT},r}^{(s_p)}] = (1-\alpha)C_{\mathcal{R},\mathrm{PR}}^{(s_p)}$ should be satisfied. Then α is computed by (15):

$$\alpha = \frac{C_{\mathcal{R},\mathrm{PR}}^{(s_p)}}{\min_{r\in\mathcal{R}}\left[C_{\mathrm{PT},r}^{(s_p)}\right] + C_{\mathcal{R},\mathrm{PR}}^{(s_p)}} \tag{15}$$

In order to achieve as good performance as possible, exhaustive searching can be adopted to select the best strategy from all possible single-relay and two-relay schemes combining with α adjustment.

It should be noted that, by adaptively adjusting α, the bottleneck can be eliminated, however synchronization complexity arises. Similarly, the adjustment of transmit power at relay node(s) can also be employed to counteract bottleneck effect, with which no synchronization complexity will be resulted. For limited space, power adaptation is not elaborated in this paper.

IV. Simulation Results

We now evaluate the performance of proposed mechanism. For conciseness, we use a general expression [*Time partition*, *Relay energy*, *Relay strategy*, *Expected data*] to indicate different cooperation schemes. The first term *time partition* is either 0.5 or α, representing equal time division and adaptive partition of frame duration, respectively. *Relay energy*, denoted by $E_{r|\mathcal{R}}^{(s_p)}$, indicates total energy consumption of relay node(s) used for forwarding primary traffic. Recall that $P_{\mathrm{SR}_1}^{(s_p)}=P_{\mathrm{ST}_2}^{(s_p)}=P_T/2$, when a single relay is adopted, $E_{r|\mathcal{R}}^{(s_p)}=(1-\alpha)TP_T/2$, whereas for the two-relay situation, $E_{r|\mathcal{R}}^{(s_p)}=(1-\alpha)TP_T$. The term *relay strategy* is chose from set $\{\mathrm{SR}_1,\mathrm{ST}_2,\mathrm{SR}_1\,\&\,\mathrm{ST}_2, Adpt\}$, in which SR_1 and ST_2 indicate single-relay cooperation, $\mathrm{SR}_1\,\&\,\mathrm{ST}_2$ denotes two-relay situation and the abbreviation *Adpt* represents for the optimal selection of cooperation schemes using exhaustive searching. *Expected data*, selected from $\{s_p,s_1,s_2\}$, denotes the desired data transmission. It should be noted that in the following figures, the x-axis, SNR is based on the transmit power at PT normalized by noise power, hence for relay links the actual SNR for both primary and secondary data is the coordinate value multiplied by 1/2 due to the fact that $P_{\mathrm{SR}_1}^{(s_p)}=P_{\mathrm{ST}_2}^{(s_p)}=P_{\mathrm{ST}_1}^{(s_1)}=P_{\mathrm{ST}_2}^{(s_2)}=P_T/2$.

Fig. 2 shows the spectral efficiency of primary transmission with fixed time division. It can be seen that the adaptive relay selection, yielding variable energy consumption, denoted by $[0.5, TP_{r|\mathcal{R}}^{(s_p)}/2, Adpt, s_p]$ achieves the best spectral efficiency. Among two single-relay schemes, employing SR_1 outperforms adopting ST_2 due to the fact that based on the filtering algorithms in Section III.A, we have $C_{\mathrm{PT,SR}_1}^{(s_p)}>C_{\mathrm{PT,ST}_2}^{(s_p)}$, and in addition $C_{\mathrm{SR}_1,\mathrm{PR}}^{(s_p)}=C_{\mathrm{ST}_2,\mathrm{PR}}^{(s_p)}$ statistically. As for two-relay situation, its spectral efficiency is inferior to that with single relay SR_1. This is because the achievable spectral efficiency in the first hop is limited by $\min_{r\in\mathcal{R}}[C_{\mathrm{PT},r}^{(s_p)}]$, or specifically speaking, by $C_{\mathrm{PT,ST}_2}^{(s_p)}$ in a statistical sense. Whereas in the second hop, PR receives two copies of s_p from SR_1 and ST_2, and combines them to decode, resulting in enhanced received SNR. Consequently, the probability that bottleneck exists in the first hop is obviously higher than that in the second hop. As a result, the spectral efficiency with two relays is inferior to that with single relay SR_1.

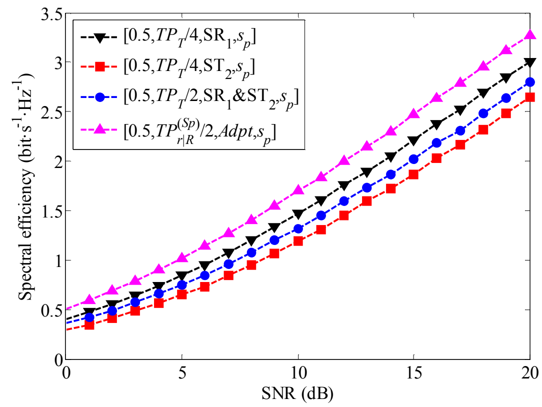

Fig. 2. Spectral efficiency of the primary without α adaptation.

In Fig. 3, the spectral efficiency of the primary with adaptable time division factor is plotted. Similar to Fig. 2, the optimal relay selection, i.e., $[\alpha,(1-\alpha)TP_{r|\mathcal{R}}^{(s_p)}/2, Adpt, s_p]$ achieves the maximum spectral efficiency. As for two single-relay mechanisms, the one employing SR_1 outperforms that using ST_2 due to different filtering processing in phase one. The spectral efficiency of two-relay strategy exceeds that with single relay ST_2 statistically, but inferior to that with SR_1, due to the bottleneck effect. Provided with same cooperation scheme, bottleneck can be eliminated by adjusting α, thus the achievable spectral efficiency is improved compared to that in Fig. 2.

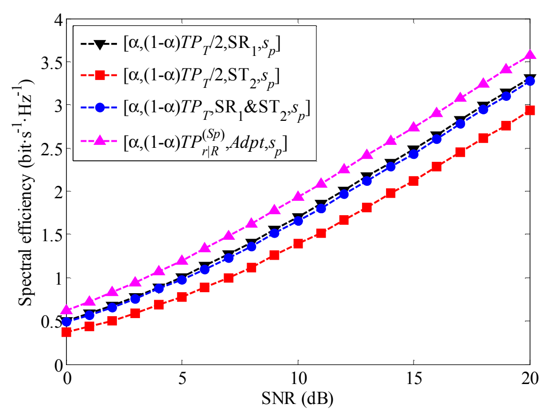

Fig. 3. Spectral efficiency of the primary with adaptable α.

In Fig. 4, the spectral efficiency of the secondary with different strategies is illustrated. When fixed time partition is employed, spectral efficiency of the secondary with single-relay scheme is the same with that adopting two relays. Thus we use $\{r, \mathrm{SR}_1\,\&\,\mathrm{ST}_2\}$ to indicate both cases. Since exhaustive searching may choose single relay to forward primary information, only the SU being recruited as relay can transmit or be transmitted to. Due to this fact, spectral efficiency with adaptive relay selection is not plotted.

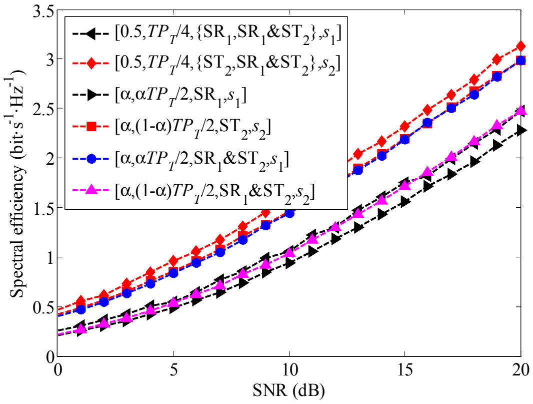

Fig. 4. Spectral efficiency of the secondary.

As the precoder and filter design for s_1 at SR_1 based on (2) results in more power loss of desired signal compared to that based on (7). Given $\alpha = 0.5$, the spectral efficiency of s_2's transmission always exceeds that of s_1. When adaptable α is employed, the end-to-end transmission is restricted by $\min_{r\in\mathcal{R}}[C_{\mathrm{PT},r}^{(s_p)}]$ with two-relay strategy, thus a large α is needed to eliminate the bottleneck. The secondary transmission from ST_1 to SR_1 is benefited from large α. However, the data volume transmitted by ST_2 reduces, yielding decreasing spectral efficiency. Among two single-relay mechanisms, since the first hop for the primary information delivery with ST_2 is statistically inferior to that with SR_1, the bottleneck elimination in the former case results in larger α, spectral efficiency of $(\mathrm{ST}_2, \mathrm{SR}_2)$ pair exceeds that of $(\mathrm{ST}_1, \mathrm{SR}_1)$. As for two-relay strategy, an even larger α is needed for bottleneck elimination compared with the single-relay situation, yielding the spectral efficiency of s_1 exceeds that of s_2.

V. Conclusion

An adaptive cooperative communication mechanism is proposed in MIMO-CCRN to achieve spatiotemporal spectrum sharing. By recruiting SU(s) as relay(s), end-to-end primary transmission is improved. In addition, SU(s) can access to the spectrum as a reward for affording help to the primary, coexisting with the primary by exploiting spatial multiplexing. With the proposed scheme, multiple SU pairs participate in relaying primary traffic, and share the spectrum access authority based on a TDMA mode. Furthermore, a time slot division adaptation is employed so as to eliminate bottleneck effect in the end-to-end primary link. Simulation results show that spectral efficiency of both the primary and secondary can be effectively enhanced.

Acknowledgment

This work was supported in part by NSFC (61102057, 61231008); National S&T Major Project (2012ZX03003005-005); the 111 Project (B08038); PCSIRT (IRT0852); the Key Project of Chinese Ministry of Education (107103); the Fundamental Research Funds for the Central Universities (K5051301014); and ISN Project (ISN1103005).

References

[1] U. S. Federal Communications Commission, "Notice of proposed rule making and order, FCC 03-322, in the matter of facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," FCC-03-322A1, 2003.

[2] J. Mitola, G. Q. J. Maguire, "Cognitive radios: making software radios more personal," IEEE Personal Communications, vol. 6, no. 4, pp. 13-18, 1999.

[3] A. Goldsmith, S. A. Jafar, et al., "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," Proceedings of the IEEE, vol. 97, no. 5, pp. 894-914, 2009.

[4] K. G. Nagananda, P. Mohapatra, et al., "Multiuser cognitive radio networks: an information-theoretic perspective," Int. Journal of Advances in Engineering Sciences and Applied Mathematics, vol. 5, no. 1, pp. 43-65, 2013.

[5] B. Cao, Q. Zhang, et al., "Toward efficient radio spectrum utilization: user cooperation in cognitive radio networking," IEEE Network, vol. 26, no. 4, pp. 46-52, Jul. 2012.

[6] O. Simeone, I. Stanojev, et al., "Spectrum leasing to cooperating secondary ad hoc networks," IEEE J. Sel. Areas in Commun., vol. 26, no. 1, pp. 203-213, 2008.

[7] J. Zhang and Q. Zhang, "Stackelberg game for utility-based cooperative cognitive radio networks," in Proceedings of ACM Mobihoc, pp. 23-31, 2009.

[8] H. Wang, L. Gao, et al., "Cooperative spectrum sharing in cognitive radio networks: a game-theoretic approach," in Proceedings of IEEE Int. Conf. on Commun. (ICC), pp. 1-5, 2010.

[9] W. Su, J. D. Matyjas, et al., "Active cooperation between primary users and cognitive radio users in cognitive ad-hoc networks," IEEE Trans. Sig. Process., vol. 60, no. 4, pp. 1796-1805, 2012.

[10] S. Hua, H. Liu, et al., "Exploiting MIMO antennas in cooperative cognitive radio networks," in Proceedings of IEEE Int. Conf. on Computer Commun. (INFOCOM), pp. 2714-2722, 2011.

[11] B. Cao, L. X. Cai, et al., "Cooperative cognitive radio networking using quadrature signaling," in Proceedings of IEEE Int. Conf. on Computer Commun. (INFOCOM), pp. 3096-3100, 2012.

[12] 3GPP TS 36.101 V11.4.0, "User equipment (UE) radio transmission and reception," 2013.

[13] J. H. Noh, S. J. Oh, "Cognitive radio channel with cooperative multi-antenna secondary systems," IEEE J. Sel. Areas in Commun., vol. 32, no. 3, pp. 1-11, 2014.

[14] R. Manna, R. H. Y. Louie, et al., "Cooperative spectrum sharing in cognitive radio networks with multiple antennas," IEEE Trans. Sig. Process., vol. 59, no. 11, pp. 5509-5522, 2011.

[15] A. S. Ibrahim, A. K. Sadek, et al., "Cooperative communications with relay-selection when to cooperate and whom to cooperate with," IEEE Trans. Wireless Commun., vol. 7, no. 7, pp. 2814-2827, 2008.

[16] N. Devroye, P. Mitran, V. Tarokh, "Achievable rates in cognitive radio channels," IEEE Trans. Inf. Theory, vol. 52, no. 5, pp. 1813-1827, 2006.

[17] Asaduzzaman, Y. K. Hyung, K. Insoo, "Opportunistic Relaying Based Spectrum Leasing for Cognitive Radio Networks," Journal of Commun. and Networks, vol. 13, no. 1, pp. 50-55, 2011.