1

The Cognitive Interference Channel with a Causal Relay in Very Strong Interference

Fernando Reátegui, Muhammad Ali Imran and Rahim Tafazolli

Abstract—We study the cognitive interference channel where an additional node (a relay) is present. In our model the relay's operation is causal rather than strictly causal, i.e., the relay's transmit symbol depends not only on its past but also on its current received symbol. We derive outer bounds for the discrete and Gaussian cases in very strong interference. A scheme for achievability based on instantaneous amplify-and-forward relaying is proposed for this model. The inner and outer bounds coincide for the special case of very strong interference.

Index Terms—Cognitive Interference Channel, Causal Relaying, Interference Relay Channel.

I. INTRODUCTION

MOTIVATED by the need to improve the transmission efficiency in modern communication systems, the idea of a radio that is aware of its surroundings and is able to accommodate its transmission scheme to achieve this goal has been proposed [1]. A radio of this type is called cognitive. In the Information theory community, the approach to cognition has taken a rather distinctive route: the cognitive radio is assumed to have causal or non-causal knowledge of the other user's current transmitted message. This assumption allows the cognitive transmitter to apply sophisticated encoding schemes in order to mitigate the interference at its receiver and simultaneously utilise part of its power in order to cooperate with the primary user. A model with such capability is known as a Cognitive Interference Channel (CIC) [2].

The capacity region of the CIC is known in very strong interference, weak interference, amongst other regimes [2]. The capacity in the very strong interference regime follows by similar techniques utilised to achieve the capacity region of the Interference Channel (IC) in strong interference [3].

Moreover, it has been shown in the literature that cooperative communication can improve the achievable rate region of several channel configurations. In [4], a relay is added to the CIC in order to help the communication; however, its operation is strictly causal. In this paper we extend this setting and assume that the relay operation is causal rather than strictly causal, i.e., the relay's transmit symbol depends not only on its past received symbols, but also on its current received symbol. We name this model the CIC with a causal relay (CIC-CR). Models with causal relays have been shown in previous works [5] to be good models for studying amplify-and-forward (AF) relaying if the overall delay spread, including the path through the relay, is much smaller than the inverse of the bandwidth.



Fig. 1. The discrete memoryless CIC with a causal relay. Cognitive and primary transmitter are denoted with subscripts 1 and 2 respectively.

We derive an outer bound to the capacity region for the discrete channel and apply the results to the Gaussian case. A scheme for achievability that utilises instantaneous AF at the relay is also presented. This simple encoding scheme is close to the capacity region in very strong interference and is tight when a relation between the channel coefficients holds. The rest of the paper is organised as follows. The discrete memoryless channel is treated in Section II and the Gaussian channel in Section III. We conclude the work in Section IV.

II. THE CIC WITH A CAUSAL RELAY

First the notation is described and the model introduced.

A. Notation

The notation of [6] is utilised throughout the paper. Lower case letters (e.g. x, y) are used to denote values of random variables. Upper case letters (e.g. X, Y) denote random variables. A sequence of random variables $(X_i, ...X_j)$ is denoted by X_i^j , for $1 \le i \le j$. When i = 1 the subscript is dropped: $X^j = (X_1, ..., X_j)$. The probability mass function (pmf) of a random variable (RV) X is denoted by $p_X(x)$. We often drop the subscript when the pmf is understood from the context, e.g. p(x). The entropy (differential entropy) of a RV and the mutual information between two RVs are indicated by $H(\cdot)$ $(h(\cdot))$ and $I(\cdot; \cdot)$ respectively. The capacity function is defined as $C(x) = 1/2 \log(1+x)$ where the logarithm is to the base 2.

B. Definitions

Definition 1. The discrete memoryless CIC-CR consists of three finite input sets \mathcal{X}_1 , \mathcal{X}_2 , \mathcal{X}_3 , three finite output sets \mathcal{Y}_1 , \mathcal{Y}_2 , \mathcal{Y}_3 and a probability transition function $p(y_1, y_2, y_3 | x_1, x_2, x_3)$. It is depicted in Fig. 1.

Definition 2. A $(2^{nR_1}, 2^{nR_2}, n)$ code for the discrete memoryless CIC-CR consists of a pair of uniformly distributed messages $M_1 \in [1:2^{nR_1}]$ and $M_2 \in [1:2^{nR_2}]$, two encoding

The authors are with the Centre for Communication Systems Research, University of Surrey, Guildford, Surrey, GU2 7XH UK e-mail: {f.reategui, m.imran, r.tafazolli}@surrey.ac.uk.

functions at the transmitters $X_1^n = f_1(M_1, M_2)$ and $X_2^n = f_2(M_2)$, an encoding function at the relay $X_{3i} = f_{3i}(Y_3^i)$ and two decoding functions $\hat{M}_t = g_t(Y_t^n)$, for t = 1, 2. The average probability of error is defined as $P_e^{(n)} = P(\bigcup_t \{\hat{M}_t \neq M_t\})$. A rate pair (R_1, R_2) is said to be achievable if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that $\lim_{n\to\infty} P_e^{(n)} = 0$. The *capacity* region of the discrete memoryless CIC-CR is the closure of the set of all achievable rate regions. We assume that the channel is memoryless, i.e., $(X_1^{i-1}, X_2^{i-1}, Y_3^{i-1}) \rightarrow$ $(X_{1i}, X_{2i}) \rightarrow Y_{3i}$, and $(X_1^{i-1}, X_2^{i-1}, X_3^{i-1}, Y_3^{i-1}, Y_1^{i-1}, Y_2^{i-1}) \rightarrow$ $(X_{1i}, X_{2i}, X_{3i}, Y_{3i}) \rightarrow (Y_{1i}, Y_{2i})$, form Markov chains.

Definition 3. The discrete memoryless CIC-CR is in the very strong interference regime if

$$I(X_1; Y_1 | X_2, Y_3, X_3) \le I(X_1; Y_2 | X_2, Y_3, X_3),$$
 (1)

$$I(X_1, X_2; Y_2 | Y_3, X_3) \le I(X_1, X_2; Y_1 | Y_3, X_3),$$
(2)

for all $p(x_1, x_2)$ and $x_{3i} = f_{3i}(y_3^i)$.

The interference regime convention of [2] for the CIC is utilised. From (1), decoding M_1 at receiver 2 should not constrain the maximum rate R_1 . Similarly from (2), requiring to decode both messages at receiver 1 should not constrain the maximum sum-rate at receiver 2, i.e., there is no rate penalty in decoding the interference at the unintended receiver. The model is said to be in strong interference if only (1) holds.

Theorem 1. An outer bound to the capacity region of the discrete memoryless CIC with a causal relay in strong interference consists of the set of rate pairs (R_1, R_2) satisfying

$$R_1 \le I(X_1; Y_3 | X_2) + I(X_1; Y_1 | X_2, Y_3, X_3),$$
(3a)

$$R_1 + R_2 \le I(X_1, X_2; Y_3) + I(X_1, X_2; Y_2 | Y_3, X_3),$$
 (3b)

for all $p(x_1, x_2)$ and $x_{3i} = f_{3i}(y_3^i)$.

Proof. See Appendix A.

III. THE GAUSSIAN CIC WITH A CAUSAL RELAY

In the Gaussian channel the relay node is equipped with one antenna for reception and another one for transmission. The antennas are isolated, therefore they do not interfere with each other. The channel model is described as follows:

$$Y_{1} = h_{11}X_{1} + h_{12}X_{2} + h_{13}X_{3} + Z_{1},$$

$$Y_{2} = h_{21}X_{1} + h_{22}X_{2} + h_{23}X_{3} + Z_{2},$$

$$Y_{3} = h_{31}X_{1} + h_{32}X_{2} + Z_{3},$$
(4)

where h_{rt} is the channel coefficient from transmitter t to receiver r for r, t = 1, 2, 3. $Z_r \sim \mathcal{N}(0, 1)$ is zero-mean unitvariance Gaussian noise. The average power constraints at the three transmitters are P_1, P_2 and P_3 respectively.

A. The Gaussian CIC-CR in very strong interference

Theorem 2. An outer bound to the capacity region of the Gaussian CIC with a causal relay in strong interference consists of the set of rate pairs (R_1, R_2) satisfying

$$R_{1} \leq \mathcal{C}\left((h_{11}^{2} + h_{31}^{2})(1 - \rho^{2})P_{1}\right),$$
(5a)
$$R_{1} + R_{2} \leq \mathcal{C}\left((h_{21}^{2} + h_{31}^{2})P_{1} + (h_{22}^{2} + h_{32}^{2})P_{2} + (h_{21}^{2}h_{32}^{2})P_{32}\right)$$

$$+ h_{22}^2 h_{31}^2) (1 - \rho^2) P_1 P_2 + 2(h_{21}h_{22} + h_{31}h_{32}) \rho \sqrt{P_1 P_2} - 2h_{21}h_{22}h_{31}h_{32} (1 - \rho^2) P_1 P_2),$$
(5b)

where $\rho \in [0, 1]$ denotes the correlation between X_1 and X_2 . *Proof.* See Appendix B.

In the following it is assumed that the relay in the Gaussian CIC-CR operates on instantaneous amplify-and-forward relying, $X_3 = \alpha Y_3$, where α is the amplification factor.

Lemma 1. The Gaussian CIC-CR is said to be in the very strong interference regime if the following holds

$$\begin{aligned} & \frac{(1-\rho^2)(h_{11}+h_{13}h_{31}\alpha)^2 P_1}{h_{13}^2\alpha^2+1} \leq \frac{(1-\rho^2)(h_{21}+h_{23}h_{31}\alpha)^2 P_1}{h_{23}^2\alpha^2+1}, \\ & \frac{1}{h_{23}\alpha^2+1} \Big((h_{21}+h_{23}h_{31}\alpha)^2 P_1 + (h_{22}+h_{23}h_{32}\alpha)^2 P_2 + \\ & (h_{21}h_{22}+(h_{21}h_{32}+h_{22}h_{31})h_{23}\alpha+h_{23}^2h_{31}h_{32}\alpha^2) 2\rho\sqrt{P_1P_2} \Big) \\ & \leq \frac{1}{h_{13}\alpha^2+1} \Big((h_{11}+h_{13}h_{31}\alpha)^2 P_1 + (h_{12}+h_{13}h_{32}\alpha)^2 P_2 + \\ & (h_{11}h_{12}+(h_{11}h_{32}+h_{12}h_{31})h_{13}\alpha+h_{13}^2h_{31}h_{32}\alpha^2) 2\rho\sqrt{P_1P_2} \Big). \end{aligned}$$

Proof. It is obtained by applying (1) and (2) to the model in (4) under the assumption of the relay operation. \Box

The instantaneous amplify-and-forward relaying has also been proposed for other models [5]. Under this assumption, the following equivalent CIC can be found:

$$Y_1 = (h_{11} + h_{13}h_{31}\alpha)X_1 + (h_{12} + h_{13}h_{32}\alpha)X_2 + h_{13}\alpha Z_3 + Z_1,$$

$$Y_2 = (h_{21} + h_{23}h_{31}\alpha)X_1 + (h_{22} + h_{23}h_{32}\alpha)X_2 + h_{23}\alpha Z_3 + Z_2.$$

As the capacity region of the cognitive interference channel in the very strong interference regime is known [2]: $R_1 \leq I(X_1; Y_1|X_2)$, $R_1 + R_2 \leq I(X_1, X_2; Y_2)$, it can be directly applied to the equivalent CIC channel. Two different values of α optimise R_1 and the sum-rate respectively. The description of the rate region is omitted due to space constraints.

Remark 1. For the symmetric channel, i.e., $P_1 = P_2$, $h_{11} = h_{22}$, $h_{31} = h_{32}$ and $h_{13} = h_{23}$, it can be verified that as long as $h_{21} = h_{11}$ and $h_{12} \ge h_{21}$, the region in (5) is the capacity region of the CIC-CR in very strong interference.

B. Comparison plots

Fig. 2 shows an example of an achievable rate region obtained by time sharing between the two values of α computed in the optimisation process and by varying the degree of cooperation at the cognitive transmitter (ρ). Fig. 2 also depicts the outer bound in (5) and an achievable rate region for this model by utilising the decode-and-forward (DF) relaying scheme. When DF is utilised the encoding function at the relay depends on the past received symbols only. This is shown for three values of the transmit power at the relay. It can be noted that instantaneous AF outperforms DF for this particular scenario. Additionally, it can be pointed out that as we are studying a scenario in very strong interference, the decoding at the relay does not constrain the achievable rates; DF performs better than partial DF. Further studies could combine DF and instantaneous AF for this setting.



Fig. 2. Inner and outer bounds for $h_{11} = h_{22} = 1$, $h_{21} = 2$, $h_{12} = 3$, $h_{31} = h_{32} = \sqrt{12}$, $h_{13} = h_{23} = \sqrt{2}$, $P_1 = P_2 = 1$.

IV. CONCLUSION

We derived outer bounds for the discrete and Gaussian cognitive interference channels with a causal relay in strong interference. We showed that a simple relaying scheme that instantaneously amplifies and forwards what it receives is optimal for the symmetric channel. In other words, not only can the transmission rates be increased by means of deploying re-transmission points in cognitive scenarios, but a rather simple relaying scheme allows to do so. We also showed that this scheme outperforms decode-and-forward at the relay as the latter only utilises the past received symbols in its encoding. The derived outer bound is general, i.e., it is a benchmark for any scheme utilised at the relay.

APPENDIX A

PROOF OF THEOREM 1

The following proposition and lemma are utilised.

Proposition 1. If $I(X_1^k; Y_1^k | X_2^k, Y_3^k) \leq I(X_1^k; Y_2^k | X_2^k, Y_3^k)$, for all $p(x_1^k, x_2^k)$, $X_3^k = f_3^k(Y_3^k)$ and $k \in [1 : n]$, then $I(X_1^k; Y_1^k | X_2^k, Y_3^k, V) \leq I(X_1^k; Y_2^k | X_2^k, Y_3^k, V)$, for all $p(v)p(x_1^k, x_2^k | v)$ and $X_3^k = f_3^k(Y_3^k)$, and V complies with the encoding Markov chains.

Proof.

$$\begin{split} I(X_1^k;Y_1^k|X_2^k,Y_3^k,V) &= \sum_v p(v)I(X_1^k,Y_1^k|X_2^k,Y_3^k,V=v),\\ &\stackrel{(a)}{\leq} \sum_v p(v)I(X_1^k,Y_2^k|X_2^k,Y_3^k,V=v),\\ &= I(X_1^k,Y_2^k|X_2^k,Y_3^k,V), \end{split}$$

where (a) follows from the hypothesis.

Lemma 2. If (1) holds, then $I(X_1^n; Y_1^n | X_2^n, Y_3^n, U) \leq I(X_1^n; Y_2^n | X_2^n, Y_3^n, U)$ for all $p(u)p(x_1^n, x_2^n | u)$ and $x_{3k} = f_{3k}(y_3^k)$ where $k \in [1:n]$ and U can be any combination of past symbols that comply with the encoding Markov chains.

Proof. By mathematical induction, for n = 1:

$$I(X_{1,1}; Y_{2,1}|X_{2,1}, Y_{3,1}, U) - I(X_{1,1}; Y_{1,1}|X_{2,1}, Y_{3,1}, U)$$

$$\stackrel{(a)}{=} I(X_{1,1}; Y_{2,1} | X_{2,1}, Y_{3,1}, X_{3,1}, U) - I(X_{1,1}; Y_{1,1} | X_{2,1}, Y_{3,1}, X_{3,1}, U) \stackrel{(b)}{\geq} 0$$

where (a) holds since $X_{3,1} = f_{3,1}(Y_{3,1})$ and (b) holds from (1) and Proposition 1. Assuming the lemma holds for n - 1, i.e.,

$$I(X_1^{n-1};Y_1^{n-1}|X_2^{n-1},Y_3^{n-1},U) \le I(X_1^{n-1};Y_2^{n-1}|X_2^{n-1},Y_3^{n-1},U), \quad (6)$$

for all $p(u)p(x_1^{n-1}, x_2^{n-1}|u)$ and $X_3^{n-1} = f_3^{n-1}(Y_3^{n-1})$. We now apply a similar procedure to the one in [3]:

$$\begin{split} I(X_1^n;Y_2^n|X_2^n,Y_3^n,U) &= I(X_1^n;Y_1^n|X_2^n,Y_3^n,U) \\ &= I(X_1^n;Y_2^{n-1}|X_2^n,Y_3^n,U) \\ &+ I(X_1^n;Y_{1n}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &- I(X_1^n;Y_1^{n-1}|X_2^n,Y_3^n,U,Y_{1n}), \\ &= I(X_1^n,Y_{1n};Y_2^{n-1}|X_2^n,Y_3^n,U,Y_{1n}), \\ &= I(X_1^n,Y_{1n};Y_2^{n-1}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &- I(X_1^n;Y_2^{n-1};Y_{1n}|X_2^n,Y_3^n,U,Y_{1n}), \\ &= I(Y_{1n};Y_2^{n-1}|X_2^n,Y_3^n,U,Y_{1n}), \\ &= I(Y_{1n};Y_2^{n-1}|X_2^n,Y_3^n,U,Y_{1n}), \\ &= I(X_1^{n-1};Y_2^{n-1}|X_2^n,Y_3^n,U,Y_{1n}) \\ &+ I(X_1^{n-1};Y_2^{n-1}|X_2^n,Y_3^n,U,Y_{1n},X_1^{n-1}) \\ &+ I(X_{1n};Y_2^{n-1}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &+ I(X_{1n};Y_{2n}|X_2^n,Y_3^n,U,Y_2^{n-1},X_{1n}) \\ &- I(Y_2^{n-1};Y_{1n}|X_2^n,Y_3^n,U,Y_2^{n-1},X_{1n}) \\ &- I(X_{1n};Y_{1n}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &- I(X_{1n};Y_{1n}|X_2^n,Y_3^n,U,Y_{2^{n-1}}), \\ &(a) I(X_1^{n-1};Y_1^{n-1}|X_2^n,Y_3^n,U,Y_{1n}) \\ &- I(X_{1n};Y_{1n}^{n-1}|X_2^n,Y_3^n,U,Y_{1n}) \\ &+ I(X_{1n};Y_{2n}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &- I(X_{1n};Y_{1n}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &- I(X_{1n};Y_{1n}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &- I(X_{1n};Y_{1n}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &- I(X_{1n};Y_{1n}|X_2^n,Y_3^n,U,Y_{2^{n-1}}) \\ &- I(X_{1n};Y_{1n}|X_2^n,Y_3^n,U,Y_{1n}) , \\ &(b) I(X_1^{n-1};Y_1^{n-1}|X_2^n,Y_3^n,U,Y_{1n}) \\ &- I(X_1^{n-1};Y_1^{n-1}|X_2^n,Y_3^n,U,Y_{1n}) \\ &- I(X_1^{n-1};Y_1^{n-1}|X_2^n,Y_3^n,U,Y_{1n}), \\ &(b) I(X_1^{n-1};Y_1^{n-1}|X_2^n,Y_3^n,U,Y_{1n}) , \\ &(c) 0, \end{aligned}$$

where (a) follows from the memoryless property of the channel and as $X_3^n = f_3^n(Y_3^n)$. (b) follows from (1) and Proposition 1 with $V = X_2^{n-1}, Y_3^{n-1}, X_3^{n-1}, U, Y_2^{n-1}$, and (c) follows from (6) and Proposition 1 with $V = X_{2n}, Y_{3n}, Y_{1n}$.

Next we prove the theorem. From Fano's inequality [6]

$$n(R_1 - \epsilon_n) \le I(M_1; Y_1^n, Y_3^n),$$
(7)

$$n(R_1 + R_2 - \epsilon_n) \le I(M_1; Y_1^n, Y_3^n) + I(M_2; Y_2^n, Y_3^n).$$
(8)

For brevity we omit ϵ_n in the following. From (7)

$$nR_1 \le I(M_1; Y_1^n, Y_3^n),$$

$$\begin{split} \stackrel{(a)}{=} &I(M_{1};Y_{1}^{n},Y_{3}^{n}|M_{2}), \\ &= \sum_{i=1}^{n} I(M_{1};Y_{1i},Y_{3i}|M_{2},Y_{3}^{i-1},Y_{1}^{i-1}), \\ \stackrel{(b)}{=} \sum_{i=1}^{n} I(M_{1},X_{1i};Y_{1i},Y_{3i}|M_{2},X_{2i},Y_{3}^{i-1},Y_{1}^{i-1}), \\ &\leq \sum_{i=1}^{n} I(M_{1},M_{2},Y_{1}^{i-1},X_{1i};Y_{1i},Y_{3i}|X_{2i},Y_{3}^{i-1}), \\ &\leq \sum_{i=1}^{n} I(M_{1},M_{2},Y_{1}^{i-1},X_{1i};Y_{3i}|X_{2i},Y_{3}^{i-1}) \\ &+ \sum_{i=1}^{n} I(M_{1},M_{2},Y_{1}^{i-1},X_{1i};Y_{1i}|X_{2i},Y_{3}^{i}), \\ \stackrel{(c)}{\leq} \sum_{i=1}^{n} I(M_{1},M_{2},Y_{1}^{i-1},Y_{3}^{i-1},X_{1i};Y_{3i}|X_{2i}) \\ &+ \sum_{i=1}^{n} I(M_{1},M_{2},Y_{1}^{i-1},Y_{3}^{i-1},X_{1i};Y_{1i}|X_{2i},Y_{3i},X_{3i}), \\ \stackrel{(d)}{=} \sum_{i=1}^{n} I(X_{1i};Y_{3i}|X_{2i}) + \sum_{i=1}^{n} I(X_{1i};Y_{1i}|X_{2i},Y_{3i},X_{3i}), \end{split}$$

where (a) follows from the independence of the messages, (b) as X_{1i} and X_{2i} are functions of (M_1, M_2) and M_2 respectively, (c) follows as X_{3i} is a function of Y_3^i , and (d) follows from the memoryless property of the channel. From (8)

$$\begin{split} &n(R_{1}+R_{2}) \\ &\leq I(M_{1};Y_{1}^{n},Y_{3}^{n}) + I(M_{2};Y_{2}^{n},Y_{3}^{n}), \\ &\stackrel{(e)}{\leq} I(M_{1};Y_{1}^{n},Y_{3}^{n}|M_{2}) + I(M_{2};Y_{2}^{n},Y_{3}^{n}), \\ &\stackrel{(f)}{=} I(M_{1},X_{1}^{n};Y_{1}^{n},Y_{3}^{n}|X_{2}^{n},M_{2}) + I(M_{2},X_{2}^{n};Y_{2}^{n},Y_{3}^{n}), \\ &= I(M_{1},X_{1}^{n};Y_{3}^{n}|X_{2}^{n},M_{2}) + I(M_{1},X_{1}^{n};Y_{1}^{n}|Y_{3}^{n},X_{2}^{n},M_{2}) \\ &\quad + I(M_{2},X_{2}^{n};Y_{3}^{n}) + I(M_{2},X_{2}^{n};Y_{2}^{n}|Y_{3}^{n}), \\ &\stackrel{(g)}{=} I(M_{1},M_{2},X_{1}^{n},X_{2}^{n};Y_{3}^{n}) + I(X_{1}^{n};Y_{1}^{n}|Y_{3}^{n},X_{2}^{n},M_{2}) \\ &\quad + I(M_{2},X_{2}^{n};Y_{2}^{n}|Y_{3}^{n}), \\ &\stackrel{(h)}{=} I(X_{1}^{n},X_{2}^{n};Y_{3}^{n}) + I(X_{1}^{n};Y_{1}^{n}|Y_{3}^{n},X_{2}^{n},M_{2}) \\ &\quad + I(M_{2},X_{2}^{n};Y_{2}^{n}|Y_{3}^{n}), \\ &\stackrel{(i)}{\leq} I(X_{1}^{n},X_{2}^{n};Y_{3}^{n}) + I(X_{1}^{n};Y_{2}^{n}|Y_{3}^{n},X_{2}^{n},M_{2}) \\ &\quad + I(M_{2},X_{2}^{n};Y_{2}^{n}|Y_{3}^{n}), \\ &\stackrel{(i)}{\leq} I(X_{1}^{n},X_{2}^{n};Y_{3}^{n}) + I(M_{2},X_{1}^{n},X_{2}^{n};Y_{2}^{n}|Y_{3}^{n}), \\ &\stackrel{(j)}{=} I(X_{1}^{n},X_{2}^{n};Y_{3}^{n}) + I(X_{1}^{n},X_{2}^{n};Y_{2}^{n}|Y_{3}^{n}), \\ &\stackrel{(j)}{=} I(X_{1}^{n},X_{2}^{n};Y_{3}^{n}) + I(X_{1}^{n},X_{2}^{n};Y_{2}^{n}|Y_{3}^{n}), \\ &\stackrel{(j)}{=} \sum_{i=1}^{n} I(X_{1}^{n},X_{2}^{n};Y_{3}^{i}|Y_{3}^{i-1}) + \sum_{i=1}^{n} I(X_{1}^{n},X_{2}^{n};Y_{2i}|Y_{2i}^{i-1},Y_{3}^{n}), \\ &\stackrel{(j)}{\leq} \sum_{i=1}^{n} I(X_{1}^{n},X_{2}^{n};Y_{3}^{i-1};Y_{3}) \\ &\quad + \sum_{i=1}^{n} I(X_{1}^{n},X_{2}^{n};Y_{3}^{i-1};Y_{3}) \\ &\stackrel{(k)}{=} \sum_{i=1}^{n} I(X_{1}^{n},X_{2}^{n};Y_{3}^{i-1};Y_{3}) \\ &\stackrel{(k)}{=} \sum_{i=1}^{n} I(X_{1i},X_{2i};Y_{3i}) + \sum_{i=1}^{n} I(X_{1i},X_{2i};Y_{2i}|Y_{3i},X_{3i}), \\ \\ &\stackrel{(k)}{=} \sum_{i=1}^{n} I(X_{1i},X_{2i};Y_{3i}) + \sum_{i=1}^{n} I(X_{1i},X_{2i};Y_{2i}|Y_{3i},X_{3i}), \\ \end{aligned}$$

where (e) follows from the independence of the messages, (f) as X_{1i} and X_{2i} are functions of (M_1, M_2) and M_2 respectively,

(g), (h), (j) and (k) from the memoryless property of the channel, (i) from Lemma 2. The rest of the proof is standard and follows by introducing a time-sharing random variable $Q \sim \text{Unif}[1:n]$, independent of (M_1, M_2, X_k^n, Y_k^n) for k = 1, 2, 3 and defining $X_k = X_{kQ}$ and $Y_k = Y_{kQ}$.

APPENDIX B Proof of Theorem 2

From (3a) we have

$$R_{1} \leq I(X_{1}, Y_{3}|X_{2}) + I(X_{1}; Y_{1}|X_{2}, Y_{3}, X_{3}),$$

$$\leq h(h_{31}X_{1} + h_{32}X_{2} + Z_{3}|X_{2}) - h(Z_{3}) - h(Z_{1}) + h(h_{11}X_{1} + h_{12}X_{2} + h_{13}X_{3} + Z_{1}|X_{2}, Y_{3}, X_{3}),$$

$$= h(h_{31}X_{1} + Z_{3}|X_{2}) - h(Z_{3}) + h(h_{11}X_{1} + Z_{1}|X_{2}, Y_{3}) - h(Z_{1}),$$

$$\stackrel{(a)}{\leq} h(h_{31}X_{1} + Z_{3}|X_{2}) - h(Z_{3}) + h(h_{11}X_{1} + Z_{1}|X_{2}, h_{31}X_{1} + Z_{3}) - h(Z_{1}),$$

$$= h(X_{2}, h_{31}X_{1} + Z_{3}, h_{11}X_{1} + Z_{1}) - h(X_{2}) - h(Z_{3}) - h(Z_{3}) - h(Z_{1}),$$

$$\stackrel{(b)}{\leq} C\left((h_{11}^{2} + h_{31}^{2})(1 - \rho^{2})P_{1}\right), \qquad (9)$$

similarly from (3b) we have

$$\begin{aligned} R_1 + R_2 \\ &\leq I(X_1, X_2; Y_3) + I(X_1, X_2; Y_2 | Y_3, X_3), \\ &\leq h(h_{31}X_1 + h_{32}X_2 + Z_3) - h(Z_3) - h(Z_2) \\ &+ h(h_{21}X_1 + h_{22}X_2 + Z_2 | h_{31}X_1 + h_{32}X_2 + Z_3), \\ &= h(h_{31}X_1 + h_{32}X_2 + Z_3, h_{21}X_1 + h_{22}X_2 + Z_2) \\ &- h(Z_3) - h(Z_2), \end{aligned}$$

$$\begin{aligned} \overset{(c)}{\leq} \mathcal{C}\Big((h_{21}^2 + h_{31}^2)P_1 + (h_{22}^2 + h_{32}^2)P_2 + (h_{21}^2h_{32}^2 \\ &+ h_{22}^2h_{31}^2)(1 - \rho^2)P_1P_2 + 2(h_{21}h_{22} + h_{31}h_{32}) \\ &\rho\sqrt{P_1P_2} - 2h_{21}h_{22}h_{31}h_{32}(1 - \rho^2)P_1P_2\Big), \end{aligned}$$

where (a) follows as conditioning reduces entropy, (b) and (c) follow as the Gaussian distribution maximises the conditional differential entropy for a given covariance constraint.

ACKNOWLEDGEMENT

This work was supported by the India-UK Advanced Technology Centre of Excellence in Next Generation Networks project jointly funded by the Department of Science and Technology, Government of India, and by the Engineering and Physical Sciences Research Council, UK.

References

- [1] J. Mitola, "Cognitive radio: an integrated agent architecture for software defined radio," Ph.D. dissertation, KTH, Stockholm, Dec. 2000.
- [2] S. Rini, D. Tuninetti, and N. Devroye, "New inner and outer bounds for the discrete memoryless cognitive interference channel and some capacity results," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4087–4109, 2011.
- [3] M. H. M. Costa and A. El Gamal, "The Capacity Region of the Discrete Memoryless Interference Channel with Strong Interference," *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 710–711, Sep. 1987.
- [4] M. Nilchian, V. Aref, and M. R. Aref, "Partial cognitive relay channel," in *IEEE Information Theory Workshop*, Jun. 2009, pp. 341–345.

- [5] H. Chang, S.-Y. Chung, and S. Kim, "Interference Channel With a Causal Relay Under Strong and Very Strong Interference," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 859–865, 2014.
 [6] A. El Gamal and Y.-H. Kim, *Network Information Theory*. New York: Cambridge University Press, 2011.