RESEARCH   **Open Access**

Capacity bounds for multiple access-cognitive interference channel

Mahtab Mirmohseni[*], Bahareh Akhbari and Mohammad Reza Aref

Abstract

Motivated by the uplink scenario in cellular cognitive radio, this study considers a communication network in which a point-to-point channel with a cognitive transmitter and a Multiple Access Channel (MAC) with common information share the same medium and interfere with each other. A Multiple Access-Cognitive Interference Channel (MA-CIFC) is proposed with three transmitters and two receivers, and its capacity region in different interference regimes is investigated. First, the inner bounds on the capacity region for the general discrete memoryless case are derived. Next, an outer bound on the capacity region for full parameter regime is provided. Using the derived inner and outer bounds, the capacity region for a class of *degraded* MA-CIFC is characterized. Two sets of strong interference conditions are also derived under which the capacity regions are established. Then, an investigation of the Gaussian case is presented, and the capacity regions are derived in the *weak* and *strong* interference regimes. Some numerical examples are also provided.

Keywords: Cognitive interference channel, Multiple access channel, Strong Interference, Weak interference, Capacity region

1. Introduction

Interference avoidance techniques have traditionally been used in wireless networks wherein multiple source-destination pairs share the same medium. However, the broadcasting nature of wireless networks may enable cooperation among entities, which ensures higher rates with more reliable communication. On the other hand, due to the increasing number of wireless systems, spectrum resources have become scarce and expensive. The exponentially growing demand for wireless services along with the rapid advancements in wireless technology has lead to cognitive radio technology which aims to overcome the spectrum inefficiency problem by developing communication systems that have the capability to sense the environment and adapt to it [1].

In overlay cognitive networks, the cognitive user can transmit simultaneously with the non-cognitive users and compensate for the interference by cooperation in sending, i.e., relaying, the non-cognitive users' messages [1]. From an information theoretic point of view, Cognitive Interference Channel (CIFC) was first introduced in [2]

to model an overlay cognitive radio and refers to a two-user Interference Channel (IFC) in which the cognitive user (secondary user) has the ability to obtain the message being transmitted by the other user (primary user), either in a non-causal or in a causal manner. An achievable rate region for the non-causal CIFC was derived in [2], by combining the Gel'fand-Pinsker (GP) binning [3] with a well-known simultaneous superposition coding scheme (rate splitting) applied to IFC [4]. For the non-causal CIFC, where the cognitive user has non-causal full or partial knowledge of the primary user's transmitted message several achievable rate regions and capacity results in some special cases have been established [5-14]. More recently a *three-user* cognitive radio network with one primary user and two cognitive users is studied in [15,16], where an achievable rate region is derived for this setup based on rate splitting and GP binning.

In the interference avoidance-based systems, i.e., when the communication medium is interference-free, uplink transmission is modeled with a Multiple Access Channel (MAC) whose capacity region has been fully characterized for independent transmitters [17,18] as well as for the transmitters with common information [19]. Recently,

* Correspondence: mirmohseni@ee.sharif.edu
Department of Electrical Engineering, Information Systems and Security Lab (ISSL) Sharif University of Technology, Tehran, Iran

Springer

taking the effects of interference into account in the uplink scenario, a MAC and an IFC have been merged into one setup by adding one more transmit-receive pair to the communication medium of a two-user MAC [20,21], where the channel inputs at the transmitters are independent and there is no cognition or cooperation.

In this paper, we introduce Multiple Access-Cognitive Interference Channel (MA-CIFC) by providing the transmitter of the point-to-point channel with *cognition capabilities* in the uplink with interference model. Moreover, transmitters of MAC have common information that enables cooperation among them. As shown in Figure 1, the proposed channe consists of three transmitters and two receivers: two-user MAC with common information as the primary network and a point-to-point channel with a cognitive transmitter that knows the message being sent by all of the transmitters in a non-causa manner. A physical example of this channel is the coexistence of cognitive users with the licensed primary users in a cellular or satellite uplink transmission, where the cognitive radios by their abilities exploit side information about the environment to maintain or improve the communication of primary users while also achieving some spectrum resources for their own communication. In this scenario, the primary non-cognitive users can be *oblivious* to the or *aware* of the cognitive users [1]. When the non-cognitive user is oblivious to the cognitive user's presence, its receiver's decoding process is independent of the interference caused by the cognitive user's transmission. In fact, the primary receiver treats interference as noise. However, in the aware cognitive user's scenario, the decoding process at the primary receiver can be adapted to improve its own rate. For example, the primary receiver can decode the cognitive user's message and cancel the interference when the interfering signal is strong enough. If the multi-antenna capability is available at the primary receiver, it can also

reduce or increase the interfering signal by beam-steering, which results in the occurrence of the weak or strong interference regimes [1].

To analyze the capacity region of MA-CIFC, we first derive three inner bounds on the capacity region (achievable rate regions). The first two bounds assume an oblivious primary receiver, which does not decode the cognitive user's message but treats it as noise. Two different coding schemes are proposed based on the superposition coding, the GP binning and the method of [6] in defining auxiliary Random Variables (RVs). Later, we show that these strategies are optimal for a *degraded* MA-CIFC and also in the Gaussian *weak* interference regime. In the third achievability scheme, we consider an aware primary receiver and obtain an inner bound on the capacity region based on using superposition coding in the encoding part and allowing both receivers to decode all messages with simultaneous joint decoding in the decoding part. This strategy is capacity-achieving in the strong interference regime. Next, we provide a general outer bound on the capacity region and derive conditions under which the first achievability scheme achieves capacity for the *degraded* MA-CIFC. We continue the capacity results by the derivation of two sets of strong interference conditions, under which the third inner bound achieves capacity. Further, we compare these two sets of conditions and identify the weaker set. We also extend the strong interference results to a network with k primary users.

Moreover, we consider the Gaussian case and find capacity results for the Gaussian MA-CIFC in both the *weak* and *strong* interference regimes. We use the second derived inner bound to show that the capacity-achieving scheme in weak interference consists of Dirty Paper Coding (DPC) [22] at the cognitive transmitter and treating interference as noise at both receivers. We also provide some numerical examples.

The rest of the paper is organized as follows. Section 2 introduces MA-CIFC model and the notations. Three inner bounds and an outer bound on the capacity region are derived in Section 3 and Section 4, respectively, for the discrete memoryless MA-CIFC. Sections 5 presents the capacity results for the discrete memoryless MA-CIFC in three special cases. In Section 6, the Gaussian MA-CIFC is investigated. Finally, Section 7 concludes the paper.

2. Channel models and preliminaries

Throughout the paper, upper case letters (e.g. X) are used to denote RVs and lower case letters (e.g. x) show their realizations. The probability mass function (p.m.f) of a RV X with alphabet set $\mathcal{X}$ is denoted by $p_X(x)$, where subscript X is occasionally omitted. $A_\varepsilon^n(X, Y)$ specifies the set of ϵ-strongly, jointly typical sequences of

Figure 1 Graphic representation for MA-CIFC.
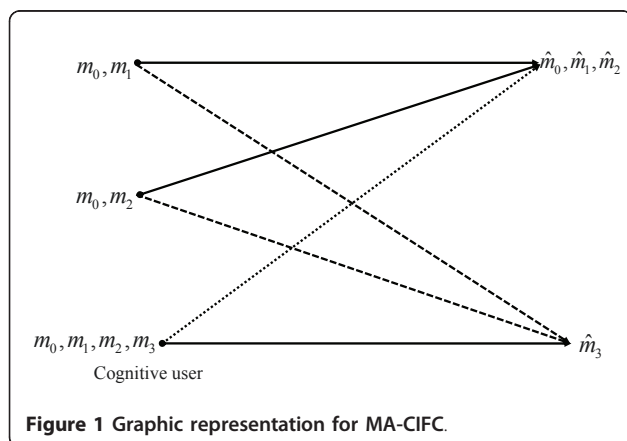
length n. The notation X_i^j indicates a sequence of RVs $(X_i, X_{i+1}, ..., X_j)$, where X^j is used instead of X_1^j, for brevity. $\mathcal{N}(0, \sigma^2)$ denotes a zero mean normal distribution with variance σ^2.

Consider the MA-CIFC in Figure 2, which is denoted by $(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3, p(y_1^n, y_3^n | x_1^n, x_2^n, x_3^n), \mathcal{Y}_1 \times \mathcal{Y}_3)$, where $X_1 \in \mathcal{X}_1, X_2 \in \mathcal{X}_2$ and $X_3 \in \mathcal{X}_3$ are channel inputs at Transmitter 1 (Tx1), Transmitter 2 (Tx2) and Transmitter 3 (Tx3), respectively; $Y_1 \in \mathcal{Y}_1$ and $Y_3 \in \mathcal{Y}_3$ are channel outputs at Receiver 1 (Rx1) and Receiver 3 (Rx3), respectively; and $p(y_1^n, y_3^n | x_1^n, x_2^n, x_3^n)$ is the channel transition probability distribution. In n channel uses, each Txj desires to send a message pair (m_0, m_j) to Rx1 where $j \in \{1,2\}$, and Tx3 desires to send a message m_3 to Rx3.

Definition 1: A $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR_3}, n)$ code for MA-CIFC consists of (i) four independent message sets $\mathcal{M}_j = \{1, ..., 2^{nR_j}\}$, where $j \in \{0, 1, 2, 3\}$; (ii) two encoding functions at the primary transmitters, $f_1 : \mathcal{M}_0 \times \mathcal{M}_1 \mapsto \mathcal{X}_1^n$ at Tx1 and $f_2 : \mathcal{M}_0 \times \mathcal{M}_2 \mapsto \mathcal{X}_2^n$ at Tx2; (iii) an encoding function at the cognitive transmitter, $f_3 : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_3 \mapsto \mathcal{X}_3^n$; and (iv) two decoding functions, $g_1 : \mathcal{Y}_1^n \mapsto \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2$ at Rx1 and $g_3 : \mathcal{Y}_3^n \mapsto \mathcal{M}_3$ at Rx3. We assume that the channel is memoryless. Thus, the channel transition probability distribution is given by

$$p(y_1^n, y_3^n | x_1^n, x_2^n, x_3^n) = \prod_{i=1}^{n} p(y_{1,i}, y_{3,i} | x_{1,i}, x_{2,i}, x_{3,i}). \quad (1)$$

The probability of error for this code is defined as

$$P_e = \frac{1}{2^n(R_0 + R_1 + R_2 + R_3)} \sum_{m_0, m_1, m_2, m_3} p[\{g_3(Y_3^n) \neq m_3\} \cup$$
$$\{g_1(Y_1^n) \neq (m_0, m_1, m_2)\} | (m_0, m_1, m_2, m_3) \text{ sent}].$$

Definition 2: A rate quadruple (R_0, R_1, R_2, R_3) is achievable if there exists a sequence of $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR_s}, n)$ codes with $P_e \to 0$ as $n \to \infty$. The capacity region $\mathcal{C}$, is the closure of the set of all achievable rates.

3. Inner bounds on the capacity region of discrete memoryless MA-CIFC

Now, we derive three achievable rate regions for the general setup. Theorems 1 and 2 assume an oblivious primary receiver (Rx1), which does not decode the cognitive user's message (m_3) and treats it as noise. The decoding procedure at the cognitive receiver (Rx3) differs in these schemes. In Theorem 1, the cognitive receiver (Rx3) decodes the primary messages (m_0, m_1, m_2), and all the transmitters use superposition coding. However, in Theorem 2, the cognitive receiver (Rx3) also treats the interference from the primary messages (m_0, m_1, m_2) as noise, while the cognitive transmitter (Tx3) uses GP binning to precode its message for interference cancelation at Rx3. We also utilize the method of [6] in defining auxiliary RVs, which helps us to achieve the outer bound in special cases. In fact, we achieve the outer bound of Theorem 4 using the region of Theorem
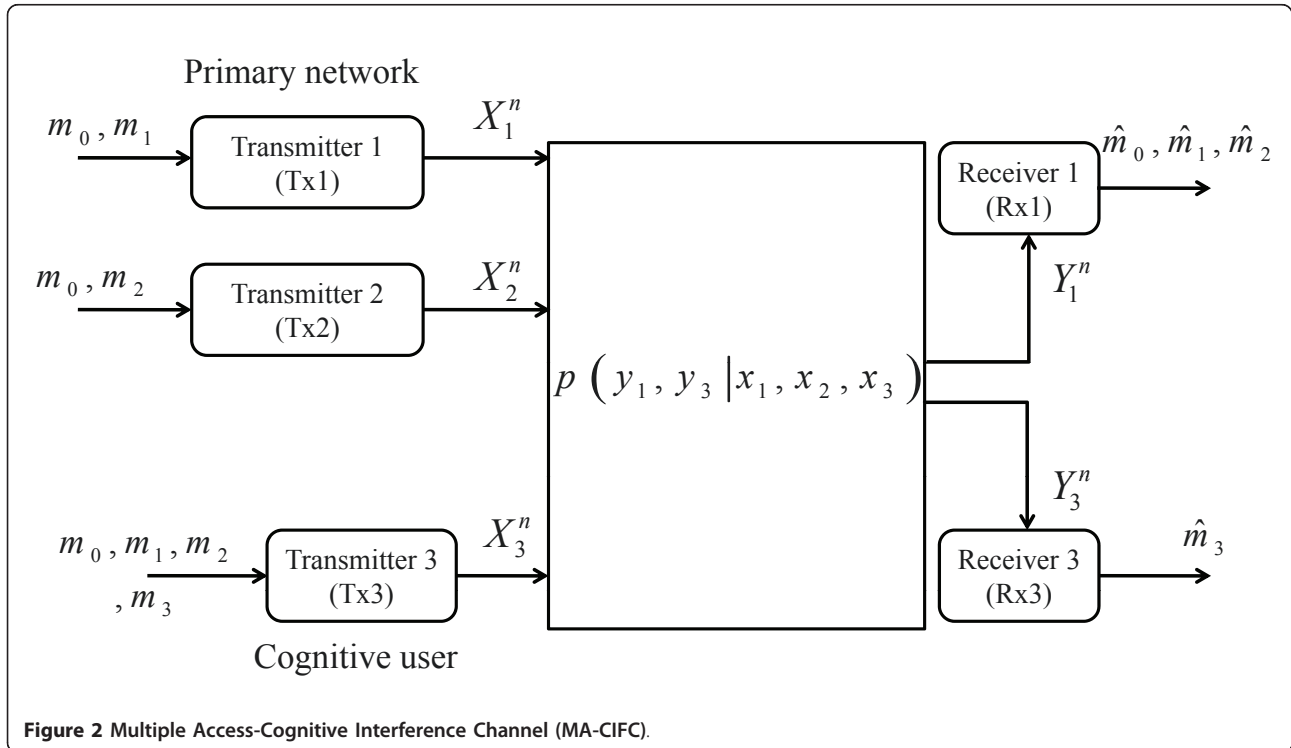
Figure 2 Multiple Access-Cognitive Interference Channel (MA-CIFC).

1 for a class of degraded MA-CIFC in Section 5. The region of Theorem 2 is used in Section 6 to derive the capacity region in the *weak* interference regime. In the scheme of Theorem 3, we consider an aware primary receiver (Rx1) which decodes the cognitive user's message (m_3). The cognitive receiver (Rx3) also decodes the primary messages (m_0, m_1, m_2). Therefore, this region is obtained based on using superposition coding in the encoding part and by allowing both receivers to decode all messages with simultaneous joint decoding in the decoding part. In Section 5, we show that this strategy is capacity-achieving in the strong interference regime. Proofs are provided in "Appendix A".

Theorem 1: The union of rate regions given by

$$R_3 \leq I(X_3; Y_3 | T, U, X_1, V, X_2) \tag{2}$$

$$R_1 \leq I(U, X_1; Y_1 | T, V, X_2) \tag{3}$$

$$R_2 \leq I(V, X_2; Y_1 | T, U, X_1) \tag{4}$$

$$R_1 + R_2 \leq I(U, X_1, V, X_2; Y_1 | T) \tag{5}$$

$$R_0 + R_1 + R_2 \leq I(T, U, X_1, V, X_2; Y_1) \tag{6}$$

$$R_1 + R_3 \leq I(U, X_1, X_3; Y_3 | T, V, X_2) \tag{7}$$

$$R_2 + R_3 \leq I(V, X_2, X_3; Y_3 | T, U, X_1) \tag{8}$$

$$R_1 + R_2 + R_3 \leq I(U, X_1, V, X_2, X_3; Y_3 | T) \tag{9}$$

$$R_0 + R_1 + R_2 + R_3 \leq I(T, U, X_1, V, X_2, X_3; Y_3) \tag{10}$$

is achievable for MA-CIFC, where the union is over all p.m.fs that factor as

$$p(t)p(u, x_1 | t)p(v, x_2 | t)p(x_3 | t, u, x_1, v, x_2). \tag{11}$$

Theorem 2: The union of rate regions given by (3)-(6) and

$$R_3 \leq I(W; Y_3) - I(W; T, U, X_1, V, X_2) \tag{12}$$

is achievable for MA-CIFC, where the union is over all p.m.fs that factor as

$$p(t)p(u, x_1 | t)p(v, x_2 | t)p(w, x_3 | t, u, x_1, v, x_2). \tag{13}$$

Theorem 3: The union of rate regions given by

$$R_3 \leq I(X_3; Y_3 | X_1, X_2, T) \tag{14}$$

$$R_1 + R_3 \leq \min\{I(X_1, X_3; Y_1 | X_2, T),$$
$$\qquad\qquad I(X_1, X_3; Y_3 | X_2, T)\} \tag{15}$$

$$R_2 + R_3 \leq \min\{I(X_2, X_3; Y_1 | X_1, T),$$
$$\qquad\qquad I(X_2, X_3; Y_3 | X_1, T)\} \tag{16}$$

$$R_0 + R_1 + R_2 + R_3 \leq \min\{I(X_1, X_2, X_3; Y_1),$$
$$\qquad\qquad I(X_1, X_2, X_3; Y_3)\} \tag{17}$$

is achievable for MA-CIFC, where the union is over all p.m.fs that factor as

$$p(t)p(x_1 | t)p(x_2 | t)p(x_3 | x_1, x_2, t). \tag{18}$$

Remark 1: We utilize the region of Theorem 1 in Section 5 to achieve capacity results for a class of degraded MA-CIFC, and the region of Theorem 2 to derive the results for the Gaussian case in Section 6. The region of Theorem 3 is also used to characterize the capacity region under strong interference conditions in Section 5.

4. An outer bound on the capacity region of discrete memoryless MA-CIFC

Here, we derive a general outer bound on the capacity region of MA-CIFC which is used to obtain the capacity region for a class of the degraded MA-CIFC in Section 5 and also to find capacity results for the Gaussian MA-CIFC in the weak interference regime in Section 6. Let $\mathcal{R}_o^1$ denote the union of all rate quadruples (R_0, R_1, R_2, R_3) satisfying (3)-(6) and

$$R_3 \leq I(X_3; Y_3, Y_1 | T, U, X_1, V, X_2), \tag{19}$$

where the union is over all p.m.fs that factor as (11).

Theorem 4: The capacity region of MA-CIFC satisfies

$$\mathcal{C} \subseteq \mathcal{R}_o^1$$

Proof: Consider a $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR_3}, n)$ code with the average error probability of $P_e^n \rightarrow 0$. Define the following RVs for $i = 1, ..., n$:

$$T_i = (M_0, Y_1^{i-1}) \tag{20}$$

$$U_i = (M_0, M_1, Y_1^{i-1}) = (M_1, T_i) \tag{21}$$

$$V_i = (M_0, M_2, Y_1^{i-1}) = (M_2, T_i) \tag{22}$$

Considering the encoding functions f_1 and f_2, defined in Definition 1, and the above definitions for auxiliary

RVs, we remark that $(X_{1,i}, U_i) \to T_i \to (X_{2,i}, V_i)$ forms a Markov chain. Thus, these choices of auxiliary RVs satisfy the p.m.f (11) of Theorem 4. Now using Fano's inequality [23], we derive the bounds in Theorem 4. For the first bound, we have:

$$
\begin{aligned}
nR_3 = H(M_3) &\overset{(a)}{=} H(M_3|M_0, M_1, M_2) \\
&= I(M_3; Y_3^n|M_0, M_1, M_2) + H(M_3|Y_3^n, M_0, M_1, M_2) \\
&\overset{(b)}{\leq} I(M_3; Y_3^n|M_0, M_1, M_2) + n\delta_{3n}
\end{aligned}
\tag{23}
$$

where (a) follows since messages are independent and (b) holds due to Fano's inequality and the fact that conditioning does not increase entropy. Hence,

$$
\begin{aligned}
nR_3 - n\delta_{3n} &\leq I(M_3; Y_3^n|M_0, M_1, M_2) \\
&\overset{(a)}{\leq} I(M_3, X_3^n; Y_3^n, Y_1^n|M_0, M_1, M_2, X_1^n, X_2^n) \\
&\overset{(b)}{=} \sum_{i=1}^{n} I\left(M_3, X_3^n; Y_{3,i}, Y_{1,i}|M_0, M_1, M_2, X_1^n, X_2^n, Y_3^{i-1}, Y_1^{i-1}\right) \\
&\overset{(c)}{\leq} \sum_{i=1}^{n} H\left(Y_{3,i}, Y_{1,i}|M_0, M_1, M_2, X_{1,i}, X_{2,i}, Y_1^{i-1}\right) \\
&\qquad - H(Y_{3,i}, Y_{1,i}|M_0, M_1, M_2, Y_1^{i-1}, X_{1,i}, X_{2,i}, X_{3,i}) \\
&\overset{(d)}{=} \sum_{i=1}^{n} I(X_{3,i}; Y_{3,i}, Y_{1,i}|T_i, U_i, X_{1,i}, V_i, X_{2,i})
\end{aligned}
\tag{24}
$$

where (a) is due to the encoding functions f_1, f_2 and f_3, defined in Definition 1, and the non-negativity of mutual information, (b) is obtained from the chain rule, (c) follows from the memoryless property of the channel and the fact that conditioning does not increase entropy, and (d) is obtained from (20)-(22).

Now, applying Fano's inequality and the independence of the messages, we can bound R_1 as:

$$
\begin{aligned}
nR_1 - n\delta_{1n} &\leq I(M_1; Y_1^n|M_0, M_2) \\
&\overset{(a)}{=} \sum_{i=1}^{n} I(M_1, X_{1,i}; Y_{1,i}|M_0, M_2, X_{2,i}, Y_1^{i-1}) \\
&= \sum_{i=1}^{n} I(M_1, X_{1,i}, M_0, Y_1^{i-1}; Y_{1,i}|M_0, M_2, X_{2,i}, Y_1^{i-1}) \\
&\overset{(b)}{=} \sum_{i=1}^{n} I(U_i, X_{1,i}; Y_{1,i}|T_i, V_i, X_{2,i}),
\end{aligned}
\tag{25}
$$

where (a) follows from the chain rule and the encoding functions f_1 and f_2, and (b) from (20)-(22). Similarly, we can show that

$$
nR_2 - n\delta_{2n} \leq \sum_{i=1}^{n} I(V_i, X_{2,i}; Y_{1,i}|T_i, U_i, X_{1,i}).
\tag{26}
$$

Next, based on similar arguments, we bound $R_1 + R_2$ as

$$
\begin{aligned}
n(R_1 + R_2) - n(\delta_{1n} + \delta_{2n}) &\leq I(M_1, M_2; Y_1^n|M_0) \\
&= \sum_{i=1}^{n} I(M_1, X_{1,i}, M_2, X_{2,i}; Y_{1,i}|M_0, Y_1^{i-1}) \\
&= \sum_{i=1}^{n} I(M_1, X_{1,i}, M_2, X_{2,i}, Y_1^{i-1}; Y_{1,i}|M_0, Y_1^{i-1}) \\
&= \sum_{i=1}^{n} I(U_i, X_{1,i}, V_i, X_{2,i}; Y_{1,i}|T_i).
\end{aligned}
\tag{27}
$$

The last sum-rate bound can be derived as follows:

$$
\begin{aligned}
n(R_0 + R_1 + R_2) - n(\delta_{0n} + \delta_{1n} + \delta_{2n}) &\leq I(M_0, M_1, M_2; Y_1^n) \\
&= \sum_{i=1}^{n} I(M_0, M_1, X_{1,i}, M_2, X_{2,i}; Y_{1,i}|Y_1^{i-1}) \\
&\overset{(a)}{\leq} \sum_{i=1}^{n} H(Y_{1,i}) - H(Y_{1,i}|M_0, M_1, X_{1,i}, M_2, X_{2,i}, Y_1^{i-1}) \\
&= \sum_{i=1}^{n} I(T_i, U_i, X_{1,i}, V_i, X_{2,i}; Y_{1,i})
\end{aligned}
\tag{28}
$$

where (a) follows since conditioning does not increase entropy. Using the standard time-sharing argument for (24)-(28) completes the proof.

5. Capacity results for discrete memoryless MA-CIFC

In this section, we characterize the capacity region of MA-CIFC under specific conditions. First, we consider a class of degraded MA-CIFC and derive conditions under which the inner bound in Theorem 1 achieves the outer bound of Theorem 4. Next, we investigate the strong interference regime by deriving two sets of strong interference conditions under which the region of Theorem 3 achieves capacity. We also compare these two sets of conditions and identify the weaker set. Finally, we extend the strong interference results to a network with k primary users.

A. Degraded MA-CIFC

Now, we characterize the capacity region for a class of MA-CIFC with a *degraded* primary receiver. We define MA-CIFC with a degraded primary receiver as a MA-CIFC where Y_1 and X_3 are independent given Y_3, X_1, X_2. More precisely, the following Markov chain holds:

$$
X_3|X_1, X_2 \to Y_3|X_1, X_2 \to Y_1|X_1, X_2,
\tag{29}
$$

or equivalently, $X_3 \to (X_1, X_2, Y_3) \to Y_1$ forms a Markov chain. This means that the primary receiver (Rx1) observes a degraded or noisier version of the cognitive user's signal (Tx3) compared with the cognitive receiver (Rx3).

Assume that the following conditions are satisfied for MA-CIFC over all p.m.fs that factor as (11):

$$I(U, X_1; Y_1 | T, V, X_2) \leq I(U, X_1; Y_3 | T, V, X_2) \tag{30}$$

$$I(V, X_2; Y_1 | T, U, X_1) \leq I(V, X_2; Y_3 | T, U, X_1) \tag{31}$$

$$I(U, X_1, V, X_2; Y_1 | T) \leq I(U, X_1, V, X_2; Y_3 | T) \tag{32}$$

$$I(T, U, X_1, V, X_2; Y_1) \leq I(T, U, X_1, V, X_2; Y_3) \tag{33}$$

Under these conditions, the cognitive receiver (Rx3) can decode the messages of the primary users with no rate penalty. If MA-CIFC with a degraded primary receiver satisfies conditions (30)-(33), the region of Theorem 1 coincides with $\mathcal{R}_o^1$ and achieves capacity, as stated in the following theorem.

Theorem 5: The capacity region of MA-CIFC with a degraded primary receiver, defined in (29), satisfying (30)-(33) is given by the union of rate regions satisfying (2)-(6) over all joint p.m.fs (11).

Remark 2: The messages of the primary users (m_0, m_1, m_2) can be decoded at Rx3 under conditions (30)-(33). Therefore, Rx3-Tx3 achieves the rate in (2). Moreover, we can see that due to the degradedness condition in (29), treating interference as noise at the primary receiver (Rx1) achieves capacity. We show in Section 6 that, in the Gaussian case the capacity is achieved by using the region of Theorem 2 based on DPC (or GP binning), where the cognitive receiver (Rx3) does not decode the primary messages and conditions (30)-(33) are not necessary.

Proof: <u>Achievability:</u> The proof follows from the region of Theorem 1. Using the condition in (30), the sum of the bounds in (2) and (3) makes the bound in (7) redundant. Similarly, conditions (31)-(33), along with the bound in (2), make the bounds in (8)-(10) redundant and the region reduces to (2)-(6).

<u>Converse:</u> To prove the converse part, we evaluate $\mathcal{R}_o^1$ of Theorem 4 with the degradedness condition in (29). It is noted that the p.m.f of Theorem 5 is the same as the one for $\mathcal{R}_o^1$. Moreover, the bounds in (3)-(6) are equal for both regions. Hence, it is only necessary to show the bound in (2). Considering (19), we obtain:

$$\begin{aligned} R_3 &\leq I(X_3; Y_3, Y_1 | T, U, X_1, V, X_2) \\ &= I(X_3; Y_3 | T, U, X_1, V, X_2) + I(X_3; Y_1 | T, U, X_1, V, X_2, Y_3) \\ &\overset{(a)}{=} I(X_3; Y_3 | T, U, X_1, V, X_2) \end{aligned}$$

where (a) is obtained by applying the degradedness condition in (29). This completes the proof.

B. Strong interference regime

Now, we derive two sets of strong interference conditions under which the region of Theorem 3 achieves capacity. First, assume that the following set of strong interference conditions, referred to as *Set1*, holds for all p.m.fs that factor as (18):

$$I(X_3; Y_3 | X_1, X_2, T) \leq I(X_3; Y_1 | X_1, X_2, T) \tag{34}$$

$$I(X_1, X_3; Y_1 | X_2, T) \leq I(X_1, X_3; Y_3 | X_2, T) \tag{35}$$

$$I(X_2, X_3; Y_1 | X_1, T) \leq I(X_2, X_3; Y_3 | X_1, T) \tag{36}$$

$$I(X_1, X_2, X_3; Y_1) \leq I(X_1, X_2, X_3; Y_3). \tag{37}$$

In fact, under these conditions, interfering signals at the receivers are strong enough that all messages can be decoded by both receivers. Condition (34) implies that the cognitive user's message (m_3) can be decoded at Rx1, while conditions (35)-(37) guarantee the decoding of the primary messages (m_0, m_1, m_2) along with m_3 at Rx3 in a MAC fashion.

Theorem 6: The capacity region of MA-CIFC satisfying (34)-(37) is given by:

$$\begin{aligned} \mathcal{C}_1^{\text{str}} = \bigcup_{p(t)p(x_1|t)p(x_2|t)p(x_3|x_1,x_2,t)} & \left\{ (R_0, R_1, R_2, R_3) : \right. \\ R_0, R_1, R_2, R_3 &\geq 0 \\ R_3 &\leq I(X_3; Y_3 | X_1, X_2, T) \tag{38} \\ R_1 + R_3 &\leq I(X_1, X_3; Y_1 | X_2, T) \tag{39} \\ R_2 + R_3 &\leq I(X_2, X_3; Y_1 | X_1, T) \tag{40} \\ R_0 + R_1 + R_2 + R_3 &\leq \left. I(X_1, X_2, X_3; Y_1) \right\}. \tag{41} \end{aligned}$$

Remark 3: The message of the cognitive user (m_3) can be decoded at Rx1, under condition (34) and (m_0, m_1, m_2) can be decoded at Rx3 under conditions (35)-(37). Hence, the bound in (38) gives the capacity of a point-to-point channel with message m_3 with side-information X_1, X_2 at the receiver. Moreover, (38)-(41) with condition (34) give the capacity region for a three-user MAC with common information where R_1 and R_2 are the common rates, R_3 is the private rate for Tx3, and the private rates for Tx1 and Tx2 are zero.

Remark 4: If we omit Tx2, i.e., $X_2 = \emptyset$, and Tx2 has no message to transmit, i.e., $R_2 = 0$, the model reduces to a CIFC, and $\mathcal{C}_1^{\text{str}}$ coincides with the capacity region of the strong interference channel with unidirectional cooperation (or CIFC), which was characterized in [8, Theorem 5]. It is noted that in this case, the common message can be ignored, i.e., $T = \emptyset$ and $R_0 = 0$.

Proof: <u>Achievability:</u> Considering (35)-(37), the proof follows from Theorem 3.

<u>Converse:</u> Consider a $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR_3}, n)$ code with an average error probability of $P_e^n \to 0$. Define the following RV for $i = 1, ..., n$:

$$T^n = M_0 \tag{42}$$

It is noted that due to the encoding functions f_1, f_2 and f_3, defined in Definition 1, the independence of messages, and the above definitions for T^n, RVs satisfy the p.m.f (18) of Theorem 6. First, we provide a useful lemma which we need in the proof of the converse part.

Lemma 1: If (34) holds for all distributions that factor as (18), then

$$I(X_3^n; Y_3^n | X_1^n, X_2, T^n, U) \leq I(X_3^n; Y_1^n | X_1^n, X_2^n, T^n, U). \quad (43)$$

Proof: The proof relies on the results in [24, Proposition 1] and [25, Lemma]. By redefining $X_2 = X_3$, $Y_2 = Y_3$, $X_1 = (X_1, X_2, T)$ in [8, Lemma 5], the proof follows.

Now, using Fano's inequality [23], we derive the bounds in Theorem 6. Using (23) provides:

$$
\begin{aligned}
nR_3 - n\delta_{3n} &\leq I(M_3; Y_3^n | M_0, M_1, M_2) \\
&\overset{(a)}{=} I(M_3, X_3^n; Y_3^n | T^n, M_1, M_2, X_1^n, X_2^n) \\
&\overset{(b)}{\leq} I(X_3^n; Y_3^n | T^n, X_1^n, X_2^n) \\
&\overset{(c)}{=} \sum_{i=1}^n I(X_3^n; Y_{3,i} | T^n, X_1^n, X_2^n, Y_3^{i-1}) \\
&\overset{(d)}{\leq} \sum_{i=1}^n I(X_{3,i}; Y_{3,i} | X_{1,i}, X_{2,i}, T_i)
\end{aligned}
\quad (44)
$$

where (a) is due to (42) and the encoding functions f_1, f_2 and f_3, defined in Definition 1, (b) follows from two facts; conditioning does not increase entropy and $(M_1, M_2, M_3) \to (X_1^n, X_2^n, X_3^n) \to Y_3^n$ forms a Markov chain, (c) is obtained from the chain rule, and (d) follows from the memoryless property of the channel and the fact that conditioning does not increase entropy.

Now, applying Fano's inequality and the independence of the messages, we can bound $R_1 + R_3$ as

$$
\begin{aligned}
n(R_1 + R_3) - n(\delta_{1n} + \delta_{3n}) &\leq \\
I(M_1; Y_1^n | M_0, M_2) &+ I(M_3; Y_3^n | M_0, M_1, M_2) \\
\overset{(a)}{=} I(M_1, X_1^n; Y_1^n | M_0, M_2, X_2^n) \\
&+ I(M_3, X_3^n; Y_3^n | M_0, M_1, M_2, X_1^n, X_2^n) \\
\overset{(b)}{=} I(M_1, X_1^n; Y_1^n | T^n, M_2, X_2^n) \\
&+ I(X_3^n; Y_3^n | T^n, M_1, M_2, X_1^n, X_2^n) \\
\overset{(c)}{\leq} I(M_1, X_1^n; Y_1^n | T^n, M_2, X_2^n) \\
&+ I(X_3^n; Y_1^n | T^n, M_1, M_2, X_1^n, X_2^n) \\
= I(M_1, X_1^n, X_3^n; Y_1^n | T^n, M_2, X_2^n) \\
\overset{(d)}{=} \sum_{i=1}^n I(M_1, X_1^n, X_3^n; Y_{1,i} | T^n, M_2, X_2^n, Y_1^{i-1}) \\
\overset{(e)}{\leq} \sum_{i=1}^n I(X_{1,i}, X_{3,i}; Y_{1,i} | X_{2,i}, T_i)
\end{aligned}
\quad (45)
$$

where (a) follows from encoding functions f_1, f_2 and f_3, (b) follows from (42) and the fact that $M_3 \to (X_1^n, X_2^n, X_3^n) \to Y_3^n$ forms a Markov chain, (c) is obtained from (43), (d) follows from the chain rule, and (e) follows from the memoryless property of the channel and the fact that conditioning does not increase entropy.

Applying similar steps, we can show that,

$$n(R_2 + R_3) - n(\delta_{2n} + \delta_{3n}) \leq \sum_{i=1}^n I(X_{2,i}, X_{3,i}; Y_{1,i} | X_{1,i}, T_1). \quad (46)$$

Finally, the sum-rate bound can be obtained as

$$
\begin{aligned}
n(R_0 + R_1 + R_2 + R_3) &- n(\delta_{0n} + \delta_{1n} + \delta_{2n} + \delta_{3n}) \\
&\leq I(M_0, M_1, M_2; Y_1^n) + I(M_3; Y_3^n | M_0, M_1, M_2) \\
&= I(M_0, M_1, M_2, X_1^n, X_2^n; Y_1^n) \\
&\quad + I(M_3, X_3^n; Y_3^n | M_0, M_1, M_2, X_1^n, X_2^n) \\
&\overset{(a)}{\leq} I(T^n, M_1, M_2, X_1^n, X_2^n; Y_1^n) \\
&\quad + I(X_3^n; Y_1^n | T^n, M_1, M_2, X_1^n, X_2^n) \\
&= I(T^n, M_1, M_2, X_1^n, X_2^n, X_3^n; Y_1^n) \\
&\overset{(b)}{=} I(T^n, X_1^n, X_2^n, X_3^n; Y_1^n) \\
&\overset{(c)}{=} \sum_{i=1}^n I(X_{1,i}, X_{2,i}, X_{3,i}; Y_{1,i})
\end{aligned}
\quad (47)
$$

where (a) follows from steps (a)-(c) in (45), (b) is due to the fact that $(M_1, M_2) \to (X_1^n, X_2^n, X_3^n) \to Y_1^n$ forms a Markov chain, and (c) follows from the memoryless property of the channel and the fact that conditioning does not increase entropy. Using a standard time-sharing argument for (44)-(47) completes the proof.

Next, we derive the second set of strong interference conditions, called *Set2*, under which the region of Theorem 3 is the capacity region. For all p.m.fs that factor as (18), *Set2* includes (34) and the following conditions:

$$I(X_1; Y_1 | X_2, T) \leq I(X_1; Y_3 | X_2, T) \quad (48)$$

$$I(X_2; Y_1 | X_1, T) \leq I(X_2; Y_3 | X_1, T) \quad (49)$$

$$I(X_1, X_2; Y_1) \leq I(X_1, X_2; Y_3). \quad (50)$$

Remark 5: Similar to the condition *Set1*, under these conditions interfering signals at the receivers are strong enough that all messages can be decoded by both receivers. The first condition in (34) is equal in the two sets under which the cognitive user's message (m_3) can be decoded at Rx1. However, conditions (48)-(50) imply that the primary messages (m_0, m_1, m_2) can be decoded at Rx3 in a MAC fashion, while in *Set1*, they can be decoded along with m_3.

Theorem 7: The capacity region of MA-CIFC, satisfying (34) and (48)-(50), referred to as $\mathcal{C}_2^{\text{str}}$, is given by the union of rate regions satisfying (14)-(17) over all p.m.fs that factor as (18).

Proof: See "Appendix B".

Remark 6: Similar to Remark 4, by omitting Tx2 ($T = X_2 = \emptyset$, $R_0 = R_2 = 0$), the model reduces to a CIFC. Moreover, $\mathcal{C}_2$ and *Set2* reduce to the capacity region and strong interference conditions which have been derived in [13] for non-causal CIFC.

Remark 7 (Comparison of two sets of conditions): In the strong interference conditions of *Set1*, the first condition in (34) is used in the converse part, while (35)-(37) are used to reduce the inner bound to $\mathcal{C}_1^{\text{str}}$. However, all the conditions of *Set2* are utilized to prove the converse part. Now, we compare the conditions in these two sets. We can write (35) as

$$I(X_1; Y_1|X_2, T) + \underbrace{[I(X_3; Y_1|X_1, X_2, T) - I(X_3; Y_3|X_1, X_2, T)]}_{I_{\text{diff}}}$$
$$\leq I(X_1; Y_3|X_2, T).$$

Considering (34), it can be seen that $I_{\text{diff}} \geq 0$. Hence, condition (35) implies condition (48), but not vice versa. Similar conclusions can be drawn for other conditions of these two sets. Therefore, *Set1* implies *Set2*, and the conditions of *Set2* are weaker compared to those of *Set1*.

C. Multiple access-cognitive interference network (MA-CIFN)

Now, we extend the result of Theorem 6 to a network with $k + 1$ transmitters and two receivers; a k-user MAC as a primary network and a point-to-point channel with a cognitive transmitter. We call it Multiple Access-Cognitive Interference Network (MA-CIFN). Consider MA-CIFN in Figure 3, denoted by $(\mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_k \times \mathcal{X}_{k+1}, p(y_1^n, y_{k+1}^n|x_1^n, x_2^n, \cdots x_k^n, x_{k+1}^n), \mathcal{Y}_1 \times \mathcal{Y}_{k+1})$, where $X_j \in \mathcal{X}_j$ is the channel input at Transmitter j (Txj), for $j \in \{1, ..., k + 1\}$; $Y_1 \in \mathcal{Y}_1$ and $Y_{k+1} \in \mathcal{Y}_{k+1}$ are channel outputs at the primary and cognitive receivers, respectively, and $p(y_1^n, y_{k+1}^n|n_1^n, x_2^n, ..., x_k^n, x_{k+1}^n)$ is the channel transition probability distribution. In n channel uses, each Txj desires to send a message pair m_j to the primary receiver where $j \in \{1, ..., k\}$, and Tx$k + 1$ desires to send a message m_{k+1} to the cognitive receiver. We ignore the common information for brevity. Definitions 1 and 2 can be simply extended to the MA-CIFN. Therefore, we state the result on the capacity region under strong interference conditions.

Corollary 1: The capacity region of the MA-CIFN, satisfying

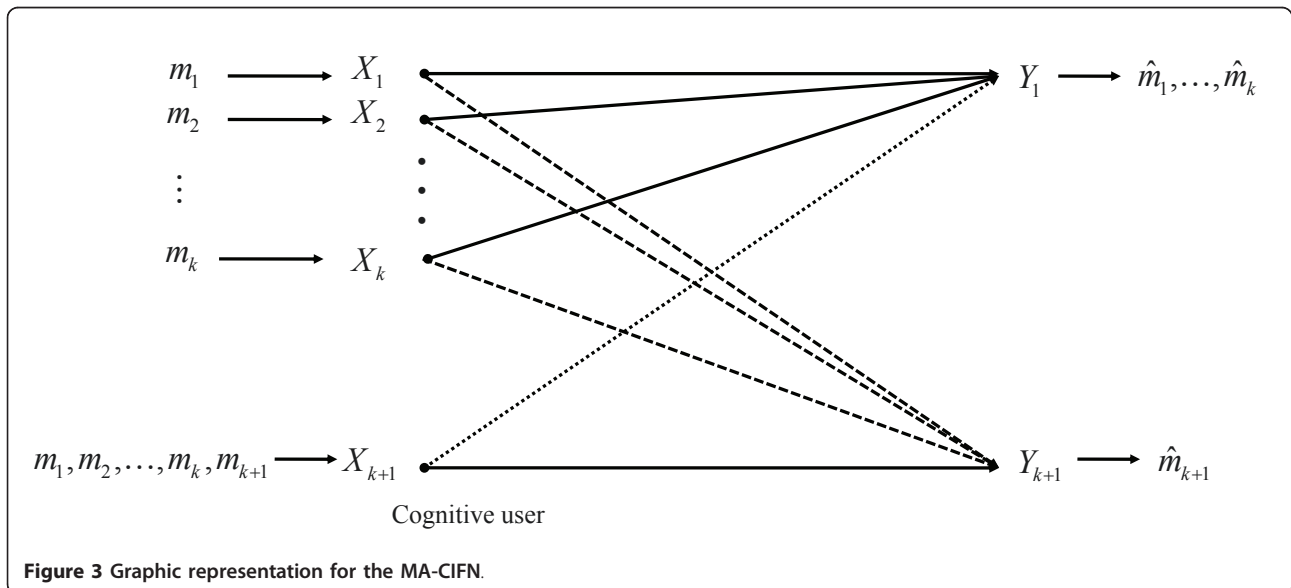$$I(X_{k+1}; Y_{k+1}|X([1:k])) \leq I(X_{k+1}; Y_1|X([1:k])) \qquad (51)$$

$$I(X_{k+1}, X(S); Y_1|X(S^c)) \leq I(X_{k+1}, X(S); Y_{k+1}|X(S^c)) \qquad (52)$$

for all $S \subseteq [1: k]$ and for every $p(x_1)p(x_2)...p(x_k) \, p(x_{k+1}|x_1,x_2,...,x_k)p(y_1,y_{k+1}|x_1,x_2,...,x_k,x_{k+1})$, is given by

$$\mathcal{C}_{net}^{\text{str}} = \bigcup_{p(x_1)p(x_2)...p(x_k)p(x_{k+1}|x_1,x_2,...,x_k)} \{$$
$$(R_1, R_2, \ldots, R_k, R_{k+1}) : R_1, R_2, \ldots, R_k, R_{k+1} \geq 0$$
$$R_{k+1} \leq I(X_{k+1}; Y_{k+1}|X([1:k])) \qquad (53)$$
$$R_{k+1} + \sum_{j \in S} R_j \leq I(X_{k+1}, X(S); Y_1|X(S^c))\} \qquad (54)$$

for all $S \subseteq [1: k]$, where $X(S)$ is the ordered vector of $X_j, j \in S$, and S^c denotes the complement of the set S.

Figure 3 Graphic representation for the MA-CIFN.

Proof: Following the same lines as the proof of Theorem 6, the proof is straightforward. Therefore, it is omitted for the sake of brevity.

Remark 8: Under condition (51), the message of the cognitive user (m_{k+1}) can be decoded at the primary receiver (Y_1). Also, the cognitive receiver (Y_{k+1}), under condition (52), can decode m_j; $j \in \{1,...,k\}$ in a MAC fashion. Therefore, the bound in (53) gives the capacity of a point-to-point channel with message m_{k+1} with side-information X_j; $j \in \{1,..., k\}$ at the cognitive receiver. Moreover, (53) and (54) with condition (51), give the capacity region for a $k + 1$-user MAC with common information at the primary receiver.

6. Gaussian MA-CIFC

In this section, we consider the Gaussian MA-CIFC and characterize capacity results for the Gaussian case in the *weak* and *strong* interference regimes. For simplicity, we assume that Tx1 and Tx2 have no common information. This means that, $R_0 = 0$ and $\mathcal{M}_0 = \emptyset$. to investigate these regions. The Gaussian MA-CIFC, as depicted in Figure 4, at time $i = 1,..., n$ can be mathematically modeled as

$$Y_{1,i} = X_{1,i} + X_{2,i} + h_{31}X_{3,i} + Z_{1,i} \qquad (55)$$

$$Y_{3,i} = h_{13}X_{1,i} + h_{23}X_{2,i} + X_{3,i} + Z_{3,i} \qquad (56)$$

where h_{31}, h_{13}, and h_{23} are known channel gains. $X_{1,i}$, $X_{2,i}$ and $X_{3,i}$ are input signals with average power constraints:

$$\frac{1}{n}\sum_{i=1}^{n} (x_{j,i})^2 \le P_j \qquad (57)$$

for $j \in \{1,2, 3\}$. $Z_{1,i}$ and $Z_{3,i}$ are independent and identically distributed (i.i.d) zero mean Gaussian noise components with unit powers, i.e., $Z_{j,i} \sim \mathcal{N}(0, 1)$ for $j \in \{1, 3\}$.

A. Strong interference regime

Here, we extend the results of Theorem 6, i.e., $\mathcal{C}_1^{str}$ and *Set1*, to the Gaussian case. The strong interference conditions of *Set1*, i.e., (34)-(37), for the above Gaussian model become:

$$h_{31}^2 \ge 1 \qquad (58)$$

$$P_1(h_{13}^2 - 1) + 2\rho_1\sqrt{P_1 P_3}(h_{13} - h_{31}) \ge P_3(1 - \rho_2^2)(h_{31}^2 - 1) \qquad (59)$$

$$P_2(h_{23}^2 - 1) + 2\rho_2\sqrt{P_2 P_3}(h_{23} - h_{31}) \ge P_3(1 - \rho_1^2)(h_{31}^2 - 1) \qquad (60)$$
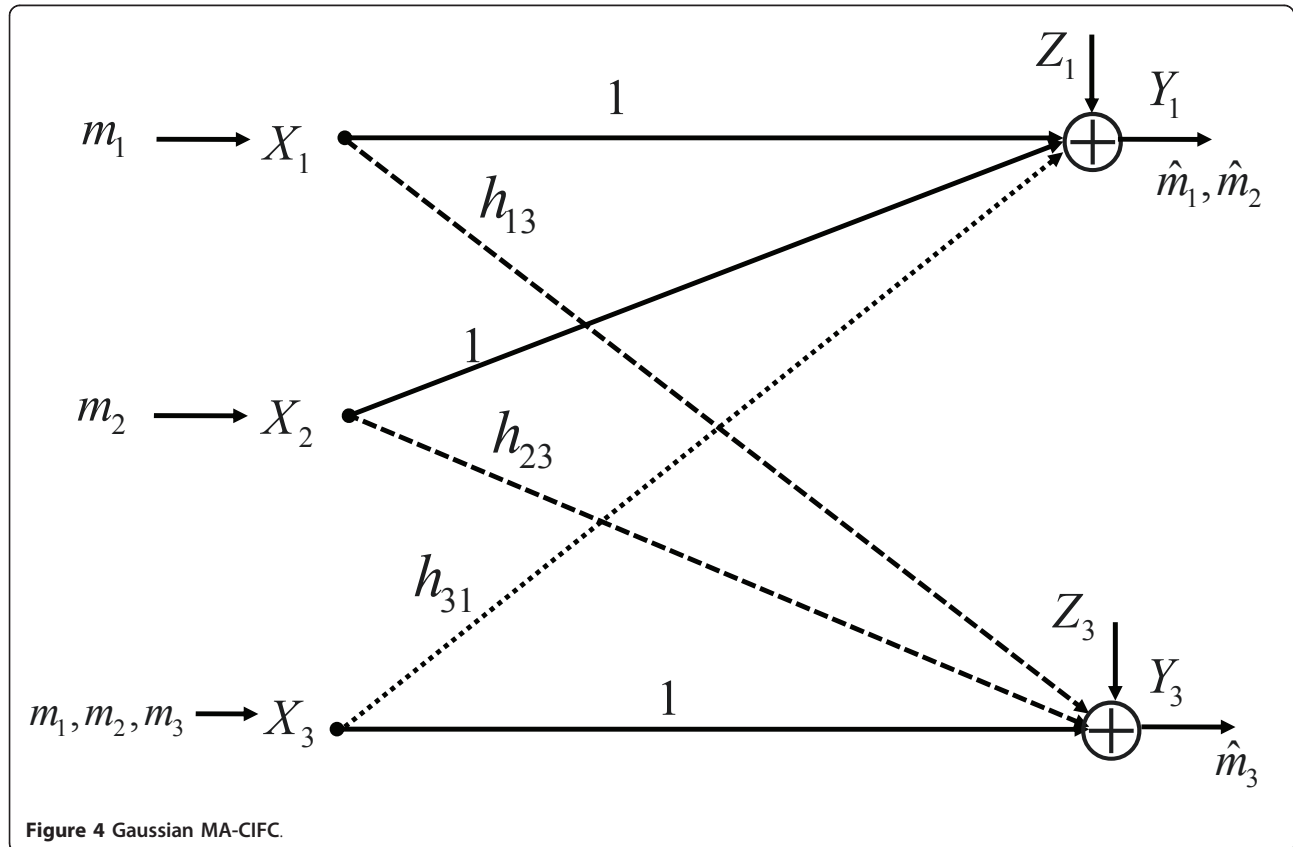

Figure 4 Gaussian MA-CIFC.

$$P_1(h_{13}^2 - 1) + P_2(h_{23}^2 - 1) + 2\rho_1\sqrt{P_1P_3}(h_{13} - h_{31})$$
$$+2\rho_2\sqrt{P_2P_3}(h_{23} - h_{31}) \geq P_3(h_{31}^2 - 1) \quad (61)$$

where $-1 \leq \rho_u \leq 1$ is the correlation coefficient between X_u and X_3, i.e., $E(X_u, X_3) = \rho_u\sqrt{P_uP_3}$ for $u \in \{1, 2\}$.

Theorem 8: For the Gaussian MA-CIFC satisfying conditions (58)-(61), the capacity region is given by

$$\mathcal{C}_1^G = \bigcup_{-1\leq\rho_1,\rho_2\leq1:\rho_1^2+\rho_2^2\leq1} \{(R_1,R_2,R_3) : R_1,R_2,R_3 \geq 0 \quad (62)$$
$$R_3 \leq \theta(P_3(1 - \rho_1^2 - \rho_2^2))$$

$$R_1 + R_3 \leq \theta\left(P_1 + h_{31}^2 P_3(1 - \rho_2^2) + 2h_{31}\rho_1\sqrt{P_1P_3}\right) \ (63)$$

$$R_2 + R_3 \leq \theta\left(P_2 + h_{31}^2 P_3(1 - \rho_1^2) + 2h_{31}\rho_2\sqrt{P_2P_3}\right) \ (64)$$

$$R_1 + R_2 + R_3 \leq$$
$$\theta\left(P_1 + P_2 + h_{31}^2 P_3 + 2h_{31}\sqrt{P_3}(\rho_1\sqrt{P_1} + \rho_2\sqrt{P_2})\right)\} \quad (65)$$

where, to simplify notation, we define

$$\theta(x) \doteq \frac{1}{2}\log(1 + x). \quad (66)$$

Remark 9: Condition (58) implies that Tx3 causes strong interference at Rx1. This enables Rx1 to decode m_3. Moreover, (59)-(61) provide strong interference conditions at Rx3, under which all messages can be decoded in Rx3 in a MAC fashion.

Proof: The achievability part follows from $\mathcal{C}_1^{str}$ in Theorem 6 by evaluating (38)-(41) with zero mean jointly Gaussian channel inputs X_1, X_2 and X_3. That is, $X_1 \sim \mathcal{N}(0, P_1), X_2 \sim \mathcal{N}(0, P_2)$, and $X_3 \sim \mathcal{N}(0, P_3)$, where $E(X_1, X_2) = 0$, $E(X_1, X_3) = \rho_1\sqrt{P_1P_3}$, and $E(X_2, X_3) = \rho_2\sqrt{P_2P_3}$. The converse proof is based on reasoning similar to that in [26] and is provided in "Appendix C".

It is noted that the channel parameters, i.e., P_1, P_2, P_3, h_{31}, h_{13}, h_{23}, must satisfy (58)-(61) for all $-1 \leq \rho_1, \rho_2 \leq 1 : \rho_1^2 + \rho_2^2 \leq 1$, to numerically evaluate the $\mathcal{C}_1^G$ using (62)-(65). Here, we choose $P_1 = P_2 = P_3 = 6, h_{31} = h_{13} = h_{23} = \sqrt{1.5}$ which satisfy strong interference conditions (58)-(61); hence, the regions are derived under strong interference conditions.

Figure 5 shows the capacity region for the Gaussian MA-CIFC of Theorem 8, for $P_1 = P_2 = P_3 = 6$, and $h_{31} = h_{13} = h_{23} = \sqrt{1.5}$, where $\rho_1 = \rho_2$ is fixed in each surface. The $\rho_1 = \rho_2 = 0$ region corresponds to the *no cooperation* case, where the channel inputs are independent. It can be seen that as $\rho_1 = \rho_2$ increases, the bound
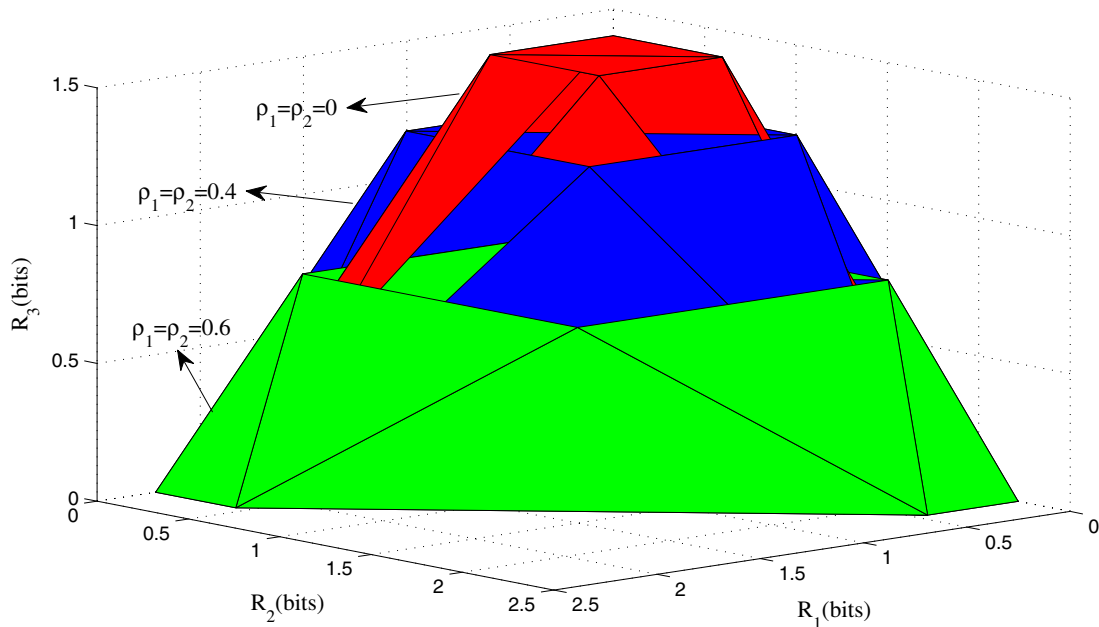
Figure 5 The capacity region for the Gaussian MA-CIFC for fixed $\rho_1 = \rho_2$ under the strong interference conditions.

on R_3 becomes more restrictive while the sum-rate bounds become looser; because Tx3 dedicates parts of its power for cooperation. This means that, as Tx3 allocates more power to relay m_1, m_2 by increasing $\rho_1 = \rho_2$, R_1 and R_2 improve, while R_3 degrades due to the less power allocated to transmit m_3. The capacity region for this channel is the union of all the regions obtained for different values of ρ_1 and ρ_2, satisfying $\rho_1^2 + \rho_2^2 \leq 1$. This union is shown in Figure 6. In order to better perceive the effect of cooperation, we let $R_2 = 0$ in Figure 7. It is seen that by increasing $\rho_1 = \rho_2$, the bound on $R_1 + R_3$ becomes looser and R_1 improves, while R_3 decreases due to the more power dedicated for cooperation.

B. Weak interference regime

Now, we consider the Gaussian MA-CIFC with weak interference at the primary receiver (Rx1), which means $h_{31} \leq 1$. We remark that, since there is no cooperation between receivers, the capacity region for this channel is the same as the one with the same marginal outputs $p(y_1^n | x_1^n, x_2^n, x_3^n)$ and $p(y_3^n | x_1^n, x_2^n, x_3^n)$. Hence, we can state the following useful lemma.

Lemma 2: The capacity region of a Gaussian MA-CIFC, defined by (55) and (56) when $h_{31} \leq 1$, is the same as the capacity region of a Gaussian MA-CIFC with the following channel outputs:

$$\hat{Y}_{1,i} = X_{1,i} + X_{2,i} + h_{31} Y'_{3,i} + Z'_{1,i} \tag{67}$$

$$\widehat{Y}_{3,i} = h_{13} X_{1,i} + h_{23} X_{2,i} + Y'_{3,i} \tag{68}$$

where $Y'_{3,i} = X_{3,i} + Z_{3,i}$ and $Z'_{1,i} \sim \mathcal{N}(0, 1 - h_{31}^2)$. Therefore, the degradedness condition in (29) holds for the Gaussian MA-CIFC when $h_{31} \leq 1$.

*Proof:* The proof follows from [6, Lemma 3.5].

Next, we use the inner bound of Theorem 2 and the outer bound of Theorem 4 to derive the capacity region, which shows that the capacity-achieving scheme in this case consists of DPC at the cognitive transmitter and treating interference as noise at both receivers.

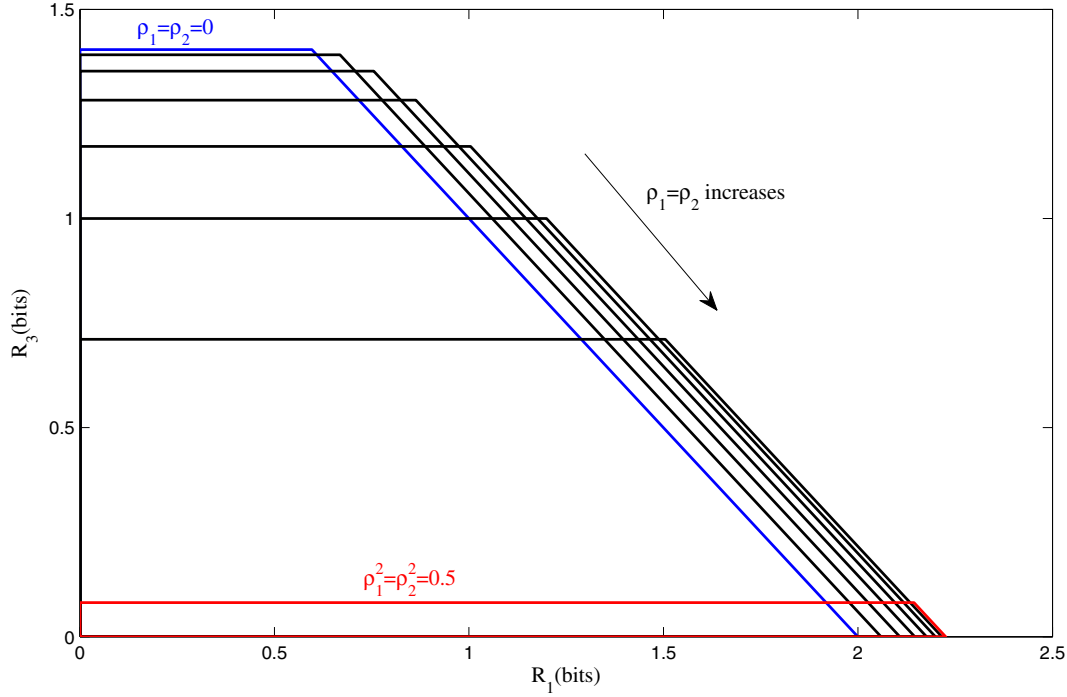*Theorem 9:* For the Gaussian MA-CIFC, defined by (55) and (56), when $h_{31} \leq 1$, the capacity region is given by

$$\mathcal{C}_2^G = \bigcup_{-1 \leq \rho_1, \rho_2 \leq 1 : \rho_1^2 + \rho_2^2 \leq 1} \{(R_1, R_2, R_3) : R_1, R_2, R_3 \geq 0 \tag{69}$$

$$R_3 \leq \theta(P_3(1 - \rho_1^2 - \rho_2^2))$$

$$R_1 \leq \theta\left(\frac{(\sqrt{P_1} + h_{31}\rho_1\sqrt{P_3})^2}{h_{31}^2 P_3(1 - \rho_1^2 - \rho_2^2) + 1}\right) \tag{70}$$
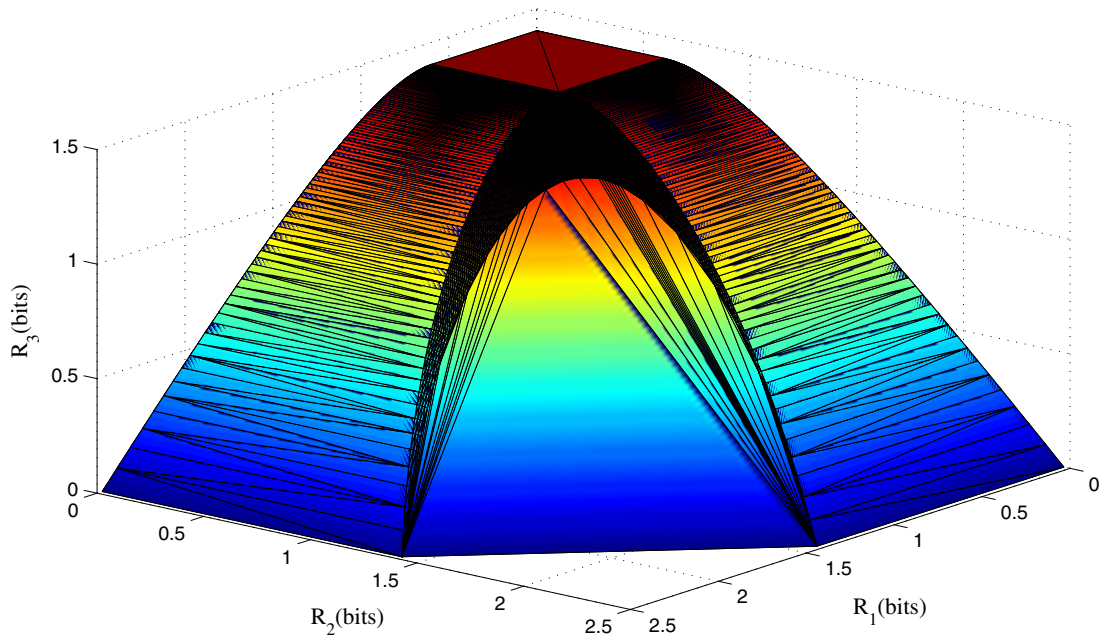
Figure 6 The capacity region for the Gaussian MA-CIFC under the strong interference conditions.

Figure 7 The capacity region for the Gaussian MA-CIFC under the strong interference conditions when $R_2 = 0$.

$$R_2 \leq \theta \left(\frac{\left(\sqrt{P_2} + h_{31}\rho_2\sqrt{P_3}\right)^2}{h_{31}^2 P_3(1 - \rho_1^2 - \rho_2^2) + 1} \right) \tag{71}$$

$$R_1 + R_2 \leq$$
$$\theta \left(\frac{\left(\sqrt{P_1} + h_{31}\rho_1\sqrt{P_3}\right)^2 + \left(\sqrt{P_2} + h_{31}\rho_2\sqrt{P_3}\right)^2}{h_{31}^2 P_3(1 - \rho_1^2 - \rho_2^2) + 1} \right) \right\} \tag{72}$$

where $\theta(\cdot)$ is defined in (66).

Remark 10: By evaluating (2) with jointly Gaussian channel inputs, one can easily achieve (69). However, this results in the Gaussian counterparts of the bounds in (7)-(10). Therefore, some conditions are necessary to make these bounds redundant, similar to the ones in (30)-(33). However, we show that (12) is also evaluated to (69), if we apply DPC with appropriate parameters. Hence, conditions (30)-(33) are unnecessary in the Gaussian case. This means that DPC completely mitigates the effects of interference for the Tx3-Rx3 pair and leaves the link between them interference-free for fixed values of ρ_1, ρ_2. Consequently, $\mathcal{C}_2^G$ is independent of h_{13} and h_{23}.

Remark 11: If we omit Tx2, the model reduces to a CIFC and by setting $P_2 = \rho_2 = R_2 = 0$, $\mathcal{C}_2^G$ coincides with the capacity region of the Gaussian CIFC with weak interference, which was characterized in [6, Lemma 3.6].

Proof: The rate region in Theorem 2 can be extended to the discrete-time Gaussian memoryless case with continuous alphabets by standard arguments [23]. Hence, it is sufficient to evaluate (3)-(6) and (12) with an appropriate choice of input distribution to reach (69)-(72). Let $R_0 = 0$, $\mathcal{M}_0 = \emptyset$, and $T = \emptyset$, since Tx1 and Tx2 have no common information. Also, let U and V be deterministic constants. We choose zero mean jointly Gaussian channel inputs X_1, X_2 and X_3. In fact, $X_j \sim \mathcal{N}(0, P_j)$ for $j \in \{1,2,3\}$, where $E(X_1,X_2) = 0$, $E(X_1,X_3) = \rho_1\sqrt{P_1 P_3}$, and $E(X_2,X_3) = \rho_2\sqrt{P_2 P_3}$. Noting the p.m.f (13), consider the following choice of input distribution for certain $\{-1 \leq \rho_1, \rho_2 \leq 1 : \rho_1^2 + \rho_2^2 \leq 1\}$:

$$X_1 \sim \mathcal{N}(0, P_1), X_2 \sim \mathcal{N}(0, P_2)$$
$$W = X_3 + \alpha_1 X_1 + \alpha_2 X_2, X_3' \sim \mathcal{N}(0, (1 - \rho_1^2 - \rho_2^2)P_3)$$
$$X_3 = X_3' + \rho_1\sqrt{\frac{P_3}{P_1}}X_1 + \rho_2\sqrt{\frac{P_3}{P_2}}X_2 \tag{73}$$

Therefore, (3)-(6) are easily evaluated to (70)-(72). In "Appendix D", we derive (69) by evaluating (12) with appropriate parameters. The converse proof follows by applying the bounds in the proof of Theorem 4 to the Gaussian case and utilizing Entropy Power Inequality (EPI) [23,27]. A detailed converse proof is provided in "Appendix D".

Remark 12: According to Theorems 8 and 9, jointly Gaussian channel inputs X_1, X_2 and X_3 are optimal for the Gaussian MA-CIFC under the strong and weak interference conditions, determined in the above theorems.

Figure 8 shows the capacity region for the Gaussian MA-CIFC of Theorem 9, for $P_1 = P_2 = P_3 = 6$, and $h_{31} = \sqrt{0.55}$, where $\rho_1 = \rho_2$ is fixed in each surface. It is noted that the capacity region is independent of h_{13} and h_{23}. The $\rho_1 = \rho_2 = 0$ region corresponds to the *no cooperation* case, where channel inputs are independent. We see that when Tx3 dedicates parts of its power for cooperation, i.e., $\rho_1 = \rho_2 = 0.5$, the rates of the primary users (R_1, R_2) increase, while R_3 decreases. The capacity region for this channel is the union of all the regions obtained for different values of ρ_1 and ρ_2 satisfying, $\rho_1^2 + \rho_2^2 \leq 1$, which is shown in Figure 9. Similar to Figure 7, we investigate the capacity region for $R_2 = 0$ in Figure 10 in the weak interference regime. It is seen that, when Tx3 dedicates more power for cooperation by increasing $\rho_1 = \rho_2$, R_1 improves and R_3 decreases.

7. Conclusion
We investigated a cognitive communication network where a MAC with common information and a point-to-point channel share a same medium and interfere with each other. For this purpose, we introduced Multiple Access-Cognitive Interference Channel (MA-CIFC) by merging a two-user MAC as a primary network and a cognitive transmitter-receiver pair in which the cognitive transmitter knows the message being sent by all of the transmitters in a non-causal manner. We analyzed the capacity region of MA-CIFC by deriving the inner and outer bounds on the capacity region of this channel. These bounds were proved to be tight in some special cases. Therefore, we determined the optimal strategy in these cases. Specifically, in the discrete memoryless case, we established the capacity regions for a class of *degraded* MA-CIFC and also under two sets of strong interference conditions. We also derive strong interference conditions for a network with k primary users. Further, we characterized the capacity region of the Gaussian MA-CIFC in the *weak* and *strong* interference regimes. We showed that DPC at the cognitive transmitter and treating interference as noise at the receivers, i.e., an oblivious primary receiver, are optimal in the weak interference. However, the receivers have to decode all messages when the interference is strong enough, which requires an aware primary receiver.

Appendix A Proofs of Theorems 1, 2 and 3
Outline of the proof for Theorem 1: We propose the following random coding scheme, which contains superposition coding and the technique of [6] in defining
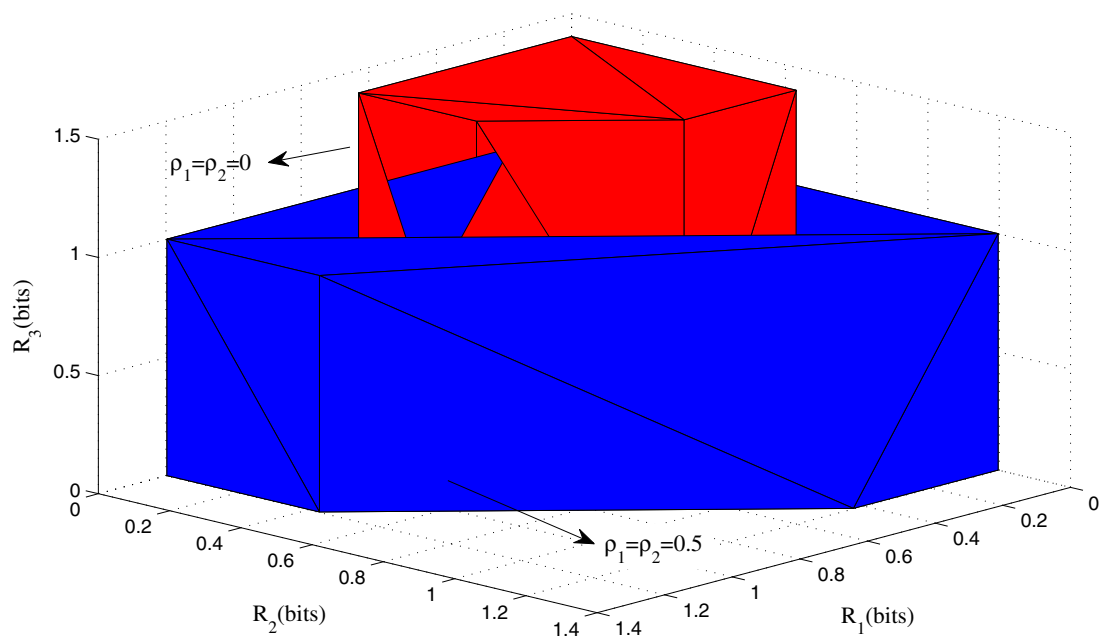
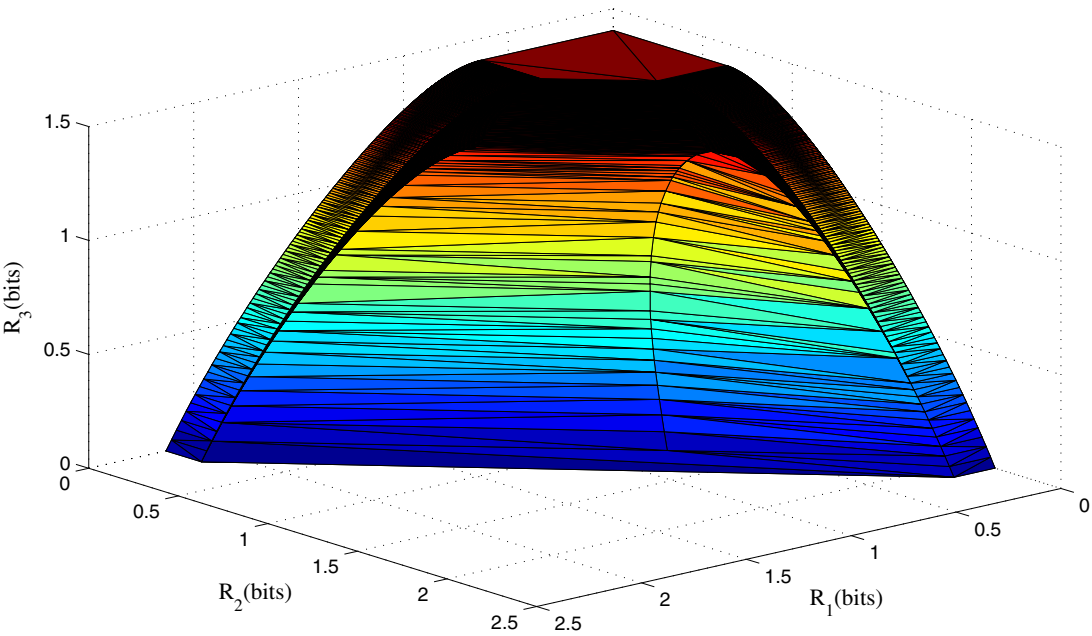**Figure 8 The capacity region for the weak Gaussian MA-CIFC for fixed $\rho_1 = \rho_2$.**

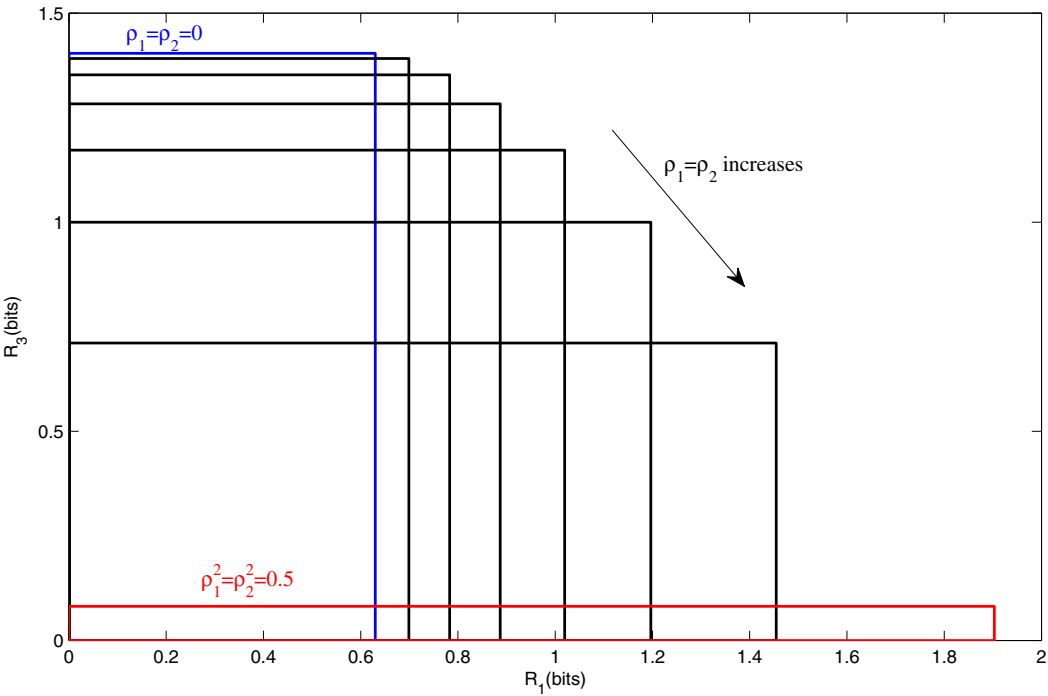**Figure 9 The capacity region for the weak Gaussian MA-CIFC**.

Figure 10 The capacity region for the weak Gaussian MA-CIFC when *R₂ = 0*.

auxiliary RVs, i.e., U and V. The cognitive receiver (Rx3) decodes the interfering signals caused by the primary messages (m_1, m_2), while the primary receiver (Rx1) does not decode the interference from the cognitive user's message (m_3) and treats it as noise.

Codebook Generation: Fix a joint p.m.f as (11). Generate 2^{nR_0} i.i.d t^n sequences, according to the probability $\prod_{i=1}^{n} p(t_i)$. Index them as $t^n(m_0)$ where $m_0 \in [1, 2^{nR_0}]$. For each $t^n(m_0)$, generate 2^{nR_1} i.i.d (u^n, x_1^n) sequences, each with probability $\prod_{i=1}^{n} p(u_i, x_{1,i}|t_i)$. Index them as $(u^n(m_0, m_1), x_1^n(m_0, m_1))$ where $m_1 \in [1, 2^{nR_1}]$. Similarly, for each $t^n(m_0)$, generate 2^{nR_2} i.i.d (v^n, x_2^n) sequences, each with probability $\prod_{i=1}^{n} p(v_i, x_{2,i}|t_i)$. Index them as $(v^n(m_0, m_2), x_2^n(m_0, m_2))$ where $m_2 \in [1, 2^{nR_2}]$. For each $(t^n(m_0), u^n(m_0, m_1), x_1^n(m_0, m_1), v^n(m_0, m_2), x_2^n(m_0, m_2))$, generate 2^{nR_3} i.i.d x_3^n sequences, according to $\prod_{i=1}^{n} p(x_{3,i}|t_i, u_i, x_{1,i}, v_i, x_{2,i})$. Index them as $x_3^n(m_0, m_1, m_2, m_3)$ where $m_3 \in [1, 2^{nR_3}]$.

Encoding: In order to transmit the messages (m_0, m_1, m_2, m_3), Txj sends $x_j^n(m_0, m_j)$ for $j \in \{1,2\}$ and Tx3 sends $x_3^n(m_0, m_1, m_2, m_3)$.

Decoding:

Rx1: After receiving y_1^n, Rx1 looks for a unique triple $(\hat{m}_0, \hat{m}_1, \hat{m}_2)$ and some $\hat{m}_2$ such that

$$(y_1^n, t^n(\widehat{m}_0), u_n(\widehat{m}_0, \widehat{m}_1), x_1^n(\widehat{m}_0, \widehat{m}_1), v^n(\widehat{m}_0, \widehat{m}_2),$$
$$x_2^n(\widehat{m}_0, \widehat{m}_2)) \in A_\epsilon^n(Y_1, T, U, X_1, V, X_2).$$

For large enough n with arbitrarily high probability, $(\hat{m}_0, \hat{m}_1, \hat{m}_2) = (m_0, m_1, m_2)$ if (3)-(6) hold.

Rx3: After receiving y_3^n, Rx3 finds a unique index $\hat{\hat{m}}_3$ and some triple $(\hat{\hat{m}}_0, \hat{\hat{m}}_1, \hat{\hat{m}}_2)$ such that

$$(y_3^n, x_3^n(\widehat{\widehat{m}}_0, \widehat{\widehat{m}}_1, \widehat{\widehat{m}}_2, \widehat{\widehat{m}}_3), t^n(\widehat{\widehat{m}}_0), u^n(\widehat{\widehat{m}}_0, \widehat{\widehat{m}}_1), x_1^n(\widehat{\widehat{m}}_0, \widehat{\widehat{m}}_1),$$
$$v^n(\widehat{\widehat{m}}_0, \widehat{\widehat{m}}_2), x_2^n(\widehat{\widehat{m}}_0, \widehat{\widehat{m}}_2)) \in A_\epsilon^n(Y_3, X_3, T, U, X_1, V, X_2).$$

With arbitrary high probability, $\hat{\hat{m}}_3 = m_3$ if n is large enough and (2), (7)-(10) hold. This completes the proof.

Outline of the proof for Theorem 2: Our proposed random coding scheme, in the encoding part, contains the methods of Theorem 1 and GP binning at the cognitive transmitter (Tx3) which is used at Tx3 to cancel the interference caused by m_0, m_1, m_2 at Rx3. In the decoding part, both receivers decode only their intended messages, treating the interference as noise. Therefore, unlike the decoding part of Theorem 1, Rx3 decodes

only its own message ($m3$), while treating the other signals as noise.

Codebook Generation: Fix a joint p.m.f as (13). Generate $t^n(m_0), u^n(m_0, m_1), x_1^n(m_0, m_1), v^n(m_0, m_2), x_2^n(m_0, m_2)$ codewords based on the same lines as in the *codebook generation* part of Theorem 1. Then, generate $2^{n(R_3+L)}$ i.i.d w^n sequences. Index them as $w^n(m_3, l)$ where $m_3 \in [1, 2^{nR_3}]$ and $l \in [1, 2^{nL}]$.

Encoding: In order to transmit the messages (m_0, m_1, m_2, m_3), Txj sends $x_j^n(m_0, m_j)$ for $j \in \{1,2\}$. Tx3 (the cognitive transmitter), in addition to m_3, knows m_0, m_1 and m_2. Hence, knowing codewords $t^n, u^n, x_1^n, v^n, x_2^n$, to send m_3, it seeks an index l such that

$$(w^n(m_3, l), t^n(m_0), u^n(m_0, m_1), x_1^n(m_0, m_1), v^n(m_0, m_2),$$
$$x_2^n(m_0, m_2)) \in A_\epsilon^n(W, T, U, X_1, V, X_2).$$

If there is more than one such index, Tx3 picks the smallest. If there is no such codeword, it declares an error. Using covering lemma [27], it can be shown that there exists such an index l with high enough probability if n is sufficiently large and

$$L \geq I(W; T, U, X_1, V, X_2). \tag{74}$$

Then, Tx3 sends x_3^n generated according to $\prod_{i=1}^{n} p(x_{3,i}|t_i, u_i, x_{1,i}, v_i, x_{2,i})$.

Decoding: The decoding procedure at Rx1 is similar to Theorem 1 and the error in this receiver can be bounded if (3)-(6) hold.

Rx3: After receiving y_3^n, Rx3 finds a unique index $\hat{\hat{m}}_3$ for some index $\hat{\hat{l}}$ such that

$$\left(y_3^n, w^n\left(\widehat{\widehat{m}}_3, \widehat{\widehat{l}}\right)\right) \in A_\epsilon^n(Y_3, W).$$

For large enough n, the probability of error can be made sufficiently small if

$$R_3 + L \leq I(W; Y_3). \tag{75}$$

Combining (74) and (75) results in (12). This completes the proof.

Outline of the proof for Theorem 3: We propose the following random coding scheme, which contains superposition coding in the encoding part and simultaneous joint decoding in the decoding part. All messages are common to both receivers, i.e., both receivers decode (m_0, m_1, m_2, m_3).

Codebook Generation: Fix a joint p.m.f as (18). Generate 2^{nR_0} i.i.d t^n sequences, according to the probability $\prod_{i=1}^{n} p(t_i)$. Index them as $t^n(m_0)$ where $m_0 \in [1, 2^{nR_0}]$. For j

$\in \{1,2\}$ and each $t^n(m_0)$, generate 2^{nR_j} i.i.d x_j^n sequences, each with probability $\prod_{i=1}^{n} p(x_{j,i}|t_i)$. Index them as $x_j^n(m_0, m_j)$ where $m_j \in [1, 2^{nR_j}]$. For each $(t^n(m_0), x_1^n(m_0, m_1), x_2^n(m_0, m_2))$, generate 2^{nR_3} i.i.d x_3^n sequences, each with probability $\prod_{i=1}^{n} p(x_{3,i}|t_i, x_{1,i}, x_{2,i})$. Index them as $x_3^n(m_0, m_1, m_2, m_3)$ where $m_3 \in [1, 2^{nR_3}]$.

Encoding: In order to transmit message (m_0, m_1, m_2, m_3), Txj sends $x_j^n(m_0, m_j)$ for $j \in \{1,2\}$ and Tx3 sends $x_3^n(m_0, m_1, m_2, m_3)$.

Decoding:

Rx1: After receiving γ_1^n, Rx1 looks for a unique triple $(\hat{m}_0, \hat{m}_1, \hat{m}_2)$ and some $\hat{m}_3$ such that

$$(\gamma_1^n, t^n(\hat{m}_0), x_1^n(\hat{m}_0, \hat{m}_1), x_2^n(\hat{m}_0, \hat{m}_2), x_3^n(\hat{m}_0, \hat{m}_1, \hat{m}_2, \hat{m}_3))$$
$$\in A_\epsilon^n(Y_1, T, X_1, X_2, X_3).$$

For large enough n, with arbitrarily high probability $(\hat{m}_0, \hat{m}_1, \hat{m}_2) = (m_0, m_1, m_2)$ if

$$R_1 + R_3 \leq I(X_1, X_3; Y_1 | X_2, T) \tag{76}$$

$$R_2 + R_3 \leq I(X_2, X_3; Y_1 | X_1, T) \tag{77}$$

$$R_0 + R_1 + R_2 + R_3 \leq I(X_1, X_2, X_3; Y_1). \tag{78}$$

Rx3: Similarly, after receiving γ_3^n, Rx3 finds a unique index $\hat{\hat{m}}_3$ and some triple $(\hat{\hat{m}}_0, \hat{\hat{m}}_1, \hat{\hat{m}}_2)$ such that

$$(\gamma_3^n, t^n(\hat{\hat{m}}_0), x_1^n(\hat{\hat{m}}_0, \hat{\hat{m}}_1), x_2^n(\hat{\hat{m}}_0, \hat{\hat{m}}_2), x_3^n(\hat{\hat{m}}_0, \hat{\hat{m}}_1, \hat{\hat{m}}_2, \hat{\hat{m}}_3))$$
$$\in A_\varepsilon^n(Y_3, T, X_1, X_2, X_3).$$

With the arbitrary high probability $\hat{\hat{m}}_3 = m_3$, if n is large enough and

$$R_3 \leq I(X_3; Y_3 | X_1, X_2, T) \tag{79}$$

$$R_1 + R_3 \leq I(X_1, X_3; Y_3 | X_2, T) \tag{80}$$

$$R_2 + R_3 \leq I(X_2, X_3; Y_3 | X_1, T) \tag{81}$$

$$R_0 + R_1 + R_2 + R_3 \leq I(X_1, X_2; Y_3 | X_1). \tag{82}$$

This completes the proof.

Appendix B Proof of Theorem 7

Since achievability directly follows from Theorem 3, we proceed to the converse part. We assume a code with the properties indicated in the converse proof of Theorem 6 and define T^n as (42). Four bounds in $\mathcal{C}_2^{str}$ are the same as the bounds in $\mathcal{C}_1^{str}$, which are shown in the converse proof of Theorem 6. Therefore, it is necessary to

prove the three bounds. Moreover, similar to Lemma 1, we have:

Lemma 3: If (48)-(50) hold for all distributions that factor as (18), then

$$I(X_1^n; Y_1^n | X_2^n, T^n) \leq I(X_1^n; Y_3^n | X_2^n, T^n) \tag{83}$$

$$I(X_2^n; Y_1^n | X_1^n, T^n) \leq I(X_2^n; Y_3^n | X_1^n, T^n) \tag{84}$$

$$I(X_1^n; X_2^n; Y_1^n) \leq I(X_1^n; Y_2^n | X_3^n). \tag{85}$$

We next prove the remaining three bounds of Theorem 7. Applying Fano's inequality, similar to (45), we have:

$$n(R_1 + R_3) - n(\delta_{1n} + \delta_{3n})$$
$$\leq I(M_1, X_1^n; Y_1^n | M_0, M_2, X_2^n) + I(X_3^n; Y_3^n | M_0, M_1, M_2, X_1^n, X_2^n)$$
$$\overset{(a)}{=} I(X_1^n; Y_1^n | T^n, M_2, X_2^n) + I(X_3^n; Y_3^n | T^n, M_1, M_2, X_1^n, X_2^n)$$
$$\overset{(b)}{\leq} I(X_1^n; Y_3^n | T^n, M_2, X_2^n) + I(X_3^n; Y_3^n | T^n, M_1, M_2, X_1^n, X_2^n) \tag{86}$$
$$\overset{(c)}{\leq} I(M_1, X_1^n, X_3^n; Y_3^n | T^n, M_2, X_2^n)$$
$$\overset{(d)}{\leq} \sum_{i=1}^{n} I(X_{1,i}, X_{3,i}; Y_{3,i} | X_{2,i}, T_i)$$

where (a) follows from (42) and the deterministic relation between X_1^n and M_1, (b) is obtained from (83), (c) is due to the fact that mutual information is non-negative, and (d) follows from the facts that conditioning does not increase entropy and the channel is memoryless.

Using a similar approach and condition (84), it can be shown that

$$n(R_2 + R_3) - n(\delta_{2n} + \delta_{3n}) \leq \sum_{i=1}^{n} I(X_{2,i}, X_{3,i} Y_{3,i} | X_{1,i}, T_i). \tag{87}$$

Finally, using (47) we obtain the last sum-rate bound as:

$$n(R_0 + R_1 + R_2 + R_3) - n(\delta_{0n} + \delta_{1n} + \delta_{2n} + \delta_{3n})$$
$$\leq I(M_0, M_1, M_2, X_1^n, X_2^n; Y_1^n)$$
$$\qquad + I(X_3^n; Y_3^n | M_0, M_1, M_2, X_1^n, X_2^n)$$
$$\overset{(a)}{\leq} I(X_1^n, X_2^n; Y_1^n) + I(X_3^n; Y_3^n | X_1^n, X_2^n) \tag{88}$$
$$\overset{(b)}{\leq} I(X_1^n, X_2^n; Y_3^n) + I(X_3^n; Y_3^n | X_1^n, X_2^n)$$
$$= I(X_1^n, X_2^n, X_3^n; Y_3^n) \leq \sum_{i=1}^{n} I(X_{1,i}, X_{2,i} X_{3,i}; Y_{3,i})$$

where (a) is obtained from the deterministic relation between X_j^n and (M_0, M_j) for $j \in \{1,2\}$, and (b) follows from (85). Using a standard time-sharing argument for (44)-(47) and (86)-(88) completes the proof.

Appendix C Proof of the converse part for Theorem 8

For any rate triple $(R_1, R_2 R_3) \in \mathcal{C}$, Rx1 is able to decode m_1 and m_2 reliably. Then, Rx1 is able to construct

$$\tilde{Y}_3 = \frac{Y_1 - X_1 - X_2}{h_{31}} + h_{13}X_1 + h_{23}X_2$$

$$= h_{13}X_1 + h_{23}X_2 + X_3 + \frac{1}{h_{31}}Z_1.$$

If condition (58) holds, then $\tilde{Y}_3$ is a less noisy version of Y_3. Since Rx3 has to decode m_3, Rx1 can decode m_3 via $\tilde{Y}_3$ with no rate penalty. Therefore, (R_1, R_2, R_3) is contained in the capacity region of a three-user MAC with common information [19] at Rx1, where R_1 and R_2 are the common rates between Tx1-Tx3 and Tx2-Tx3, respectively; R_3 is the private rate for Tx3; and the private rates for Tx1 and Tx2 are zero. From the maximum entropy theorem [23], this region is largest for Gaussian inputs and is evaluated to (63)-(65). The bound in (62) follows by applying the standard methods as in (44).

Appendix D Detailed proof for Theorem 9

First, we derive (69) by evaluating (12) with appropriate parameters. Consider the mapping in (73). We choose

$$\alpha_1 = \beta\left(h_{13} + \rho_1\sqrt{\frac{P_3}{P_1}}\right) - \rho_1\sqrt{\frac{P_3}{P_1}} \tag{89}$$

$$\alpha_2 = \beta\left(h_{23} + \rho_2\sqrt{\frac{P_3}{P_2}}\right) - \rho_2\sqrt{\frac{P_3}{P_2}} \tag{90}$$

$$\beta = \frac{P_3(1 - \rho_1^2 - \rho_2^2)}{P_3(1 - \rho_1^2 - \rho_2^2) + 1} \tag{91}$$

It is noted that (91) makes $(1 - \beta)X'_3 - \beta Z_3$ and $X'_3 + Z_3$ uncorrelated and hence independent since they are jointly Gaussian. Using (12) and (73), we obtain

$$I(W; Y_3) - I(W; T, U, X_1, V, X_2) = h(W|X_1, X_2) - h(W|Y_3)$$
$$= h(X'_3) - h(W|Y_3). \tag{92}$$

Now, we evaluate $h(W|Y_3)$:

$$h(W|Y_3) = h(X'_3 + (\alpha_1 + \rho_1\sqrt{\frac{P_3}{P_1}})X_1 + (\alpha_2 + \rho_2\sqrt{\frac{P_3}{P_2}})X_2|Y_3)$$

$$= h(X'_3 + (\alpha_1 + \rho_1\sqrt{\frac{P_3}{P_1}})X_1 + (\alpha_2 + \rho_2\sqrt{\frac{P_3}{P_2}})X_2 - \beta Y_3|Y_3)$$

$$\overset{(a)}{=} h((1 - \beta)X'_3 - \beta Z_3|(h_{13} + \rho_1\sqrt{\frac{p_3}{P_1}})X_1 \tag{93}$$

$$+ (h_{23} + \rho_2\sqrt{\frac{P_3}{P_2}})X_2 + X'_3 + Z_3$$

$$\overset{(b)}{=} h((1 - \beta)X'_3 - \beta Z_3|X'_3 + Z_3) = h(X'_3|X'_3 + Z_3)$$

where (a) follows from (89), (90), and $Y_{13} + \rho_1\sqrt{\frac{P_3}{P_1}})X_1 + (h_{23} + \rho_2\sqrt{\frac{P_3}{P_2}})X_2 + X'_3 + Z_3$ is obtained because, by applying (91), $(1 - \beta)X'_3 - \beta Z_3, X'_3 + Z_3, X_1, X_1$ and X_2 are jointly independent. By applying (93) to (92), we can evaluate (12) as:

$$h(X'_3) - h(W|Y_3) = h(X'_3) - h(X'_3|X'_3 + Z_3)$$
$$= I(X'_3; X'_3 + Z_3) = \theta(P_3(1 - \rho_1^2 - \rho_2^2))$$

This completes the achievability proof.

Converse: Utilizing the power constraint (57), we derive the bounds of $\mathcal{R}_o^1$ for the Gaussian MA-CIFC when $h_{31} < 1$. It can be easily shown that we can set $T = \emptyset$, since $R_0 = 0$ and $\mathcal{M}_0 = \emptyset$. Applying the degradedness condition (29) to (19), which holds due to Lemma 2, we obtain:

$$R_3 \leq I(X_3; Y_3, Y_1|U, X_1, V, X_2) = I(X_3; Y_3|U, X_1, V, X_2)$$
$$= h(Y_3|U, X_1, V, X_2) - \frac{1}{2}\log 2\pi e. \tag{94}$$

Using the fact that conditioning does not increase the entropy and (57) with $j = 3$ yields:

$$h(Y_3|U, X_1, V, X_2) \geq h(Y_3|U, X_1, V, X_2, X_3) = \frac{1}{2}\log 2\pi e$$

$$h(Y_3|U, X_1, V, X_2) = h(X_3 + Z_3|U, X_1, V, X_2, X_3)$$

$$\leq h(X_3 + Z_3) \leq \frac{1}{2}\log 2\pi e(P_3 + 1).$$

Hence, there exists $0 \leq \lambda \leq 1$ such that

$$h(Y_3|U, X_1, V, X_2) = \frac{1}{2}\log 2\pi e(\lambda P_3 + 1). \tag{95}$$

Combining (94) and (95) results in

$$R_3 \leq I(X_3; Y_3|U, X_1, V, X_2) = \frac{1}{2}\log(\lambda P_3 + 1) \tag{96}$$

for some $0 \leq \lambda \leq 1$. Now, considering that the Gaussian distribution maximizes the entropy of a RV for a given value of the second moment, we have:

$$h(Y_3|U, X_1, V, X_2) \leq h(X_3 + Z_3|X_1, X_2)$$
$$\leq \frac{1}{2}\log 2\pi e(P_3(1 - \rho_1^2 - \rho_2^2) + 1) \tag{97}$$

where $-1 \leq \rho_1, \rho_2 \leq 1 : \rho_1^2 + \rho_2^2 \leq 1, \rho_1 = \frac{E(X_1, X_3)}{\sqrt{P_1 P_3}}$, and $\rho_2 = \frac{E(X_2, X_3)}{\sqrt{P_2 P_3}}$. Considering (95) and (97), we obtain:

$$\lambda \leq 1 - \rho_1^2 - \rho_2^2$$

Moreover, since (69) is achievable and R_3 does not appear in (70)-(72), it can be shown that

$$\lambda = 1 - \rho_1^2 - \rho_2^2. \tag{98}$$

Consequently, (69) follows from (96) and (98). Now, consider the bound in (3):

$$R_1 \le h(Y_1|V, X_2) - h(Y_1|U, X_1, V, X_2). \tag{99}$$

We can compute the first term as

$$
h(Y_1|V, X_2) \\
\le \frac{1}{2}\log\left(P_1 + h_{31}^2 P_3(1-\rho_2^2) + 2h_{31\rho_1}\sqrt{P_1 P_3} + 1\right) \tag{100}
$$

which incorporates the fact that Gaussian distribution maximizes the entropy of a RV for a given value of the second moment.

To compute the second term in (99), the Entropy Power Inequality (EPI) [23,27] is used to obtain:

$$
\begin{aligned}
2^{2h(Y_1|U,X_1,V,X_2)} &\overset{(a)}{=} 2^{2h(h_{31}Y'_3+Z'_1|U,X_1,V,X_2)} \\
&\overset{(b)}{\ge} 2^{2h(h_{31}Y'_3|U,X_1,V,X_2)} + 2^{2h(Z'_1|U,X_1,V,X_2)} \\
&= h_{31}^2 2^{2h(Y_3|U,X_1,V,X_2)} + 2\pi e(1-h_{31}^2) \\
&\overset{(c)}{=} 2\pi e(1 + h_{31}^2 P_3(1-\rho_1^2-\rho_2^2))
\end{aligned} \tag{101}
$$

where (a) follows from (67), (b) is obtained from EPI, and (c) follows from (95) and (98). Therefore, (3) is evaluated to (70) by combining (99)-(101).

Utilizing (101), one can easily evaluate (4) and (5) to (71) and (72), respectively. This completes the converse proof for Theorem 9.

Competing interests
The authors declare that they have no competing interests.

References
1. A Goldsmith, SA Jafar, I Maric, S Srinivasa, Breaking spectrum gridlock with cognitive radios: an information theoretic perspective. Proc ieee, invited. **97**(5), 894–914 (2009)
2. N Devroye, P Mitran, V Tarokh, Achievable rates in cognitive radio channels. ieee Trans Inf Theory. **52**, 1813–1827 (2006)
3. S Gel'fand, M Pinsker, Coding for channels with random parameters. Probl Control Info Theory. **9**(1), 19–31 (1980)
4. TS Han, K Kobayashi, A new achievable rate region for the interference channel. ieee Trans Inf Theory. **27**, 49–60 (1981). doi:10.1109/TIT.1981.1056307
5. A Jovicic, P Viswanath, Cognitive radio: an information-theoretic perspective. ieee Trans Inf Theory. **55**, 3945–3958 (2009)
6. W Wu, S Vishwanath, A Arapostathis, Capacity of a class of cognitive radio channels: interference channels with degraded message sets. ieee Trans Inf Theory. **53**(11), 4391–4399 (2007)
7. N Devroye, M Vu, V Tarokh, Achievable rates and scaling Laws in cognitive radio channels. EURASIP J Wirel Commun Netw Special issue on cognitive radio and dynamic spectrum sharing systems. **2008**(896246) (2008)
8. I Maric, RD Yates, G Kramer, Capacity of interference channels with partial transmitter cooperation. IEEE Trans Inf Theory. **53**, 3536–3548 (2007)
9. I Maric, A Goldsmith, G Kramer, S Shamai (Shitz), On the capacity of interference channels with one cooperating transmitter. Eur Trans Telecommun. **19**, 405–420 (2008). doi:10.1002/ett.1298
10. S Rini, D Tuninetti, N Devroye, State of the cognitive interference channel: a new unified inner bound, and capacity to within 1.87 bits, in *Proceedings of the 2010 International Zurich Seminar on Communications* http://arxiv.org/abs/0910.3028v1
11. S Rini, D Tuninetti, B Devroye, New inner and outer bounds for the discrete memoryless cognitive interference channel and some capacity results. http://arxiv.org/abs/1003.4328v1 (2010)
12. M Mirmohseni, B Akhbari, MR Aref, On the capacity of interference channel with causal and non-causal generalized feedback at the cognitive transmitter. IEEE Trans Inf Theoryhttp://ee.sharif.ir/~mirmohseni/IT-prep.pdf. submitted to, April 2010, Revised, Oct. 2011.
13. M Mirmohseni, B Akhbari, MR Aref, Capacity regions for some classes of causal cognitive interference channels with delay, in *Proceedings of the IEEE Information Theory Workshop (ITW)*, (Dublin, 2010)
14. J Jiang, Y Xin, On the achievable rate regions for interference channels with degraded message sets. IEEE Trans Inf Theory. **54**, 4707–4712 (2008)
15. KG Nagananda, CR Murthy, Three-user cognitive channels with cumulative message sharing: an achievable rate region, in *Proceedings of the IEEE Information Theory Workshop (ITW)*, (Greece, 2009), pp. 291–295
16. KG Nagananda, CR Murthy, Information theoretic results for three-user cognitive channels, in *Proceedings of the IEEE Globecom* (2009)
17. R Ahlswede, Multiway communication channels. In: Proceedings of the IEEE International Symposium Information Theory (ISIT), (Tsahkadsor, Armenian S.S.R, 1971), pp. 23–52
18. HHJ Liao, Multiple access channels, (University of Hawaii, Honolulu, 1972) Ph.D. Thesis
19. D Slepian, JK Wolf, A coding theorem for multiple access channels with correlated sources. Bell Syst Tech J. **52**, 1037–1076 (1973)
20. A Chaaban, A Sezgin, On the capacity of the 2-user Gaussian MAC interfering with a P2P link. http://arxiv.org/abs/1010.6255v1 (2010)
21. F Zhu, X Shang, B Chen, HV Poor, On the capacity of multiple-access-Z-interference channels. http://arxiv.org/abs/1011.1225v1 (2010)
22. MHM Costa, Writing on dirty paper. IEEE Trans Inf Theory. **29**, 439–441 (1983). doi:10.1109/TIT.1983.1056659
23. TM Cover, JA Thomas, Elements of Information Theory, 2nd edn. (Wiley Series in Telecommunications, 2006)
24. J Koörner, K Marton, Comparison of two noisy channels, in *Topics in information theory*, ed. by I Csiszar, P Elias (North-Holland, Amsterdam, 1977). Colloquia Mathematica Societatis Janos Bolyai
25. MHM Costa, A El Gamal, The capacity region of the discrete mem-oryless interference channel with strong interference. IEEE Trans Inf Theory. **33**(5), 710–711 (1987). doi:10.1109/TIT.1987.1057340
26. H Sato, The capacity of the Gaussian interference channel under strong interference. IEEE Trans Inf Theory. **27**(6), 786–788 (1981). doi:10.1109/TIT.1981.1056416
27. A El Gamal, YH Kim, Lecture notes on network information theory. http://arxiv.org/abs/1001.3404 (2010). [Online]. Available