Dirty Paper Coding for the MIMO Cognitive Radio Channel with Imperfect CSIT

Chinmay S. Vaze and Mahesh K. Varanasi

Abstract—A Dirty Paper Coding (DPC) based transmission scheme for the Gaussian multiple-input multiple-output (MIMO) cognitive radio channel (CRC) is studied when there is imperfect and perfect channel knowledge at the transmitters (CSIT) and the receivers, respectively. In particular, the problem of optimizing the sum-rate of the MIMO CRC over the transmit covariance matrices is dealt with. Such an optimization, under the DPC-based transmission strategy, needs to be performed jointly with an optimization over the inflation factor. To this end, first the problem of determination of inflation factor over the MIMO channel $Y = H_1X + H_2S + Z$ with imperfect CSIT is investigated. For this problem, two iterative algorithms, which generalize the corresponding algorithms proposed for the channel $Y = H(X + S) + Z$, are developed. Later, the necessary conditions for maximizing the sum-rate of the MIMO CRC over the transmit covariances for a given choice of inflation factor are derived. Using these necessary conditions and the algorithms for the determination of the inflation factor, an iterative, numerical algorithm for the joint optimization is proposed. Some interesting observations are made from the numerical results obtained from the algorithm. Furthermore, the high-SNR sum-rate scaling factor achievable over the CRC with imperfect CSIT is obtained.

Index Terms—Cognitive radio, dirty paper coding, inflation factor, covariance optimization.

I. INTRODUCTION

THE cognitive radio channel (CRC) was introduced in [1]. A cognitive radio is a device that can sense its environment in real time and can accordingly adapt its transmission strategy. These are of current interest because of the dramatically high spectral efficiency they can achieve [1]. In [1], the authors introduced a more general cognitive protocol under which the CRC is an interference channel with degraded message sets [2].

The Gaussian multiple-input multiple-output (MIMO) interference channel consists of two transmitter-receiver pairs with each transmitter having a message for its paired receiver and the received signals are defined via equations

$$Y_1 = H_1X_1 + H_2X_2 + Z_1, \quad Y_2 = H_1X_1 + H_2X_2 + Z_2.$$ 

Here, $(H_{ij})_{ij=1}^2$ are the fading channel matrices of dimensions $r_i \times t_i$; the transmitted signals $X_1 \sim \mathcal{CN}(0, \Sigma_1)$ and $X_2 \sim \mathcal{CN}(0, \Sigma_2)$ are subject to the power constraint of $P_1$ and $P_2$; and $Z_1 \sim \mathcal{CN}(0, I_{r_1})$ and $Z_2 \sim \mathcal{CN}(0, I_{r_2})$ are the additive noises [2]. The Gaussian MIMO CRC is defined as the Gaussian MIMO interference channel in which the second transmitter (corresponding to signal $X_2$) or the cognitive transmitter (CT) knows the message (or the codeword) of the first or the primary transmitter (corresponding to signal $X_1$) non-causally [1].

An achievable-rate region for the Gaussian MIMO CRC has been proposed in [3]. In their coding scheme, the CT, because of its non-causal knowledge, acts as a relay to aid the primary receiver and also transmits its own message to its paired (or cognitive) receiver. It employs dirty paper coding (DPC) [4] to cancel the interference at the cognitive receiver due to the signals intended for the primary receiver. With the assumption that the required channel matrices are known at the transmitters and the receivers, it is further shown in [3] that their achievable-rate region includes points corresponding to the sum-capacity of the MIMO CRC under certain conditions. Now, to achieve the sum-capacity, an optimization over the transmit covariances is required. This problem is studied in [5] where the authors propose the so-called adaptive sum-power iterative waterfilling algorithm which computes the sum-capacity and the optimal transmit covariances.

Although the MIMO CRC is increasingly being studied under the assumption of perfect transmitter-channel knowledge (CSIT) [6], [8], [9], etc., not many papers exist which deal with the practically important scenario of imperfect-CSIT CRC. We find it timely to consider the aforementioned problem of covariance optimization under imperfect CSIT. Towards this end, one needs to seek answers to the following two questions:

1) DPC at the CT under imperfect CSIT: Since the channel seen by the cognitive receiver is of the form $Y = H_1X_1 + H_2S + Z$ (this will become more clear in Section III), where $S$ is the interference known non-causally at the CT but not at its receiver, it is imperative to first study the problem of DPC over this channel when there is imperfect CSIT of $H_1$ and $H_2$. This problem is equivalent to the determination of the optimal inflation factor (see [1]) under imperfect CSIT. We studied a similar problem for the fading dirty paper channel (FDPC) $Y = H(X + S) + Z$ in [9] and developed two iterative algorithms for determination of inflation factor. These algorithms significantly improve the prior attempts mentioned therein. The same problem for the channel $Y = H_1X_1 + H_2S + Z$, which we call the Generalized FDPC (G-FDPC), has been considered in [10], but only in the special case of (all) single-antenna terminals, and a suboptimal solution is proposed. We study this important problem in Section III of this paper.

2) Covariance optimization under imperfect CSIT: The problem of covariance optimization is considerably in-

This work was supported in part by NSF Grants CCF-0431170 and CCF-0728955 and the authors are with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309-0425 USA (e-mail: Chinmay.Vaze, varanasi@colorado.edu, Ph: 001-303-492-7327).

1 Notation: For a square matrix $A$, $tr(A)$, $|A|$, and $rank(A)$ denote its trace, determinant, and rank, respectively. For any general matrix $A$, $A^*$ and $A^{-*}$ denote the complex-conjugate transpose and pseudo-inverse of matrix $A$, respectively. vec$(A)$ denotes the vector obtained by stacking the columns of $A$. $I_m$ is the $m \times m$ identity matrix. $E_{HS}$ denotes expectation over the random variable $H$. 

arXiv:0906.2252v1 [cs.IT] 12 Jun 2009
volved, even under perfect CSIT. In [5], rather than using the achievable-rate region of [3], the authors formulate the problem in terms of an outer-bound (obtained in [3]) to the capacity region which includes points corresponding to the sum-capacity of the MIMO CRC. This is a non-convex optimization problem and is converted into an equivalent convex-concave game using the ‘MAC-BC’ transformations [11]. Then, for the resulting optimization, an iterative numerical algorithm is proposed. Unfortunately, the algorithm can not always guarantee the optimum solution.

The use of the outer-bound or the ‘MAC-BC’ transformations is not possible under imperfect CSIT. Also, unlike the perfect-CSIT case (under which the interference can be assumed to be canceled perfectly by DPC), under imperfect CSIT, an additional optimization over the inflation factor needs to be performed jointly with the transmit covariances. Furthermore, the problem becomes more complicated because the sum-rate optimal solution need not necessarily have the power constraints satisfied with equal variances. Further, the problem becomes more complicated because the transmitted signal knowledge but imperfect CSIT and let tr(Σ) = |Σ|. The above rate expression is valid only the auxiliary random variable (see [13] for definition) as for the achievable-rate region of [3], the authors formulate the the rate achievable over the G-FDPC in absence of interference (i.e., when $Q = 0$) or $R_{\text{nos}} = E[H] \log \frac{\Sigma + H S H^*}{|\Sigma|}$.

The problem of determination of inflation factor, i.e., the maximization in (1) is equivalent to min$_W \log |M| = 0$. Using the obtained necessary conditions, an iterative algorithm is proposed.

**Numerical Results:** Here, ‘lb’ denotes the rate achievable using $W = 0$, i.e., by treating the interference as noise. $R$ denotes the rate achievable using the algorithms. In Fig. 1, we take $H_1$ and $H_2$ to be independent with their elements $i.i.d.$ $\mathcal{CN}(\mu, \sigma^2)$ with $|\mu|^2 + \sigma^2 = 1$ and $\mu = |\mu| \sqrt{\frac{1}{2}}$. When $\mu = 0$, the all-zero inflation factor performs almost as well as the inflation factor obtained using the algorithms. However, as $|\mu|$ increases, the algorithms outperform the simple choice of $W = 0$. This type of observation was also made in [10] in the case of SISO G-FDPC. It is generalized here to the MIMO case. In Fig. 2, the fac-

A slightly different (due to a constraint on the rate of primary user) version of the problem is considered in [8]. [12] for the CRC with all single-antenna terminals. The authors of [8], [12] consider the amplify-and-forward strategy for relaying at the CT, and in this sense, their scheme is less general than the one studied here.

**II. DPC over the G-FDPC**

Motivation to study this problem will become more clear in Section III. But, as noted before, this is an important step in the overall joint optimization. The G-FDPC is defined via equation $Y = H_1 X + H_2 S + Z$. Here, $H_1$ and $H_2$ are the channel matrices of dimensions $r \times t_s$ and $r \times t_s$, respectively; the transmitted signal $X \sim \mathcal{CN}(0, \Sigma_X)$ has a power constraint of $P$; the interference $S \sim \mathcal{CN}(0, \Sigma_S)$ is known non-causally at the transmitter but not at the receiver; $Z \sim \mathcal{CN}(0, \Sigma_Z)$ is the additive noise; and $X$, $S$, and $Z$ are independent. Assume perfect receiver channel knowledge but imperfect CSIT [3]. Assume $|\Sigma_X|, |\Sigma_Z| > 0$; let tr($\Sigma_S$) = $Q$, tr($\Sigma_Z$) = $N$. Define SNR = $\frac{Q}{N}$. Select the auxiliary random variable (see [13] for definition) as $U = X + W S$, i.e., Costa’s choice [4] extended to the MIMO case, where the $t_x \times t_s$ matrix $W$ is the inflation factor [4]. Similar to [4], we obtain the achievable rate as given by

$$R = \max_{W} \log \frac{E_H \log \frac{\Sigma_X + H_1 \Sigma_S H_1^* + H_2 \Sigma_S H_2^*}{|M|}}{2}$$

with $M = \begin{bmatrix} \Sigma_X + W S \Sigma_S W^* & \Sigma_X H_1^* + W S \Sigma_S H_2^* \\ H_1 \Sigma_X H_1^* + H_2 \Sigma_S W^* \Sigma_Z + H_1 \Sigma_X H_1^* + H_2 \Sigma_S H_2^* \end{bmatrix}$, and $H = [H_1 \ H_2]$. The above rate expression is valid only if $|\Sigma_X| > 0$. The case of $|\Sigma_X| = 0$ can be handled as in [13]. We define the non-interference upper-bound $R_{\text{nos}}$ as

$$R_{\text{nos}} = \log \frac{E[H] \log \frac{\Sigma + H S H^*}{|\Sigma|}}{2}$$

$\Sigma$ is total noise power.

$\Sigma$ is total noise power.

$\Sigma$ is total noise power.

$\Sigma$ is total noise power.
Achievable rates vs. SNR: $H_1$ and $H_2$ correlated.

Fig. 2

III. Optimization over the Transmit Covariances

As per the coding scheme of [3], let $X_2 = X_{21} + X_{22}$ where the signal $X_{21}$ corresponds to relaying and is correlated with $X_1$ while $X_{22}$ is the signal intended for the cognitive receiver. Let $\begin{bmatrix} X_1 \\ X_{21} \end{bmatrix} \sim \mathcal{CN}\left(0, \Sigma = \begin{bmatrix} \Sigma_{11} & V \\ V^* & \Sigma_{22} \end{bmatrix} \right)$, and $X_{22} \sim \mathcal{CN}(0, \Sigma_{22})$. Also let $\Sigma = T_1T_1^* + \Sigma_{22} = T_2T_2^*$ for some $T_1$ and $T_2$; and $X_{22} = T_2X_{22}'$ with $X_{22}' \sim \mathcal{CN}(0, I_{t_2})$. The CT would choose the auxiliary random variable as $U = X_{2}' + W \begin{bmatrix} X_1 \\ X_{21} \end{bmatrix}$, where $X_{22}'$ is independent of $X_1$ and $X_{21}$. Hence $\Sigma_2 = \Sigma_{22} + \Sigma_{21}$.

Now the channel between the CT-receiver pair is $Y_2 = H_{22}X_{22} + [H_{12}H_{22}] \begin{bmatrix} X_1 \\ X_{21} \end{bmatrix} + Z_2$ which resembles the G-FDPC. Therefore, using the algorithms of Section III, we can determine the inflation factor to be used at the CT once $\Sigma$ (or $T_1$) and $\Sigma_2$ (or $T_2$) are specified. This explains the reason to first study DPC over the G-FDPC.

Denote $H_1 = [H_{11}, H_{21}]$, $H_2 = [H_{12}, H_{22}]$, and $\tilde{H} = [H_{11}^* H_{22}]^*$. Then the achievable sum-rate $R_{sum} = R_p + R_c$ under no CSIT and perfect receiver channel knowledge is given by equation (2) at the bottom of the page.

Since $W$ depends on $T_1$ and $T_2$, we need to optimize $R_{sum}$ jointly over $T_1$, $T_2$, and $W$, as mentioned earlier. However, since $W$ can be determined given $T_1$ and $T_2$, let us first consider the optimization of $R_{sum}$ over $T_1$ and $T_2$ for a given value of $W$; later the algorithm for the joint optimization can be formulated. Let us consider: $\max_{T_1, T_2} R_{sum}$, subject to $\text{tr}(\Sigma_1) \leq P_p$ and $\text{tr}(\Sigma_2 + \Sigma_{22}) \leq P_c$. This is a non-convex optimization problem. To obtain the necessary conditions, we form the lagrangian $J$; and set $\frac{\partial J}{\partial T_1} = 0$ and $\frac{\partial J}{\partial T_2} = 0$. We omit the details of differentiation and directly state the necessary conditions, as given by equations (3) and (4) at the bottom of the page, where $\lambda_p^{-1}$ and $\lambda_c^{-1}$ are the lagrange multipliers.

**Algorithm for the Joint Optimization (Alg. 1):**

1. Start with some initial choices $T_1(0)$ and $T_2(0)$. For these choices, determine $W(0)$ using the algorithms discussed in Section III.
2. At the $n$th iteration,
   - Determine the transmit covariances: to this end, we set $T_1(n) = \begin{bmatrix} \lambda_p^{-1}I_t & 0 \\ 0 & \lambda_c^{-1}I_{t_2} \end{bmatrix} g_1(T_1(n-1), T_2(n-1), W(n-1))$ and $T_2(n) = \lambda_c g_2(T_1(n-1), T_2(n-1), W(n-1))$. The required expectations are evaluated numerically. Find Lagrange multipliers so as to meet the power const.
Theorem 1: G-FDPC with $|\Sigma X| > 0$, or equivalently, the G-FDPC with $H_1 = H_2 = H$ and $t_x = t_x = t$. The high-SNR scaling factor of $\min(t, r)$, which is equal to that of the corresponding no-interference upper-bound, is achievable with the choice of $W = I_t$. Next consider the FDPC with $|\Sigma X| = 0$ (let $\Sigma_X = T T^*$); this channel is then equivalent to the G-FDPC $Y = H_1 X' + H_2 S + Z$ with $H_1 = HT$, $H_2 = H$, and $X' \sim C N(0, I)$ (so $H_1$ and $H_2$ not equal). In this case, the achievable scaling factor may not always be equal to that of the no-interference upper-bound; but in most cases, by making an appropriate choice for $W$ (say, $W = T^+$ [14]), one can achieve a better scaling factor than that achievable with $W = 0$. Note, the G-FDPCs in the two cases above do not satisfy the assumption regarding $H_1$ and $H_2$ made in Theorem 1. Finally, consider the G-FDPC that satisfies the assumption of Theorem 1 (for example, $H_1$ and $H_2$ are independent and Rayleigh-faded). Then, as per Theorem 1, there is no advantage in optimizing over $W$ as far as the scaling factor is concerned. Thus, the ‘more highly’ $H_1$ and $H_2$ are correlated, the ‘larger’ is the increase in the scaling factor over that achievable by treating the interference as noise.

Theorem 2: CRC: Assume that the channel matrices $\{H_{ij}\}$ are full rank and independent; and the ratio $P_t$ remains constant. The high-SNR$(= P_t)$ sum-rate scaling factor achievable over the no-CSIT CRC is given by

$$\gamma_{\text{sum}} = \max_{\text{rank}(\Sigma'), \text{rank}(\Sigma_2)} \gamma_p + \gamma_c \text{ with }$$

$$\gamma_p = \min(r, \text{rank}(\Sigma')) - \min(r, \text{rank}(\Sigma_2)),$$

$$\gamma_c = \min(r, \text{rank}(\Sigma')) - \min(r, \text{rank}(\Sigma') - \text{rank}(\Sigma_2)),$$

where $\Sigma'$ is the covariance matrix of $[X_1 \ X_2]$; and the maximization is under the constraints of $0 < \text{rank}(\Sigma') \leq t_1 + t_2$, and $0 \leq \text{rank}(\Sigma_2) \leq \min(t_2, \text{rank}(\Sigma'))$.

Proof: We present the outline here and omit the details. Given $\Sigma$ and $\Sigma_2$, $P_t$ achieves the scaling factor of $\gamma_p$ because the primary receiver treats the interference $X_{22}$ as noise. $\gamma_c$ is achieved by the choice of $W = [0 \ T_2^+]$ (see Theorem 1 above and Theorem 1 of [14]).

The maximization in Theorem 2 is over only finitely many values; thus, can be done via exhaustive search. Note, the power constraints are not solved as strict equalities here.

IV. High-SNR Analysis: Scaling Factor

Theorem 1: G-FDPC: Assume that the ratio $P_t / N$ is constant as $P \to \infty$, and the fading processes are such that for any positive semi-definite $A$, $\text{rank}([H_1, H_2] A [H_1, H_2]^*) = \min(r, \text{rank}(A))$ with probability 1. The high-SNR scaling factor achievable over the no-CSIT G-FDPC using DPC is independent of the choice of $W$, as long as $W$ is chosen such that the term $\log |\Sigma X + \Sigma \Sigma W^*|$ scales in the high-SNR regime as $t_2 \log \text{SNR}$.

Thus, the naive scheme of treating the interference as noise (i.e., $W = 0$) achieves the optimal scaling factor, which is given by $\min(r, \text{rank}(\Sigma X) + \text{rank}(\Sigma S)) - \min(r, \text{rank}(\Sigma S))$. Also note that there is no loss of generality in choosing $W$ to satisfy the condition stated in Theorem 1 because $W$’s that do not satisfy this condition can achieve only a suboptimal scaling factor.

The intuition detailed in the paragraph just preceding Section IIII can explain the result of Theorem 1. Considering the FDPC with $|\Sigma X| > 0$, or equivalently, the G-FDPC with $H_1 = H_2 = H$ and $t_x = t_x = t$. The high-SNR scaling factor of $\min(t, r)$, which is equal to that of the corresponding no-interference upper-bound, is achievable with the choice of $W = I_t$. Next consider the FDPC with $|\Sigma X| = 0$ (let $\Sigma_X = T T^*$); this channel is then equivalent to the G-FDPC $Y = H_1 X' + H_2 S + Z$ with $H_1 = HT$, $H_2 = H$, and $X' \sim C N(0, I)$ (so $H_1$ and $H_2$ not equal). In this case, the achievable scaling factor may not always be equal to that of the no-interference upper-bound; but in most cases, by making an appropriate choice for $W$ (say, $W = T^+$ [14]), one can achieve a better scaling factor than that achievable with $W = 0$. Note, the G-FDPCs in the two cases above do not satisfy the assumption regarding $H_1$ and $H_2$ made in Theorem 1. Finally, consider the G-FDPC that satisfies the assumption of Theorem 1 (for example, $H_1$ and $H_2$ are independent and Rayleigh-faded). Then, as per Theorem 1, there is no advantage in optimizing over $W$ as far as the scaling factor is concerned. Thus, the ‘more highly’ $H_1$ and $H_2$ are correlated, the ‘larger’ is the increase in the scaling factor over that achievable by treating the interference as noise.

Theorem 2: CRC: Assume that the channel matrices $\{H_{ij}\}$ are full rank and independent; and the ratio $P_t$ remains constant. The high-SNR$(= P_t)$ sum-rate scaling factor achievable over the no-CSIT CRC is given by

$$\gamma_{\text{sum}} = \max_{\text{rank}(\Sigma'), \text{rank}(\Sigma_2)} \gamma_p + \gamma_c \text{ with }$$

$$\gamma_p = \min(r, \text{rank}(\Sigma')) - \min(r, \text{rank}(\Sigma_2)),$$

$$\gamma_c = \min(r, \text{rank}(\Sigma')) - \min(r, \text{rank}(\Sigma') - \text{rank}(\Sigma_2)),$$

where $\Sigma'$ is the covariance matrix of $[X_1 \ X_2]$; and the maximization is under the constraints of $0 < \text{rank}(\Sigma') \leq t_1 + t_2$, and $0 \leq \text{rank}(\Sigma_2) \leq \min(t_2, \text{rank}(\Sigma'))$.

Proof: We present the outline here and omit the details. Given $\Sigma$ and $\Sigma_2$, $P_t$ achieves the scaling factor of $\gamma_p$ because the primary receiver treats the interference $X_{22}$ as noise. $\gamma_c$ is achieved by the choice of $W = [0 \ T_2^+]$ (see Theorem 1 above and Theorem 1 of [14]).

The maximization in Theorem 2 is over only finitely many values; thus, can be done via exhaustive search. Note, the power constraints are not solved as strict equalities here.

V. Numerical Results for the CRC

In figures, ‘ub’ denotes the sum-rate achievable by optimizing over the transmit covariances under the assumption that the interference is perfectly canceled at the cognitive receiver (i.e., the sum-rate with $R_c = E_{H_{22}} \log |I_{r_2} + H_{222}\Sigma_2 H_{222}^*|$). We quantize each element of the fading matrices separately using an ‘equally spaced level’ quantizer as defined in [13]. In figures, if $B = [B_{ij}]$, then $B_{ij}$ denotes the number of feedback bits used per element of matrix $H_{ij}$. Further, $H_{ij}$ are independent. Alg. 1 is unfortunately sensitive to the initial choices. We take 4 to 5 initial choices in these examples and then select the best solution.

In Fig. 3 we consider the CRC with elements of $\{H_{ij}\} \sim i.i.d. N(0, 0.6, 0.64)$. The improvement in the achievable
VI. Conclusion

This paper is one of the earliest works that studies the imperfect-CSIT MIMO CRC. To the best of the authors’ knowledge, it proposes for the first time a transmission strategy for the multi-antenna CRC with imperfect CSIT. En-route, brings into focus the problem of determination of $\lambda$’s. Furthermore, the paper derives an achievable high-SNR sum-rate scaling factor. It would be worthwhile to obtain the highest-achievable sum-rate scaling factor. This problem can be interesting; recall its counterpart for the Gaussian MIMO broadcast channel, a problem that is open even after serious attempts. More efforts are needed to answer these two open questions.

References

[1] N. Devroye, P. Mitran, and V. Tarokh, “Achievable rates in cognitive radio channels,” IEEE Trans. Inform. Theory, vol. 52, no. 5, pp. 1813–1827, May 2006.
[2] A. Carleial, “Interference channels,” IEEE Trans. Inform. Theory, vol. 24, no. 1, pp. 60–70, Jan. 1978.
[3] S. Sritharan and S. Vishwanath, “On the capacity of a class of MIMO cognitive radios,” in IEEE Inform. Theory Workshop, Lake Tahoe, CA, USA, Sep. 2007, pp. 384–389.
[4] A. Jovicic and P. Viswanath, “Cognitive radio: An information-theoretic perspective,” 2006. [Online.] Available: http://arxiv.org/pdf/0802.4233.
[5] A. Somekh-Baruch, S. Vishwanath, S. Verdu, and S. Shamai, “On the capacity of cognitive radios in multiple access networks,” in 42nd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, Oct. 2008.
[6] P. H. Lin, S. C. Lin, and H. Su, “Cognitive radio with partial channel state information at the transmitter,” in IEEE Int. Conf. Commun., May 2008, pp. 1065–1071.
[7] S. Vishwanath, N. Jindal, and A. Goldsmith, “Duality, achievable rates, and sum-rate capacity of gaussian MIMO broadcast channels,” in 23rd IEEEINFOCOM, 2004.

Fig. 3
Achievable sum-rate vs. $P_p$.

Fig. 4
Achievable sum-rate vs. $P_p$.

Fig. 5
Comparison of Alg. 1 and Alg. 2.

sum-rate with the introduction of partial CSIT is evident. Here, the scaling factor of 1 is achieved for $R_{sum}$ by letting the CT to use its entire power for relaying. Hence, the curve corresponding to B2 merges with that corresponding to the no CSIT at high SNR. For the CRC of Fig. 4, we have the elements of $\{H_{ij}\} \sim$ i.i.d. Unif[0,1]. For this CRC, as per Theorem 2, the optimal solution should achieve $\lambda_{sum} = 0$. This fact can be easily seen from the plot. In Fig. 5, we have the CRC with elements of $\{H_{ij}\} \sim$ i.i.d. $\mathcal{N}(0,1)$. It can be seen that Alg. 1 outperforms Alg. 2. However, in some cases, for example, the CRCs in Figs. 8 and 9, Alg. 2 does provide a relatively tight lower-bound. Coming back to Fig. 3, again, $R_{sum}$ achieves the scaling factor of 1 whereas according Theorem 2, $\gamma_{sum} = 2$. This is achieved by setting $\Sigma_1 = 0$, i.e., the primary transmitter needs to turn off its power. The apparent inconsistency here is because we have considered a suboptimal solution of solving the power constraints as strict equalities. This example emphasizes the importance of the problem of determination of $\lambda$’s.
channels,” IEEE Trans. Inform. Theory, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[12] P. H. Lin, S. C. Lin, and H. Su, “Design of the cognitive radio with partial channel state information at the transmitter,” in 43rd Annual Conf. on Information Sciences and Systems, Baltimore, MD, USA, Mar. 2009.

[13] S. Gelfand and M. Pinsker, “Coding for channel with random parameters,” Problems of control and information theory, vol. 9, no. 1, pp. 19–31, 1980.

[14] C. S. Vaze and M. K. Varanasi, “On the achievable rate of the fading dirty paper channel with imperfect csit,” in 43rd Annual Conf. on Information Sciences and Systems, Baltimore, MD, USA, Mar. 2009.

[15] J. Max, “Quantizing for minimum distortion,” IEEE Trans. Inform. Theory, vol. 6, no. 1, pp. 7–12, Mar. 1960.