Achievable Rates for Cognitive Radios Opportunistically Permitting Excessive Secondary-to-Primary Interference

Woohyuk Chang, *Student Member, IEEE*, Sae-Young Chung, *Senior Member, IEEE*, and Yong H. Lee, *Senior Member, IEEE*

***Abstract*—We propose a cognitive radio system in fading environments where the secondary sender opportunistically violates the secondary-to-primary (S-P) interference power limit. Assuming a slowly-varying fading interference channel with two-senders and two-receivers, the primary sender adjusts its rate and power depending on its own channel, while ignoring the secondary user. Through an optimization process, the secondary sender's rate and power are determined and the decision is made on whether the secondary sender violates the S-P interference power limit. The primary receiver removes the effect of excessive S-P interference by multiuser decoding (MUD) which jointly decodes the primary and secondary senders' data. Achievable rates of the proposed system are examined by computer simulation. It was observed that the proposed system tends to violate the S-P interference power limit when the S-P channel gain is larger than the other channel gains. Remarkably, the proposed system can provide a considerably high data rate for the secondary user even without sacrificing the data rate for the primary user over a wide range of signal-to-noise ratios.**

***Index Terms*—Cognitive radio, spectrum sharing, interference channel, interference-temperature, power control, water-filling.**

I. Introduction

COGNITIVE radios (CRs) can considerably improve spectral efficiency by allowing secondary users to share a frequency band with primary users [1]–[3]. In a shared band, secondary users are allowed if either the channel is unused or interference to the primary receivers, called secondary-to-primary (S-P) interference, can be maintained below a certain threshold. This type of opportunistic channel allocation requires CR systems to be equipped with techniques for accurate spectrum sensing, dynamic frequency selection, and power/rate control in a distributed environment. Various techniques for opportunistic channel access in CR systems have been introduced, and some recent results can be found in the special issues on CRs [4], [5].

In this paper, we propose a CR system that opportunistically permits excessive S-P interference, violating the S-P interference power limit. Our research shows that under certain channel conditions, excessive S-P interference is as harmless as no interference for the primary user and can increase the secondary user's data rate. Assuming an interference channel model with two-senders and two-receivers (Fig. 1), which has been considered in [6] and [7] for deriving an achievable rate region of a CR system, performance degradation of the primary user caused by excessive S-P interference can be avoided as follows. The primary sender $(\mathcal{S}_1)$ controls its transmission rate and power depending on its own channel (h_{11}) while ignoring the secondary user. Whenever the secondary sender $(\mathcal{S}_2)$ causes excessive S-P interference, the primary receiver $(\mathcal{D}_1)$ performs multiuser decoding (MUD) and $\mathcal{S}_2$ adjusts its transmission rate and power so that $\mathcal{D}_1$ can reliably decode data from both the primary and secondary senders ($\mathcal{S}_1$ and $\mathcal{S}_2$) via MUD. MUD at $\mathcal{D}_1$ is shown not to require any additional control over $\mathcal{S}_1$'s behavior. Through the cooperation between $\mathcal{S}_2$ and $\mathcal{D}_1$ for MUD, $\mathcal{S}_2$'s transmission rate can be increased without degrading the primary user's performance.[1]

One scenario corresponding to the proposed CR system is the uplink of a cellular communication system, where the primary pair $\mathcal{S}_1$ and $\mathcal{D}_1$ stand for a mobile station (MS) and a base station (BS), respectively. In this case, the BS $(\mathcal{D}_1)$ controls the secondary pair as well as $\mathcal{S}_1$ and performs MUD whenever necessary. By providing this extra service, the BS can increase the cellular system capacity.

The interference channel in Fig. 1 consists of two multiple access channels (MACs) between $(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{D}_i$, $i \in \{1, 2\}$ [8]. By simultaneously considering the capacity regions of the two MAC channels, we formulate an optimization problem that maximizes $\mathcal{S}_2$'s average transmission rate under an average transmission power constraint and derive the optimal power/rate control policy for $\mathcal{S}_2$. Achievable rates of the proposed CR system are obtained through computer simulation and compared with those of conventional CR systems that always maintain S-P interference below a threshold. The results indicate that the proposed system can provide a considerably high data rate for the secondary user without sacrificing the primary user's data rate and can outperform conventional CR systems.

The organization of this paper is as follows. In Section II, we introduce the CR system model. In Section III, the

Manuscript received August 12, 2008; revised March 7, 2009 and August 19, 2009; accepted October 29, 2009. The associate editor coordinating the review of this paper and approving it for publication was R. Nabar.

This work was supported by the IT R&D program of MKE/KEIT [2008-F-004-02].

The authors are with the Department of Electrical Engineering, KAIST, Daejeon 305-701, Republic of Korea (e-mail: whchang@stein.kaist.ac.kr; {sychung, yohlee}@ee.kaist.ac.kr).

Digital Object Identifier 10.1109/TWC.2010.02.081083

[1]Because the primary receiver contributes to an increase of the system capacity, the communications service provider may offer rewards in the form of discounts for the extra service. Such a reward would motivate primary receivers to help secondary communications.

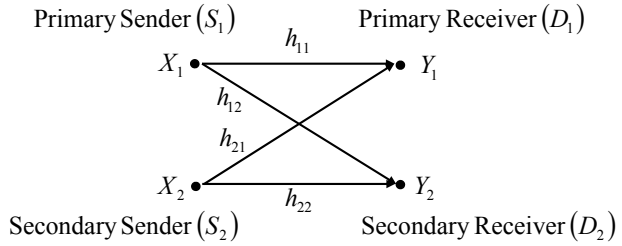

Fig. 1. A slowly varying fading interference channel.

optimization problem for maximizing $\mathcal{S}_2$'s transmission rate under an average power constraint is formulated and solved using the method of Lagrange multipliers. This section also presents an operation scenario for the proposed CR system. In Section IV, average achievable rates of the proposed system are obtained through computer simulation and are compared with those of conventional CR systems. Finally, Section V presents the conclusion.

II. System Model

We consider the interference channel shown in Fig. 1, where the primary user $\mathcal{S}_1$ and secondary user $\mathcal{S}_2$ transmit complex-valued codewords $\underline{x}_1 = (x_{11}, \cdots, x_{1n})$ and $\underline{x}_2 = (x_{21}, \cdots, x_{2n})$ with powers P_1 and P_2, respectively. Assuming slowly varying flat fading channels whose gains h_{11}, h_{12}, h_{21}, and h_{22} are fixed during a coding block, the signals received by the primary and secondary receivers ($\mathcal{D}_1$ and $\mathcal{D}_2$) are given by

$$\underline{y}_1 = h_{11}\underline{x}_1 + h_{21}\underline{x}_2 + \underline{z}_1, \tag{1}$$

$$\underline{y}_2 = h_{22}\underline{x}_2 + h_{12}\underline{x}_1 + \underline{z}_2, \tag{2}$$

where h_{ij} are complex random variables, $\underline{z}_1$ and $\underline{z}_2$ are noise vectors consisting of n zero-mean, independently identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables with variances N_1 and N_2, respectively. To simplify notations, we define $g_{ij} \triangleq |h_{ij}|^2$ for $i, j \in \{1, 2\}$ and $\underline{g} \triangleq (g_{11}, g_{12}, g_{21}, g_{22})$. A BS controlling the terminals $(\mathcal{S}_i, \mathcal{D}_i)$, $i \in \{1, 2\}$ is assumed to exist. The BS collects the channel state information (CSI), $\underline{g}$, and their probability density functions (pdfs), and then it optimizes some design parameters for the terminals. The CSI $\underline{g}$ and the design parameters are delivered to the terminals. In the case where the proposed CR system is applied to the uplink of a cellular communication system, the BS becomes $\mathcal{D}_1$ and the overhead for collecting the CSI and delivering the channel/design parameters to the terminals can be considerably reduced.

The transmitters ($\mathcal{S}_1$ and $\mathcal{S}_2$) and receivers ($\mathcal{D}_1$ and $\mathcal{D}_2$) operate as follows:

- $\mathcal{S}_1$ transmits regardless of $\mathcal{S}_2$'s instantaneous behavior, but allows some interference from $\mathcal{S}_2$ that is limited by a certain threshold Q, referred to as the S-P interference power limit. Its average transmission rate is maximized as

$$\bar{R}_1^* = \max_{P_1 \geq 0} \mathbb{E}_{g_{11}} \left[\log_2 \left(1 + \frac{g_{11} P_1}{N_1 + Q} \right) \right] \quad \text{[bits/sec/Hz]} \tag{3a}$$

subject to

$$\mathbb{E}_{g_{11}} [P_1] \leq \bar{P}_1, \tag{3b}$$

where P_1 is the instantaneous transmission power of $\mathcal{S}_1$ and $\bar{P}_1$ denotes the average transmission power limit. The optimal power that achieves the maximum rate in (3a) can be found by water-filling [9]:

$$P_1^* = \left[\frac{1}{\mu \ln 2} - \frac{N_1 + Q}{g_{11}} \right]^+, \tag{4}$$

where μ is chosen to satisfy (3b) and $[x]^+ = \max\{x, 0\}$. Given g_{11}, $\mathcal{S}_1$ transmits with rate

$$R_1^* = C \left(\frac{g_{11} P_1^*}{N_1 + Q} \right), \tag{5}$$

where $C(x) = \log_2(1 + x)$ bits/sec/Hz. Note that $\mathcal{S}_1$ neglects (g_{12}, g_{21}, g_{22}). This can be fully justified when $Q = 0$, because, in this case, the primary user can achieve the channel capacity if g_{11} is given. We suggest that Q be sufficiently small so that the capacity loss of the primary user will be almost negligible.[2]
- $\mathcal{S}_2$ opportunistically violates the S-P interference power limit. Whether $\mathcal{S}_2$ violates the limit or not is determined via an optimization process for maximizing the average data rate of $\mathcal{S}_2$ under an average transmit power constraint. The optimization process also determines $\mathcal{S}_2$'s transmission rate and power.
- $\mathcal{D}_1$ jointly decodes data from $\mathcal{S}_1$ and $\mathcal{S}_2$ when $\mathcal{S}_2$ violates the S-P interference power limit ($g_{21} P_2 > Q$); otherwise, $\mathcal{D}_1$ decodes only $\mathcal{S}_1$'s data while considering the signal from $\mathcal{S}_2$ as interference.
- $\mathcal{D}_2$ jointly decodes data from $\mathcal{S}_1$ and $\mathcal{S}_2$ if both R_1^* in (5) and $\mathcal{S}_2$'s data rate lie inside the capacity region of the MAC channel between $(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{D}_2$; otherwise, $\mathcal{D}_2$ decodes only $\mathcal{S}_2$'s data while considering the signal from $\mathcal{S}_1$ as interference.

III. Optimal Control for Rate, Power, and Decoding Modes

Our objective is to maximize the average transmission rate of $\mathcal{S}_2$ under the constraint of average transmission power. To this end, we derive the maximum achievable rates of $\mathcal{S}_2$ for all possible $\underline{g}$ vectors and examine conditions for P_2 to maximize the transmission rate while satisfying the S-P interference power limit. The derivation is started by examining the capacity regions of the two MACs between $(\mathcal{S}_1, \mathcal{S}_2)$ and $(\mathcal{D}_i)$, $i \in \{1, 2\}$.

Consider the channels between $(\mathcal{S}_1, \mathcal{S}_2)$ and $(\mathcal{D}_1, \mathcal{D}_2)$. Assuming that each $\mathcal{D}_i$ decodes data from both $\mathcal{S}_1$ and $\mathcal{S}_2$ (in other words, each $\mathcal{D}_i$ performs MUD), the capacity region of the MAC between $(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{D}_i$ is given by all $\left(R_1^{(i)}, R_2^{(i)}\right)$

[2]When $Q > 0$, it would be possible to increase the primary user's rate through cooperation between the primary and secondary senders. However, this type of cooperative communication is outside the scope of this paper and we focus on the case where $Q \simeq 0$. In Section IV, it is shown that the proposed system can provide a substantial data rate for the secondary user even when $Q = 0$.

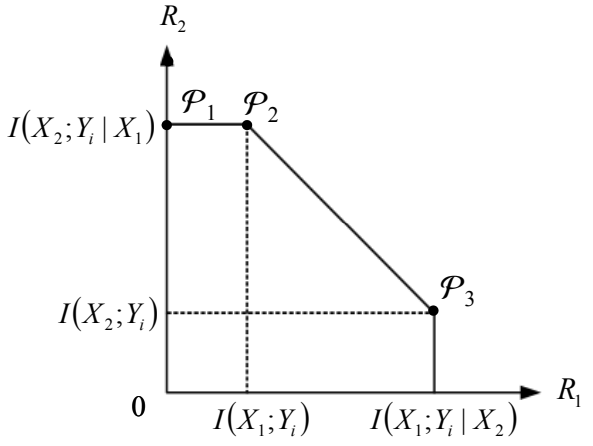

Fig. 2. Capacity region of the multiple access channel from $(\mathcal{S}_1, \mathcal{S}_2)$ to $\mathcal{D}_i$, $i = 1, 2$.

values satisfying

$$R_1^{(i)} \le I(X_1; Y_i|X_2) = C\left(\frac{g_{1i}P_1}{N_i}\right), \quad (6)$$

$$R_2^{(i)} \le I(X_2; Y_i|X_1) = C\left(\frac{g_{2i}P_2}{N_i}\right), \quad (7)$$

$$R_1^{(i)} + R_2^{(i)} \le I(X_1, X_2; Y_i) = C\left(\frac{g_{1i}P_1 + g_{2i}P_2}{N_i}\right), \quad (8)$$

where random variables X_i and Y_i denote the channel input and output, respectively, and $I(\cdot;\cdot)$ denotes the mutual information (Fig. 2) [8]. On the other hand, if $\mathcal{D}_i$ only decodes $\mathcal{S}_i$'s data while treating the signal from the other user as interference ($\mathcal{D}_i$ performs single user decoding (SUD)), the maximum achievable rate is given by

$$R_i^{(i)} \le I(X_i; Y_i) = C\left(\frac{g_{ii}P_i}{N_i + g_{ji}P_j}\right), \quad (9)$$

where $j = 3 - i$. In the CR system under consideration, the transmission power and rate of $\mathcal{S}_1$ are fixed at P_1^* in (4) and R_1^* in (5), respectively, once $\underline{g}$ is given. Thus, the maximum achievable rate of the secondary user's transmission rate, R_2, for a given $\underline{g}$ can be obtained by using P_1^* and R_1^* in (6)–(8). Next, the maximum achievable rates are derived by considering the two MACs between $(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{D}_i$, $i \in \{1, 2\}$. The case for $\mathcal{D}_2$ is described first, and then the case for $\mathcal{D}_1$ follows.

A. Maximum Achievable Rates for the Channel between $(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{D}_2$

Consider the MAC between $(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{D}_2$. Referring to (4) and Fig. 2, we define the following sets of $\underline{g}$: $\mathcal{G}_0 \triangleq \{\underline{g} : P_1^* = 0\}$, $\mathcal{G}_1 \triangleq \{\underline{g} : P_1^* > 0 \text{ and } R_1^* > I(X_1; Y_2|X_2)\}$, and $\mathcal{G}_2 \triangleq \{\underline{g} : P_1^* > 0 \text{ and } R_1^* \le I(X_1; Y_2|X_2)\}$. From (4), $P_1^* = 0$ if $g_{11} < \mu(N_1 + Q)\ln 2$. When $P_1^* = 0$ $(\underline{g} \in \mathcal{G}_0)$, the primary user $(\mathcal{S}_1, \mathcal{D}_1)$ is inactive and the secondary user can transmit with its maximum available power. The secondary user's transmission rate is bounded as

$$R_2 \le C\left(\frac{g_{22}P_2}{N_2}\right) \triangleq R_{2,0}(P_2), \quad (10)$$

where $R_{2,0}(\cdot)$ denotes the upper bound of R_2 when $\underline{g} \in \mathcal{G}_0$. When $P_1^* > 0$, the set of $\underline{g}$'s are partitioned into two sets $\mathcal{G}_1$ and $\mathcal{G}_2$ depending on $I(X_1; Y_2|X_2) = C\left(\frac{g_{12}P_1^*}{N_2}\right)$ which is a constant for a given P_1^*. The channel vectors in $\mathcal{G}_2$ must be partitioned further depending on $I(X_1; Y_2) = C\left(\frac{g_{12}P_1^*}{N_2 + g_{22}P_2}\right)$ (see Fig. 2). However, this partitioning cannot be fixed because $I(X_1; Y_2)$ is a function of P_2, which is to be determined. To consider $I(X_1; Y_2)$, we define an indicator function J which is equal to one if $I(X_1; Y_2) < R_1^*$, and zero otherwise. It can be seen that $R_1^* \gtrless I(X_1; Y_2)$ is equivalent to

$$P_2 \gtrless \gamma, \quad (11)$$

where $\gamma = \frac{1}{g_{22}}\left(\frac{g_{12}(N_1+Q)}{g_{11}} - N_2\right)$. Thus, $J = 1$ if $P_2 > \gamma$, and 0 otherwise. In our CR system, J is a design parameter to be optimized.

The cases for $P_1^* > 0$ are summarized as follows:

- If $\underline{g} \in \mathcal{G}_1$, then $\mathcal{D}_2$ cannot decode $\mathcal{S}_1$'s data and R_2 is bounded by (9):
$$R_2 \le C\left(\frac{g_{22}P_2}{N_2 + g_{12}P_1^*}\right) \triangleq R_{2,1}(P_2), \quad (12)$$
where $R_{2,1}(\cdot)$ denotes the upper bound of R_2 when $\underline{g} \in \mathcal{G}_1$.
- If $\underline{g} \in \mathcal{G}_2$ and $J = 1$, then $\mathcal{D}_2$ performs MUD and R_2 is maximized when (R_1^*, R_2) lies on the line connecting $\mathcal{P}_2$ and $\mathcal{P}_3$ in Fig. 2. In this case, R_2 is bounded by (8):
$$R_2 \le C\left(\frac{g_{12}P_1^* + g_{22}P_2}{N_2}\right) - R_1^* \triangleq R_{2,2,1}(P_2), \quad (13)$$
where $R_{2,2,1}(\cdot)$ denotes the upper bound of R_2 when $\underline{g} \in \mathcal{G}_2$ and $J = 1$.
- If $\underline{g} \in \mathcal{G}_2$ and $J = 0$, then $\mathcal{D}_2$ performs MUD and R_2 is maximized when (R_1^*, R_2) lies on the line connecting $\mathcal{P}_1$ and $\mathcal{P}_2$ in Fig. 2. In this case R_2 is bounded by (7):
$$R_2 \le C\left(\frac{g_{22}P_2}{N_2}\right) \triangleq R_{2,2,0}(P_2), \quad (14)$$
where $R_{2,2,0}(\cdot)$ denotes the upper bound of R_2 when $\underline{g} \in \mathcal{G}_2$ and $J = 0$.
- $\mathcal{D}_2$ performs MUD if and only if $\underline{g} \in \mathcal{G}_2$.

From (13) and (14), it can be shown that $R_{2,2,1}(\gamma + \Delta) > R_{2,2,0}(\gamma)$ for any $\Delta > 0$. Furthermore, as shown in Fig. 3, the capacity region for $J = 1$ encompasses that for $J = 0$. This occurs because the case with $J = 1$ $(P_2 > \gamma)$ requires larger transmission power P_2 than the case with $J = 0$ $(P_2 \le \gamma)$. The values of J and P_2 will be determined through an optimization process maximizing $\mathcal{S}_2$'s average transmission rate under an average transmission power constraint.

B. Maximum Achievable Rates for the Channel between $(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{D}_1$

We now examine how R_2 and P_2 are bounded by the behavior of $\mathcal{D}_1$ which operates depending on the condition

$$g_{21}P_2 \gtrless Q. \quad (15)$$

As in (11), this condition is a function of P_2, which is to be determined, and thus we define another indicator function K

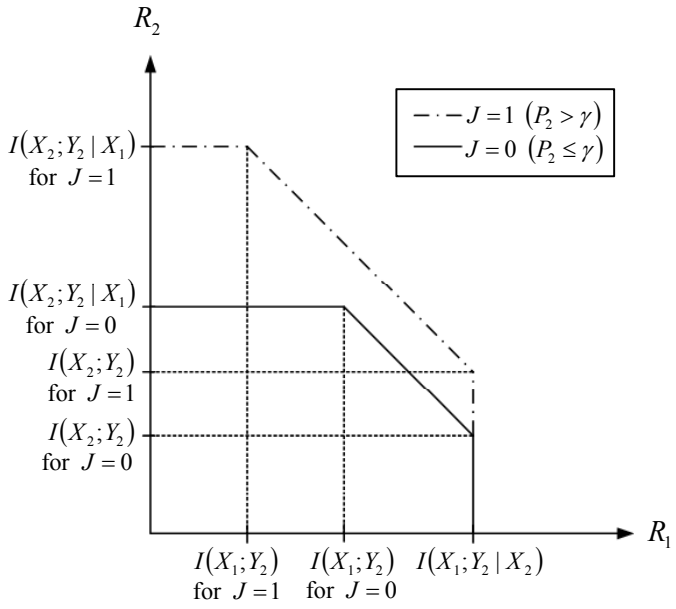

Fig. 3. Capacity regions for $J=1$ and $J=0$ when the channel is the MAC between $(\mathcal{S}_1,\mathcal{S}_2)$ and $\mathcal{D}_2$.

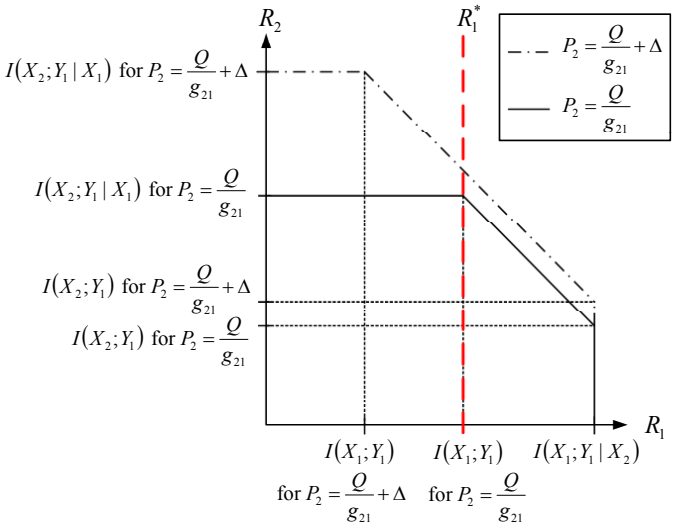

Fig. 4. Capacity regions for $P_2=\frac{Q}{g_{21}}$ and $P_2=\frac{Q}{g_{21}}+\Delta$ when the channel is the MAC between $(\mathcal{S}_1,\mathcal{S}_2)$ and $\mathcal{D}_1$. Here, $Q>0$ and $\Delta>0$.

which is equal to one if $g_{21}P_2>Q$, and zero otherwise. When $K=1$, $\mathcal{S}_2$ violates the S-P interference power limit, resulting in $g_{21}P_2>Q$, and $\mathcal{D}_1$ is asked to perform MUD. It should be noted that $R_1^*=C\left(\frac{g_{11}P_1^*}{N_1+Q}\right)\le C\left(\frac{g_{11}P_1^*}{N_1}\right)=I(X_1;Y_1|X_2)$, and thus R_1^* always lies inside the capacity region of the MAC between $(\mathcal{S}_1,\mathcal{S}_2)$ and $\mathcal{D}_1$. Therefore, setting K at one does not require any additional control on $\mathcal{S}_1$'s behavior, and K can be set at one whenever R_2 lies in the capacity region. Like the indicator J, K is a parameter to be determined optimally. According to (15), the following two cases are considered:

- If $K=0$ $(g_{21}P_2\le Q)$ and $P_1^*>0$, then $\mathcal{D}_1$ performs SUD while ignoring the secondary user. Since R_1^* can take any positive value, the maximum achievable rates of R_2 are given by (12)–(14).
- If $K=1$ $(g_{21}P_2>Q)$ and $P_1^*>0$, then $\mathcal{D}_1$ performs MUD. In this case $I(X_1;Y_1)=C\left(\frac{g_{11}P_1^*}{N_1+g_{21}P_2}\right)<R_1^*$. Therefore, $I(X_1;Y_1)<R_1^*\le I(X_1;Y_1|X_2)$ and R_2 is maximized when (R_1^*,R_2) lies on the line connecting $\mathcal{P}_2$ and $\mathcal{P}_3$ in Fig. 2 and is bounded by (8):

$$R_2\le C\left(\frac{g_{11}P_1^*+g_{21}P_2}{N_1}\right)-R_1^*\triangleq R_{2,3}(P_2), \quad (16)$$

  where $R_{2,3}(\cdot)$ denotes the upper bound of R_2 when $K=1$ and $P_1^*>0$.

It is not recommended to perform MUD at $\mathcal{D}_1$ when the interference power limit is not violated $(g_{21}P_2\le Q)$. This is because performing MUD at $\mathcal{D}_1$ requires R_2 to lie in the capacity region, indicating that R_2 should also be upper bounded by $I(X_2;Y_1|X_1)=C\left(\frac{g_{21}P_2}{N_1}\right)$ in addition to the upper bounds in (12)–(14), and the performance of MUD can be worse than that of SUD. The reason why violating the S-P interference power limit results in performing MUD at $\mathcal{D}_1$ can be more explicitly explained as follows. When $P_2=\frac{Q}{g_{21}}$, $R_1^*=I(X_1;Y_1)$ and $\mathcal{S}_1$'s data can be still successfully decoded by performing SUD at $\mathcal{D}_1$ (see (9)). However, as shown in Fig. 4, if P_2 is slightly increased by $\Delta>0$, then $I(X_1;Y_1)$ decreases, resulting in $I(X_1;Y_1)<R_1^*\le I(X_1;Y_1|X_2)$, and hence MUD should be performed to successfully decode $\mathcal{S}_1$'s data.

These results indicate that the maximum achievable rates of R_2 are given by (10) and (12)–(14) unless both $K=1$ and $P_1^*>0$ are met. When $K=1$ and $P_1^*>0$, the maximum achievable rate is written as

$$R_2\le \begin{cases}\min\{R_{2,1}(P_2),R_{2,3}(P_2)\} & \text{for } \underline{g}\in\mathcal{G}_1\\ \min\{R_{2,2,J}(P_2),R_{2,3}(P_2)\} & \text{for } \underline{g}\in\mathcal{G}_2 \text{ and } J\in\{0,1\}.\end{cases} \quad (17)$$

The rates in the right-hand-side (RHS) of (17) are increasing functions of P_2, as shown in Fig. 5 for $\underline{g}\in\mathcal{G}_1$. From (12)–(14) and (16), it can be seen that $R_{2,1}(0)=R_{2,2,0}(0)=0$ and $R_{2,3}(0)\ge 0$ $\left(R_{2,2,1}(0)\right.$ can take any real value$\left.\right)$. Let $R_{2,1}(P_2)$ and $R_{2,3}(P_2)$ intersect at $P_2=P_{C,1}$, and $R_{2,2,J}(P_2)$ and $R_{2,3}(P_2)$ intersect at $P_2=P_{C,2,J}$. When $\underline{g}\in\mathcal{G}_1$, due to the fact that $0=R_{2,1}(0)\le R_{2,3}(0)$, there are only two cases to consider when identifying the minimum of $R_{2,1}(P_2)$ and $R_{2,3}(P_2)$ (Fig. 5). We define two disjoint subsets of $\mathcal{G}_1$ as $\mathcal{G}_{1,1}\triangleq\left\{\underline{g}:\underline{g}\in\mathcal{G}_1 \text{ and } P_{C,1}>0\right\}$ and $\mathcal{G}_{1,2}\triangleq\left\{\underline{g}:\underline{g}\in\mathcal{G}_1 \text{ and } P_{C,1}<0\right\}$. A channel gain which results in $P_{C,1}=0$ will be included in $\mathcal{G}_{1,1}$ if $R_{2,1}(P_2)>R_{2,3}(P_2)$ for $P_2>0$, and in $\mathcal{G}_{1,2}$ if $R_{2,1}(P_2)<R_{2,3}(P_2)$ for $P_2>0$. Similarly, when $\underline{g}\in\mathcal{G}_2$ and $J=0$, it is sufficient to define $\mathcal{G}_{2,0,1}\triangleq\left\{\underline{g}:\underline{g}\in\mathcal{G}_2 \text{ and } P_{C,2,0}>0\right\}$ and $\mathcal{G}_{2,0,2}\triangleq\left\{\underline{g}:\underline{g}\in\mathcal{G}_2 \text{ and } P_{C,2,0}<0\right\}$. When $P_{C,2,0}=0$, $\underline{g}\in\mathcal{G}_{2,0,1}$ if $R_{2,2,0}(P_2)>R_{2,3}(P_2)$ for $P_2>0$, and $\underline{g}\in\mathcal{G}_{2,0,2}$ if $R_{2,2,0}(P_2)<R_{2,3}(P_2)$ for $P_2>0$.

On the other hand, if $\underline{g}\in\mathcal{G}_2$ and $J=1$, then $R_{2,2,1}(0)$ can take any real value and there are four cases to consider. In this case, we define the following four disjoint subsets of $\mathcal{G}_2$: $\mathcal{G}_{2,1,1}\triangleq\left\{\underline{g}:\underline{g}\in\mathcal{G}_2,\ R_{2,2,1}(0)<R_{2,3}(0) \text{ and } P_{C,2,1}>0\right\}$, $\mathcal{G}_{2,1,2}\triangleq\left\{\underline{g}:\underline{g}\in\mathcal{G}_2,\ R_{2,2,1}(0)>R_{2,3}(0) \text{ and } P_{C,2,1}>0\right\}$, $\mathcal{G}_{2,1,3}\triangleq\left\{\underline{g}:\underline{g}\in\mathcal{G}_2,\ R_{2,2,1}(0)<R_{2,3}(0) \text{ and } P_{C,2,1}<0\right\}$,

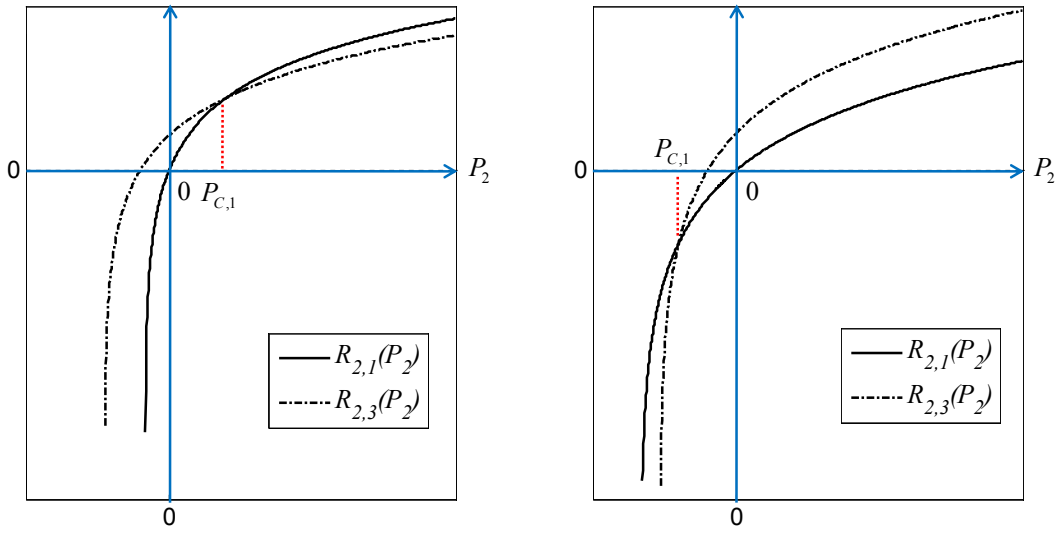

Fig. 5. $R_{2,1}(P_2)$ and $R_{2,3}(P_2)$ against P_2, where the two curves intersect at $P_2 = P_{C,1}$.

and $\mathcal{G}_{2,1,4} \triangleq \{\underline{g} : \underline{g} \in \mathcal{G}_2,\ R_{2,2,1}(0) > R_{2,3}(0)$ and $P_{C,2,1} < 0\}$. A channel gain $\underline{g} \in \mathcal{G}_2$ which results in $P_{C,2,1} = 0$ will be included in $\mathcal{G}_{2,1,3}$ if $R_{2,2,1}(P_2) < R_{2,3}(P_2)$ for $P_2 > 0$, and in $\mathcal{G}_{2,1,4}$ if $R_{2,2,1}(P_2) > R_{2,3}(P_2)$ for $P_2 > 0$. It should be pointed out that the sets for $J = 0$, $\{\mathcal{G}_{2,0,u} : u = 1, 2\}$, and those for $J = 1$, $\{\mathcal{G}_{2,1,v} : v = 1, 2, 3, 4\}$, are *not* disjoint, i.e., $\mathcal{G}_{2,0,u} \cap \mathcal{G}_{2,1,v} \neq \emptyset$. To divide $\mathcal{G}_2$ into disjoint subsets when $K = 1$, we define $\mathcal{E}_{u,v} \triangleq \mathcal{G}_{2,0,u} \cap \mathcal{G}_{2,1,v}$; then $\{\mathcal{E}_{u,v}\}$ are disjoint with each other and $\bigcup_{u,v} \mathcal{E}_{u,v} = \mathcal{G}_2$. For a given $\underline{g} \in \mathcal{E}_{u,v}$, we shall consider both the cases with $J = 0$ and $J = 1$. The minimum value in (17) can be identified by examining whether $\underline{g}$ belongs to either $\mathcal{G}_{1,i}$, $i \in \{1, 2\}$ or $\mathcal{E}_{u,v}$ and by determining $J \in \{0, 1\}$ when $\underline{g} \in \mathcal{E}_{u,v}$.

To proceed further, we also define an additional indicator function L which is equal to zero if $P_2 \leq P_{C,1}$ ($P_2 \leq P_{C,2,J}$), and one otherwise. (When $P_{C,1} = 0$ ($P_{C,2,J} = 0$), L is fixed at one). Using these definitions and notations, the minimum values in (17) are obtained as follows:

- If $\underline{g} \in \mathcal{G}_{1,1}$ and $K = 1$, then $\min\{R_{2,1}(P_2), R_{2,3}(P_2)\} = R_{2,1}(P_2)$ for $L = 0$ and $R_{2,3}(P_2)$ otherwise.
- If $\underline{g} \in \mathcal{G}_{1,2}$ and $K = 1$, then $\min\{R_{2,1}(P_2), R_{2,3}(P_2)\} = R_{2,1}(P_2)$ for all $P_2 \geq 0$.
- If $\underline{g} \in \mathcal{E}_{u,v}$, $J = 0$ and $K = 1$, then $\min\{R_{2,2,0}(P_2), R_{2,3}(P_2)\} = R_{2,2,0}(P_2)$ for $(u, L) \in \{(1, 0), (2, 1)\}$ and $R_{2,3}(P_2)$ for $(u, L) = (1, 1)$.
- If $\underline{g} \in \mathcal{E}_{u,v}$, $J = 1$ and $K = 1$, then $\min\{R_{2,2,1}(P_2), R_{2,3}(P_2)\} = R_{2,2,1}(P_2)$ for $(v, L) \in \{(1, 0), (2, 1), (3, 1)\}$ and $R_{2,3}(P_2)$ for $(v, L) \in \{(1, 1), (2, 0), (4, 1)\}$.

Summarizing these results, the maximum achievable rates of R_2 are selected from $\{R_{2,0}(P_2), R_{2,1}(P_2), R_{2,2,1}(P_2), R_{2,2,0}(P_2), R_{2,3}(P_2)\}$ depending on $\underline{g}$, J, K and L, which result in a total of 16 cases (Table I).

C. Optimization: Problem Formulation and Solution

The average achievable rate for the secondary user is maximized by solving the following optimization problem:

$$\begin{aligned}
\bar{R}_2^* = \max_{\substack{P_2 \geq 0 \\ J,K,L \in \{0,1\}}} \ & \mathbb{E}_{\underline{g} \in \mathcal{G}_0}\left[R_{2,0}(P_2)\right] \\
& + \mathbb{E}_{\underline{g} \in \mathcal{G}_1}\left[R_{2,1}(P_2) \cdot (1 - K)\right] \\
& + \mathbb{E}_{\underline{g} \in \mathcal{G}_2}\left[\{R_{2,2,0}(P_2) \cdot (1 - J) + R_{2,2,1}(P_2) \cdot J\} \cdot (1 - K)\right] \\
& + \mathbb{E}_{\underline{g} \in \mathcal{G}_{1,1}}\left[\{R_{2,1}(P_2) \cdot (1 - L) + R_{2,3}(P_2) \cdot L\} \cdot K\right] \\
& + \mathbb{E}_{\underline{g} \in \mathcal{G}_{1,2}}\left[R_{2,1}(P_2) \cdot L \cdot K\right] \\
& + \sum_{v=1}^{4} \Big\{ \mathbb{E}_{\underline{g} \in \mathcal{E}_{1,v}}[\{R_{2,2,0}(P_2) \cdot (1 - L) \\
& \qquad + R_{2,3}(P_2) \cdot L\} \cdot (1 - J) \cdot K] \\
& \quad + \mathbb{E}_{\underline{g} \in \mathcal{E}_{2,v}}\left[R_{2,2,0}(P_2) \cdot L \cdot (1 - J) \cdot K\right] \Big\} \\
& + \sum_{u=1}^{2} \Big\{ \mathbb{E}_{\underline{g} \in \mathcal{E}_{u,1}}[\{R_{2,2,1}(P_2) \cdot (1 - L) \\
& \qquad + R_{2,3}(P_2) \cdot L\} \cdot J \cdot K] \\
& \quad + \mathbb{E}_{\underline{g} \in \mathcal{E}_{u,2}}\left[\{R_{2,2,1}(P_2) \cdot L + R_{2,3}(P_2) \cdot (1 - L)\} \cdot J \cdot K\right] \\
& \quad + \mathbb{E}_{\underline{g} \in \mathcal{E}_{u,3}}\left[R_{2,2,1}(P_2) \cdot L \cdot J \cdot K\right] \\
& \quad + \mathbb{E}_{\underline{g} \in \mathcal{E}_{u,4}}\left[R_{2,3}(P_2) \cdot L \cdot J \cdot K\right] \Big\} \qquad (18)
\end{aligned}$$

subject to $\mathbb{E}_{\underline{g}}[P_2] \leq \bar{P}_2$; $P_2 \leq \gamma$ for $\underline{g} \in \mathcal{G}_2$ and $J = 0$; $P_2 > \gamma$ for $\underline{g} \in \mathcal{G}_2$ and $J = 1$; $P_2 \leq \frac{Q}{g_{21}}$ for $\underline{g} \in \mathcal{G}_0^c$ and $K = 0$; $P_2 > \frac{Q}{g_{21}}$ for $\underline{g} \in \mathcal{G}_0^c$ and $K = 1$; $P_2 \leq P_{C,1}$ for $\underline{g} \in \mathcal{G}_1$, $K = 1$ and $L = 0$; $P_2 > P_{C,1}$ for $\underline{g} \in \mathcal{G}_1$, $K = 1$ and $L = 1$; $P_2 \leq P_{C,2,J}$ for $\underline{g} \in \mathcal{G}_2$, $K = 1$ and $L = 0$; $P_2 > P_{C,2,J}$ for $\underline{g} \in \mathcal{G}_2$, $K = 1$ and $L = 1$.

Here $\mathcal{G}_0^c$ is the complement of $\mathcal{G}_0$ indicating that $P_1^* > 0$. Although not explicitly expressed in (18), it should be noted that P_2, J, K and L are functions of $\underline{g}$, i.e., $P_2(\underline{g})$, $J(\underline{g})$, $K(\underline{g})$ and $L(\underline{g})$. The optimal transmission power P_2^* can be obtained using the method of Lagrange multipliers. The Lagrangian is written by incorporating the 16 terms in (18)

TABLE I

EXPRESSIONS FOR OPTIMAL TRANSMISSION POWER P_2^{LO}, CORRESPONDING RATE R_2^{LO} AND DECODING MODES OF $\mathcal{D}_1/\mathcal{D}_2$ FOR A GIVEN $\underline{g}$ AND ALL POSSIBLE (J,K,L). HERE "–" MEANS "DON'T CARE"; $<x;y>$ IS DEFINED IN (20); $\bar{P}_{C,1}$ AND $P_{C,2,J}$ ARE DEFINED DIRECTLY BELOW (17) (SEE ALSO FIG. 5); γ IS GIVEN IN (11); α_J AND β ARE DEFINED DIRECTLY BELOW (20).

	K	J	u	v	L	P_2^{LO}	R_2^{LO}	$\mathcal{D}_1$	$\mathcal{D}_2$
$\underline{g}\in\mathcal{G}_0$	–	–	–	–	–	α_0	$R_{2,0}(P_2)$	–	SUD
$\underline{g}\in\mathcal{G}_{1,u}\subset\mathcal{G}_1$	0	–	1,2	–	–	$\min\left\{\alpha_1,\frac{Q}{g_{21}}\right\}$	$R_{2,1}(P_2)$	SUD	SUD
	1	–	1	–	0	$\left\langle \min\{\alpha_1,P_{C,1}\};\frac{Q}{g_{21}}\right\rangle$	$R_{2,1}(P_2)$	MUD	SUD
					1	$\left\langle \beta;\max\left\{\frac{Q}{g_{21}},P_{C,1}\right\}\right\rangle$	$R_{2,3}(P_2)$	MUD	SUD
			2	–	1	$\left\langle \alpha_1;\frac{Q}{g_{21}}\right\rangle$	$R_{2,1}(P_2)$	MUD	SUD
$\underline{g}\in\mathcal{E}_{u,v}\subset\mathcal{G}_2$	0	0	1,2	1,2,3,4	–	$\min\left\{\alpha_0,\gamma,\frac{Q}{g_{21}}\right\}$	$R_{2,2,0}(P_2)$	SUD	MUD
		1	1,2	1,2,3,4	–	$\left\langle \min\left\{\alpha_1,\frac{Q}{g_{21}}\right\};\gamma\right\rangle$	$R_{2,2,1}(P_2)$	SUD	MUD
	1	0	1	1,2,3,4	0	$\left\langle \min\{\alpha_0,\gamma,P_{C,2,0}\};\frac{Q}{g_{21}}\right\rangle$	$R_{2,2,0}(P_2)$	MUD	MUD
					1	$\left\langle \min\{\beta,\gamma\};\max\left\{\frac{Q}{g_{21}},P_{C,2,0}\right\}\right\rangle$	$R_{2,3}(P_2)$	MUD	MUD
			2	1,2,3,4	1	$\left\langle \min\{\alpha_0,\gamma\};\frac{Q}{g_{21}}\right\rangle$	$R_{2,2,0}(P_2)$	MUD	MUD
		1	1,2	1	0	$\left\langle \min\{\alpha_1,P_{C,2,1}\};\max\left\{\frac{Q}{g_{21}},\gamma\right\}\right\rangle$	$R_{2,2,1}(P_2)$	MUD	MUD
					1	$\left\langle \beta;\max\left\{\frac{Q}{g_{21}},\gamma,P_{C,2,1}\right\}\right\rangle$	$R_{2,3}(P_2)$	MUD	MUD
				2	0	$\left\langle \min\{\beta,P_{C,2,1}\};\max\left\{\frac{Q}{g_{21}},\gamma\right\}\right\rangle$	$R_{2,3}(P_2)$	MUD	MUD
					1	$\left\langle \alpha_1;\max\left\{\frac{Q}{g_{21}},\gamma,P_{C,2,1}\right\}\right\rangle$	$R_{2,2,1}(P_2)$	MUD	MUD
				3	1	$\left\langle \alpha_1;\max\left\{\frac{Q}{g_{21}},\gamma\right\}\right\rangle$	$R_{2,2,1}(P_2)$	MUD	MUD
				4	1	$\left\langle \beta;\max\left\{\frac{Q}{g_{21}},\gamma\right\}\right\rangle$	$R_{2,3}(P_2)$	MUD	MUD

with the corresponding constraints. The expression for the Lagrangian is lengthy, and describing the overall optimization process is a tedious task. Fortunately, however, the overall optimization problem can be decomposed into 16 subproblems that optimize each term in (18) under $\mathbb{E}_{\underline{g}}[P_2]\le\bar{P}_2$ and the corresponding constraints because of the following facts: the subsets $\{\mathcal{G}_0,\mathcal{G}_{1,u},\mathcal{E}_{u,v}|u\in\{1,2\},v\in\{1,2,3,4\}\}$ are disjoint with each other; the terms associated with the subsets which are not disjoint can be treated separately because they correspond to different (J,K,L) values; and $\mathbb{E}_{\underline{g}}[P_2]\le\bar{P}_2$ is the only constraint that should be considered by all 16 terms in (18). The solution of each subproblem will be given by a function of λ_1, which is the Lagrange multiplier for $\mathbb{E}_{\underline{g}}[P_2]\le\bar{P}_2$. Then the optimal value for λ_1 is obtained by simultaneously considering all 16 cases. The procedure for formulating and solving the subproblems is illustrated in the following example for the fourth term $\mathbb{E}_{\underline{g}\in\mathcal{G}_2}\left[R_{2,2,1}(P_2)\cdot J\cdot(1-K)\right]$ in (18) when $J=1$ and $K=0$.

Example (Subproblem for the fourth term in (18)): The Lagrangian for the fourth term is given by $\mathbb{E}_{\underline{g}\in\mathcal{G}_2}\Big[R_{2,2,1}(P_2)-\lambda_1\left(P_2-\bar{P}_2\right)+\lambda_2(P_2-\gamma)-\lambda_3\left(P_2-\frac{Q}{g_{21}}\right)\Big]$, where λ_1, λ_2, and λ_3 are Lagrange multipliers associated with the constraints $\mathbb{E}_{\underline{g}}[P_2]\le\bar{P}_2$, $P_2>\gamma$ and $P_2\le\frac{Q}{g_{21}}$, respectively. The Karush-Khun-Tucker (KKT) conditions [10] indicate that $\lambda_i\ge0$ for $i\in\{1,2,3\}$, and

$$\frac{\partial R_{2,2,1}(P_2^{LO})}{\partial P_2^{LO}}-\lambda_1+\lambda_2-\lambda_3=0, \tag{19}$$

for $\underline{g}\in\mathcal{G}_2$, $J=1$ and $K=0$, where P_2^{LO} denotes the optimal transmission power for this case and the superscript "LO" stands for locally optimal. Since (19) together with KKT conditions derived from other terms in (18) constitutes the KKT conditions for the overall expression of the Lagrangian of (18), (19) is eventually the necessary condition for the optimal P_2 of (18). An expression for P_2^{LO} can be obtained from (19) as follows. Since $\bar{R}_2^*$ increases with P_2, the constraint $\mathbb{E}_{\underline{g}}[P_2]=\bar{P}_2$ should be met, and thus $\lambda_1>0$. For λ_2, due to the strict inequality $P_2>\gamma$, $\lambda_2=0$. Finally, λ_3 can take either $\lambda_3=0$ or $\lambda_3>0$. When $P_2^{LO}<\frac{Q}{g_{21}}$, λ_3 becomes zero ($\lambda_3=0$), and, in this case, solving (19) for P_2^{LO} after setting λ_2 and λ_3 at zero results in $P_2^{LO}=\alpha_1$ where $\alpha_1\triangleq\left[\frac{1}{\lambda_1\ln 2}-\frac{N_2+g_{12}P_1^*}{g_{22}}\right]^+$. Note that α_1 is eligible for P_2^{LO} only when $\alpha_1<\frac{Q}{g_{21}}$. For the case where $P_2^{LO}=\frac{Q}{g_{21}}$ ($\lambda_3>0$), it is not necessary to solve (19) because P_2^{LO} is fixed at $\frac{Q}{g_{21}}$. Summarizing these results, P_2^{LO} is given by $P_2^{LO}=\min\left\{\alpha_1,\frac{Q}{g_{21}}\right\}$ if $\min\left\{\alpha_1,\frac{Q}{g_{21}}\right\}>\gamma$. When $\min\left\{\alpha_1,\frac{Q}{g_{21}}\right\}\le\gamma$, $P_2^{LO}=\emptyset$ where $\emptyset$ denotes the null set indicating that no solution exists. The expression for P_2^{LO} incorporating all possible cases may be written as

$$P_2^{LO}=\left\langle \min\left\{\alpha_1,\frac{Q}{g_{21}}\right\};\gamma\right\rangle \tag{20}$$

where $\langle x;y\rangle=x$ if $x>y$, and $\emptyset$ otherwise. □

In this manner, each of the 16 Lagrangian terms from (18) is solved to derive the corresponding optimal transmission power. The results are summarized in Table I, where $\alpha_J\triangleq\left[\frac{1}{\lambda_1\ln 2}-\frac{N_2+J\cdot g_{12}P_1^*}{g_{22}}\right]^+$, $\beta\triangleq\left[\frac{1}{\lambda_1\ln 2}-\frac{N_1+g_{11}P_1^*}{g_{21}}\right]^+$, and "–" means "don't care," indicating that it is not necessary to determine the corresponding indicator value. Note in the table that some (J,K,L) values are associated with multiple (u,v) pairs. For example, $(J,K,L)=(0,1,0)$ is associated with $\{(u,v)|u=1,v\in\{1,2,3,4\}\}$. This, in turn, indicates that multiple (J,K,L) values can be considered for a given

(u, v) pair. For example, if $(u, v) = (1, 1)$, then $(J, K, L) \in \{(0, 0, -), (1, 0, -), (0, 1, 0), (0, 1, 1), (1, 1, 0), (1, 1, 1)\}$. This fact is used for the optimization process described below. Each P_2^{LO} is a *nonincreasing* function of λ_1 which is to be determined to satisfy $\mathbb{E}_{\underline{g}}[P_2^*] = \bar{P}_2$. The optimal λ_1, denoted as λ_1^*, is found through a Monte Carlo approach in conjunction with a grid search.[3] The procedure is described in the following steps, where $[\lambda_{1,0}(k), \lambda_{1,S}(k)]$ denotes an interval for a grid search in the k-th iteration, $\lambda_{1,0}(k) < \lambda_{1,S}(k)$, and S is a positive integer.

Optimization Process for Obtaining λ_1^*

Step 1. Generate M sample vectors of $\underline{g}$ following a given distribution. Set $k = 0$ and determine an initial interval $[\lambda_{1,0}(0), \lambda_{1,S}(0)]$ which includes the optimal λ_1^*.

Step 2. In the k-th iteration, obtain $(S-1)$ points $\{\lambda_{1,1}(k), \cdots, \lambda_{1,S-1}(k)\}$ which are equally spaced in between $\lambda_{1,0}(k)$ and $\lambda_{1,S}(k)$ $\big(\lambda_{1,0}(k) < \lambda_{1,1}(k) < \cdots < \lambda_{1,S}(k)\big)$.

Step 3. For each $\lambda_{1,s}(k)$, $0 \le s \le S$, we estimate $\left|\mathbb{E}_{\underline{g}}[P_2^*(\lambda_{1,s}(k))] - \bar{P}_2\right|$ by evaluating $\left|\frac{1}{M}\sum_{\underline{g}} P_2^*(\lambda_{1,s}(k)) - \bar{P}_2\right|$. The procedure for obtaining $P_2^*(\lambda_{1,s}(k))$ is as follows. For each $\underline{g}$, find the subset of $\mathcal{G}$ including $\underline{g}$ from $\{\mathcal{G}_0, \mathcal{G}_{1,u}, \mathcal{E}_{u,v} | u \in \{1, 2\}, v \in \{1, 2, 3, 4\}\}$ and evaluate $P_2^{LO}(\lambda_{1,s}(k))$ and its corresponding transmission rate $R_2^{LO}(\lambda_{1,s}(k))$ for all possible (J, K, L) values associated with the subset. For example, if $\underline{g} \in \mathcal{E}_{1,1}$ then all $P_2^{LO}(\lambda_{1,s}(k))$ and $R_2^{LO}(\lambda_{1,s}(k))$ values associated with $(J, K, L) \in \{(0, 0, -), (1, 0, -), (0, 1, 0), (0, 1, 1), (1, 1, 0), (1, 1, 1)\}$ are calculated. The maximum among $\{R_2^{LO}(\lambda_{1,s}(k))\}$ is the optimal transmission rate $R_2^*(\lambda_{1,s}(k))$ for this $\underline{g}$, and the corresponding transmit power becomes $P_2^*(\lambda_{1,s}(k))$.

Step 4. Select $\lambda_{1,s^*}(k)$, $1 \le s^* \le S-1$, associated with the minimum among $\left\{\left|\frac{1}{M}\sum_{\underline{g}} P_2^*(\lambda_{1,s}(k)) - \bar{P}_2\right|, 1 \le s \le S-1\right\}$ and update the interval: $[\lambda_{1,0}(k+1), \lambda_{1,S}(k+1)] = [\lambda_{1,s^*-1}(k), \lambda_{1,s^*+1}(k)]$.

Step 5. Stop if $|\lambda_{1,S}(k+1) - \lambda_{1,0}(k+1)| < \epsilon$, where ϵ is a small positive number. Otherwise, go to Step 2 and continue for the $(k+1)$-th iteration.

The optimization process is stable in the sense that $\lambda_{1,s^*}(k)$ in Step 4 approaches the optimal λ_1^* as k increases. This is true because $\{P_2^{LO}\}$ in Table I are nondecreasing functions of λ_1. The maximum average rate $\bar{R}_2^*$ in (18) can be estimated by evaluating $\frac{1}{M}\sum_{\underline{g}} R_2^*(\lambda_1^*)$, where $R_2^*(\lambda_1^*)$ for each $\underline{g}$ is obtained in Step 3 above. The computational complexity of the optimization process is $O(MSI)$, where M is the number of generated sample vectors of $\underline{g}$, S is the number of equally spaced points for the grid search, and I is the total number of iterations.

[3]The grid search is employed to simplify the algorithm. More elegant line search algorithms such as the golden section method [10] can be employed instead of the grid search.

The proposed CR system is based on the interference channel in Fig. 1, and thus it is worthwhile to compare the achievable rates of the proposed system with the interference channel capacity region. Unfortunately, however, the interference channel capacity region is still an open problem for general Gaussian/fading channels. The capacity region is known in the case of strong interference [11]–[13], in which channel parameters satisfy both $\frac{g_{12}}{N_2} \ge \frac{g_{11}}{N_1}$ and $\frac{g_{21}}{N_1} \ge \frac{g_{22}}{N_2}$, and for this case we make the following observation.

Observation 1: If all channel realizations satisfy the strong interference channel condition, the proposed system with $Q = 0$ achieves the boundary point of the capacity region at which the primary user's rate is maximized.

Proof: Under the condition of strong interference and the average transmission power constraints, the boundary of the capacity region for a fading interference channel can be achieved if the transmission powers of the two senders are optimally allocated and the two receivers perform MUD whenever both transmission powers are positive [13]. Suppose that $P_1^* > 0$. Due to the strong interference channel condition, $R_1^* = C\left(\frac{g_{11}P_1^*}{N_1+Q}\right) \le I(X_1; Y_2|X_2) = C\left(\frac{g_{12}P_1^*}{N_2}\right)$, and thus $\underline{g} \in \mathcal{G}_2$. Then, referring to Table I, $\mathcal{D}_2$ always performs MUD and $\mathcal{D}_1$ also performs MUD whenever $P_2^* > 0$. Hence both $\mathcal{D}_1$ and $\mathcal{D}_2$ perform MUD whenever both P_1^* and P_2^* are positive. In our scheme, the transmission powers are optimally allocated in such a way that P_1 is optimized first to maximize the primary user's rate and then, given P_1^*, P_2 is optimized to maximize the secondary user's rate. Therefore, its achievable rate pair $(\bar{R}_1^*, \bar{R}_2^*)$ is located at the boundary point of the capacity region at which the primary user's rate is maximized. ■

D. Operation Scenario

The proposed CR system is operated as follows.

1) During a preliminary stage, the BS, which controls the terminals $(\mathcal{S}_i, \mathcal{D}_i)$, $i \in \{1, 2\}$, collects information on the pdf of $\underline{g}$ and performs the optimization process for obtaining λ_1^*. Then λ_1^* and the information on the pdf of $\underline{g}$ are delivered to $\mathcal{S}_2$, $\mathcal{D}_1$, and $\mathcal{D}_2$.
2) During a communication stage, elements of the channel vector $\underline{g}$ are periodically estimated by $(\mathcal{D}_1, \mathcal{D}_2)$. The BS collects the instantaneous channel estimates and evaluates the optimal value of (J, K, L), denoted by (J^*, K^*, L^*), for a given channel estimate. Because (J^*, K^*, L^*) corresponds to $R_2^*(\lambda_1^*)$, which is considered in Step 3 of the optimization process, (J^*, K^*, L^*) can be obtained by evaluating $R_2^*(\lambda_1^*)$ for the channel estimate, following the procedure described in Step 3. The estimate of $\underline{g}$ and (J^*, K^*, L^*) are delivered to $\mathcal{S}_2$, $\mathcal{D}_1$, and $\mathcal{D}_2$, while only the estimate of g_{11} is delivered to $\mathcal{S}_1$. Then, the optimal transmission power P_2^* and the optimal rate R_2^* for $\mathcal{S}_2$ are given by $P_2^{LO}(\lambda_1^*)$ and $R_2^{LO}(\lambda_1^*)$, respectively, associated with (J^*, K^*, L^*) in Table I. The behavior of $\mathcal{D}_i$s, whether they perform SUD or MUD, is determined once (J^*, K^*, L^*) are given.

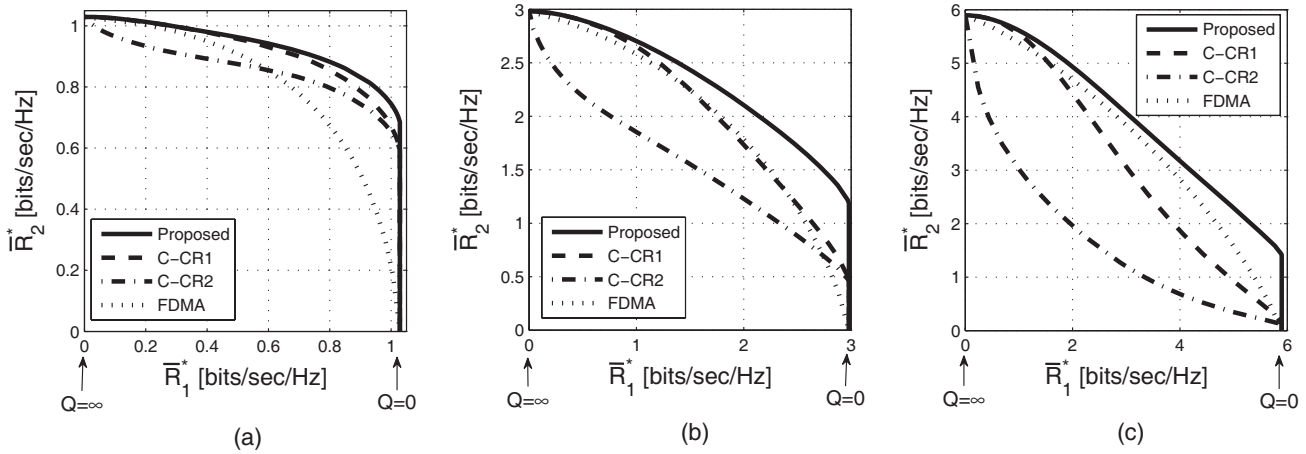

Fig. 6. Average achievable rates $\bar{R}_1^*$ versus $\bar{R}_2^*$. (a) $\frac{\bar{P}_1}{N_1} = \frac{\bar{P}_2}{N_2} = 0$ dB, (b) $\frac{\bar{P}_1}{N_1} = \frac{\bar{P}_2}{N_2} = 10$ dB, and (c) $\frac{\bar{P}_1}{N_1} = \frac{\bar{P}_2}{N_2} = 20$ dB. The maximum values of $\bar{R}_1^*$ and $\bar{R}_2^*$ are achieved when $Q = 0$ and ∞, respectively.

E. Conventional CR Systems

Table I can also be used for conventional CR systems that always maintain S-P interference power below Q. Because $\mathcal{D}_1$ always performs SUD in conventional CRs, K is fixed at zero $(K = 0)$. For a CR system performing only SUD at $\mathcal{D}_1$ and performing either SUD or MUD at $\mathcal{D}_2$ depending on $\underline{g}$, a CR system which will be called conventional CR1 (C-CR1) [14], only those $\left\{P_2^{LO}\right\}$ associated with $K = 0$ in Table I are considered during Step 3 of the optimization process described above.

If we consider a CR system performing only SUD at both $\mathcal{D}_1$ and $\mathcal{D}_2$, which will be called C-CR2, then $K = 0$ and $\mathcal{D}_2$ always assumes that $\underline{g} \in \mathcal{G}_1$ unless $P_1^* = 0$ $\left(\underline{g} \in \mathcal{G}_0\right)$. Thus, P_2^{LO} for $\underline{g} \in \mathcal{G}_0^c$ is always given by $\min\left\{\alpha_1, \frac{Q}{g_{21}}\right\}$, which is shown in the third row of Table I. It is interesting to note that the C-CR1 becomes identical to the C-CR2 when the S-P interference power limit Q is set at zero $(Q = 0)$. This is true because in C-CR1, if $Q = 0$, $\mathcal{S}_2$ transmits only when $\mathcal{S}_1$ is in an idle stage and $\mathcal{D}_2$ decodes only $\mathcal{S}_2$'s data.

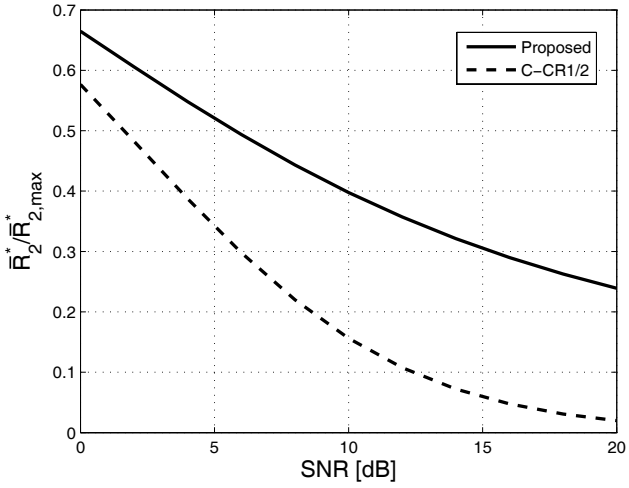

Fig. 7. $\bar{R}_2^*/\bar{R}_{2,\max}^*$ versus SNR $\left(\frac{\bar{P}_1}{N} = \frac{\bar{P}_2}{N}\right)$ for $Q = 0$. In this case, C-CR1 is identical to C-CR2. For FDMA, $\bar{R}_2^* = 0$ for $Q = 0$.

IV. Simulation Results

The performances of the proposed CR and C-CR1/2 are examined through a computer simulation. The simulation environments are as follows. The average transmission rates of $\mathcal{S}_1$ and $\mathcal{S}_2$, denoted by $\bar{R}_1^*$ and $\bar{R}_2^*$, respectively, are obtained based on 10^6 realizations of $\underline{g}$, where g_{11}, g_{12}, g_{21}, and g_{22} are independent of each other and follow a Rayleigh distribution. The average channel gains are fixed at one, i.e., $\mathbb{E}[g_{11}] = \mathbb{E}[g_{22}] = \mathbb{E}[g_{12}] = \mathbb{E}[g_{21}] = 1$, and $\bar{P}_1 = \bar{P}_2 = 1$. The noise powers N_1 and N_2 are varied so that the signal-to-noise ratio (SNR) $\frac{\bar{P}_i}{N_i}$ for $i \in \{1, 2\}$, can take 0 dB, 10 dB, and 20 dB. For simplicity, it is assumed that the SNRs of the primary and secondary users are identical $\left(\frac{\bar{P}_1}{N_1} = \frac{\bar{P}_2}{N_2}\right)$. The S-P interference power limit Q varies from zero to infinity. The maximum values of $\bar{R}_1^*$ and $\bar{R}_2^*$, denoted as $\bar{R}_{1,\max}^*$ and $\bar{R}_{2,\max}^*$, are obtained for $Q = 0$ and ∞, respectively, and we are interested in Q values in the vicinity of zero because the secondary user of a CR system can transmit data without sacrificing the primary user's transmission rate $\bar{R}_1^*$ when $Q = 0$. For comparison, the average achievable rates for frequency division multiple access (FDMA) are also considered.

The achievable rates $\bar{R}_1^*$ and $\bar{R}_2^*$ are shown in Fig. 6. In addition, to emphasize the advantage of the CR systems over FDMA, the ratio between $\bar{R}_2^*$ and $\bar{R}_{2,\max}^*$ for $Q = 0$ is shown in Fig. 7. (Obviously, $\bar{R}_2^*$ for FDMA corresponding to the case in which $Q = 0$ is equal to zero $(\bar{R}_2^* = 0)$.) As expected, among the three CR systems, the proposed system performs the best and C-CR2 performs the worst. When the SNR is 0 dB (Fig. 6 (a)), the advantage of the CR systems over FDMA is significant. For example, the proposed CR attains as much as 66% of $\bar{R}_{2,\max}^*$, even when $Q = 0$. This occurs because the primary user $\mathcal{S}_1$ frequently stops transmission due to its poor channel quality. The performance gain achieved by the CRs reduces as the SNR increases. For example, referring to Fig. 7, $\bar{R}_2^*$ for the proposed CR with $Q = 0$ becomes about 40% (24%) of $\bar{R}_{2,\max}^*$ when the SNR is 10 dB (20 dB). These $\bar{R}_2^*$ values are achieved without sacrificing $\bar{R}_1^*$, and

thus the proposed CR would be useful even for SNR $= 20$ dB. On the other hand, use of the C-CR$1/2$ may not be recommended for high SNR because their performance gain over FDMA is rather minor, even for $Q = 0$. For example, they achieve only 2% of $\bar{R}^*_{2,\max}$ when $Q = 0$ and the SNR is 20 dB (Fig. 7). Furthermore, for SNR $= 20$ dB (Fig. 6 (c)) the C-CR$1/2$ almost always show worse performance than FDMA, while the proposed CR always outperforms FDMA. This indicates the importance of the opportunistic violation of the S-P interference power limit and the MUD at $\mathcal{D}_1$, especially for high SNR.

To get some additional insight into the behavior of the proposed CR system, channel gains resulting in MUD at $\mathcal{D}_1$ $(K = 1)$ were examined, and such channels were observed to have relatively large g_{21} values. Specifically, the following observations have been made for $Q = 0$.

Observation 2: If g_{21} is either the largest or the second-to-largest among $\{g_{11}, g_{12}, g_{21}, g_{22}\}$, then the probability that $\mathcal{D}_1$ performs MUD is 55%, 82%, and 88% for SNR $= 0$ dB, 10 dB, and 20 dB, respectively.

The reason why the probability for SNR $= 0$ dB is considerably smaller than the others is because when SNR $= 0$ dB, $\mathcal{S}_2$ frequently stops transmission due to poor channel quality.

Observation 3: Among the channel gains resulting in MUD at $\mathcal{D}_1$, about 70% have g_{21}, which is either the largest or the second-to-largest among $\{g_{11}, g_{12}, g_{21}, g_{22}\}$.

Similar observations can be made for $Q > 0$. The reason why $\mathcal{D}_1$ performs MUD when g_{21} is large can be explained as follows. When $\mathcal{D}_1$ performs SUD and $Q > 0$, the interference power limit, $P_2 g_{21} \leq Q$, severely limits P_2 for large values of g_{21}. To overcome this difficulty, P_2 tends to violate the S-P interference power limit and $\mathcal{D}_1$ performs MUD. On the other hand, if $Q = 0$ and $P_1^* > 0$, S_2 is allowed to transmit only when $\mathcal{D}_1$ performs MUD. To maintain $R_1^* = I(X_1; Y_1|X_2)$ while performing MUD at $\mathcal{D}_1$, R_2 should be upper bounded by $I(X_2; Y_1) = C\left(\frac{g_{21}P_2}{N_1 + g_{11}P_1^*}\right)$ (see Fig. 2). Therefore, to increase R_2, $\mathcal{S}_2$ tends to allocate more power to channels with a large g_{21} and a small g_{11}; this policy will increase $I(X_2; Y_1)$. The impact of a small g_{11} on $\mathcal{D}_1$'s behavior is considerably less than that of a large g_{21} because $\mathcal{S}_1$ stops transmission if g_{11} is too small.

Finally, in this section, we examine channel gains which result in MUD at $\mathcal{D}_1$ but have a small g_{21}. As indicated by Observation 2, about 30% of the channels resulting in MUD at $\mathcal{D}_1$ have g_{21}, which is either the smallest or the second-to-smallest among $\{g_{11}, g_{12}, g_{21}, g_{22}\}$. We evaluate the average of such channel gains and listed them in Table II. It is interesting to note that for all three SNR values, $\bar{g}_{22}$ is the largest and $\bar{g}_{21}$ is the smallest, where $\bar{g}_{ij}$ denotes the average of g_{ij}. The proposed system tends to violate the S-P interference power limit when $g_{22} \gg g_{21}$ because the increase of R_2 that can be achieved by such violation can be significant.

TABLE II
AVERAGE OF THE CHANNEL GAINS RESULTING IN MUD AT $\mathcal{D}_1$ AND HAVING g_{21} WHICH IS EITHER THE SMALLEST OR THE SECOND-TO-SMALLEST AMONG $\{g_{11}, g_{12}, g_{21}, g_{22}\}$. HERE $\bar{g}_{ij}$ DENOTES THE AVERAGE OF g_{ij} IN SUCH CHANNELS.

	$\bar{g}_{11}$	$\bar{g}_{12}$	$\bar{g}_{21}$	$\bar{g}_{22}$
SNR $= 0$ dB	0.91	1.43	0.86	1.67
SNR $= 10$ dB	0.70	1.37	0.64	1.43
SNR $= 20$ dB	0.68	1.35	0.60	1.39

V. CONCLUSION

A CR system that opportunistically permits excessive S-P interference was proposed, and an optimal power/rate control policy for maximizing the secondary user's average transmission rate was derived. The advantage of the proposed CR over conventional CR systems was shown by comparing their achievable rates, which are obtained by computer simulation. The results demonstrated that the proposed CR achieved considerably high data rates for the secondary user without sacrificing the data rate for the primary user and outperformed conventional CRs.

The proposed system is derived under the assumption that the primary sender $\mathcal{S}_1$ operates independently, while ignoring $\mathcal{S}_2$. Further work in this area will include extensions of this CR system for joint optimization of the two senders and for supporting multiple primary/secondray users.

REFERENCES

[1] Federal Communications Commission, "Spectrum policy task force report," ET Docket, no. 02-135, Nov. 2002.
[2] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
[3] Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE 802.22 Working Draft Proposed Standard, Rev. 0.2, Nov. 2006.
[4] Y.-C. Liang, H.-H. Chen, J. Mitola, P. Mahonen, R. Kohno, J. H. Reed, and L. Milstein, "Cognitive radio: theory and application," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, Jan. 2008.
[5] A. Swami, R. A. Berry, A. M. Sayeed, V. Tarokh, and Q. Zhao, "Signal processing and networking for dynamic spectrum access," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, Feb. 2008.
[6] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.
[7] S. Sridharan and S. Vishwanath, "On the capacity of a class of MIMO cognitive radios," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 103–117, Feb. 2008.
[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: John Wiley & Sons, Inc., 2006.
[9] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.
[10] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.
[11] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.
[12] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 6, pp. 786–788, Nov. 1981.
[13] S. T. Chung and J. M. Cioffi, "The capacity region of frequency-selective Gaussian interference channels under strong interference," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1812–1821, Sep. 2007.
[14] P. Popovski, H. Yomo, K. Nishimori, R. D. Taranto, and R. Prasad, "Opportunistic interference cancellation in cognitive radio systems," in *Proc. IEEE Int. Symp. Dynamic Spectrum Access Networks (DySPAN)*, Dublin, Ireland, Apr. 2007.

Woohyuk Chang (S'03) received the B.S. degree in electrical engineering from Ajou University, Suwon, Korea, in 2001. He received the M.S. degree in electrical engineering from KAIST, Daejeon, Korea, in 2003, and is currently working toward the Ph.D. degree at KAIST. His research interests include information theory, coding theory, and signal processing for cooperative relaying and cognitive radio networks. Mr. Chang is the recipient of the Silver Prize in the 15th Samsung Humantech Paper Contest.

Sae-Young Chung (S'89-M'00-SM'07) received the B.S. and the M.S. degrees in electrical engineering from Seoul National University in 1990 and 1992, respectively. He received the Ph.D. degree in the Department of EECS at MIT in 2000.

From June to August 1998 and from June to August 1999, he was with Lucent technologies. From September 2000 to December 2004, he was with Airvana, Inc., where he conducted research on the third generation wireless communications. Since January 2005, he has been with KAIST, where he is now an associate professor in the Department of EE. He is an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS. He served as a guest editor of JCN, EURASIP journal, and Telecommunications review. He served as a TPC co-chair of WiOpt 2009 and a tutorial co-chair of IEEE ISIT 2009. He is a IEEE Information Theory Society membership/chapters committee member, IEEE Communication Theory Technical Committee member, and IEEE Communications Society Asia Pacific Board member. His research interests include network information theory, coding theory, and their applications to wireless communications. He is a senior member of the IEEE.

Yong H. Lee (S'81-M'84-SM'98) was born in Seoul, Korea, on July 12, 1955. He received the B.S. and M.S. Degrees in electrical engineering from Seoul National University, Seoul, Korea, in 1978 and 1980, respectively, and the Ph.D. degree in electrical engineering from the University of Pennsylvania, Philadelphia, U.S.A., in 1984.

From 1984 to 1988, he was an Assistant Professor with the Department of Electrical and Computer Engineering, State University of New York, Buffalo. Since 1989, he has been with the Department of Electrical Engineering, KAIST, where he is currently a Professor and the Dean of College of Information Science and Technology. His research activities are in the area of communication signal processing, which includes interference management, resource allocation, synchronization, estimation and detection for CDMA, TDMA, OFDM and MIMO systems. He is also interested in designing and implementation of transceivers.