Fifty-second Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
October 1 - 3, 2014

The Capacity of More Capable Cognitive Interference Channels

Mojtaba Vaezi
McGill University, Montreal, Quebec H3A 0E9, Canada
Email: mojtaba.vaezi@mail.mcgill.com

***Abstract*—By introducing a new outer bound, we establish the capacity region for a class of discrete memoryless *cognitive interference channel* (DM-CIC) called *cognitive-more-capable* channel, and we show that superposition coding is the optimal encoding technique. The new capacity region is the largest capacity region for the two-user DM-CIC to date, as all existing capacity results are explicitly shown to be its subsets. In the effort to prove the latter, we also clarify the relation among the existing capacity results of the DM-CIC. Besides, as a byproduct of the introduced outer bound, we find a more tractable presentation of the outer bound introduced in [1, Theorem 3.2].**

I. Introduction

Wireless communication systems are limited in performance and capacity by interference. The capacity region of the simplest *interference channel* [2], i.e., the two-user interference channel, is still an open problem despite having been studied for several decades. It is known only for a few special classes of interference channels, e.g., the *strong* and *very strong* interference [3].

With ever-increasing demand for radio spectrum, improving the spectral utilization in wireless communication systems is unavoidable. *Cognitive radio* is recognized as a key enabling technology for this purpose [4]. Owing to the nodes which can sense the environment and adapt their strategy based on the network setup, cognitive radio technology is aimed at increasing the spectral efficiency in wireless communication systems. With this development in technology, networks with cognitive users are gaining prominence. Such a communication channel can be modeled by interference channel with cognition, which is simply known as the *cognitive channel* [5].

Most of the recent work on the cognitive interference channel has focused on the two user channel with one cognitive transmitter [1], [5]–[10]. In this channel setting, one transmitter, known as the cognitive transmitter, has *non-causal* access to the message transmitted by the other transmitter (the primary transmitter). This is used to model an "ideal" cognitive radio. We study the two-user *discrete memoryless* cognitive interference channel, too. Fundamental limits of this channel have been explored for several years now. However, the capacity of this channel remains unknown except for some special classes, e.g., in the "weak interference" [1] and "strong interference" [7] regimes.

Inspired by the concept of the less noisy broadcast channel (BC) [11, Chapter 5], the author introduced the notion of *less noisy* DM-CIC in [10]. Because of the inherent asymmetry of the cognitive channel, two different less noisy channels are distinguishable; these are dubbed the *primary-less-noisy* and *cognitive-less-noisy* DM-CIC. In the former, the primary receiver is less noisy than the secondary receiver, whereas it is the opposite in the latter. In this paper, we extend the work on *less noisy* DM-CIC [10] to the *more capable* DM-CIC. The notion of more capable DM-CIC first appeared in [12], in which the primary receiver is more capable than the secondary one. It is easy to see that, similar to the less noisy DM-CIC, two different more capable cognitive channels are conceivable: the *primary-more-capable* and *cognitive-more-capable* DM-CIC. The former was studied in [12]; the latter is the subject of study in this work.

The main contribution of this paper is to establish the capacity region for the *cognitive-more-capable DM-CIC*. To this end, in Theorem 1, we first establish a new outer bound on the capacity region of this channel. Then, we present a simpler representation of this outer bound in Theorem 2. As a corollary of this new presentation of the outer bound, we further obtain a more tractable representation of the outer bound introduced in [1, Theorem 3.2], for the DM-CIC. We then show that the outer bound in Theorem 2 is the same as an inner bound which is based on superposition coding. Therefore, we characterize the capacity region for this class of the DM-CIC and prove that *superposition coding* is the capacity-achieving technique.

In the second part of this work, we prove that the introduced capacity result is the largest capacity region for the DM-CIC to date. To this end, we explicitly show that the existing capacity results are subsets of this new capacity region. Looking from a different perspective, the capacity of the cognitive-more-capable DM-CIC reduces to the capacity regions in the "weak interference" [1], "strong interference" [7], and "less noisy" [10] regimes once the corresponding channel conditions are satisfied. Finally, the relation among different capacity results of the DM-CIC is clarified in light of this work and [13]. Further, the analysis we provide in this paper sheds more light on the existing capacity results of the DM-CIC; it makes clear how superposition coding is the capacity achieving technique in the previous results, too.

The rest of the paper is organized as follows. The system model and definitions are presented in Section II. Section III provides the main result of this paper, which includes the capacity region for the cognitive-more-capable DM-CIC. In Section IV, we show that the new capacity result includes all existing capacity results as subsets. This is followed by conclusions in Section V.

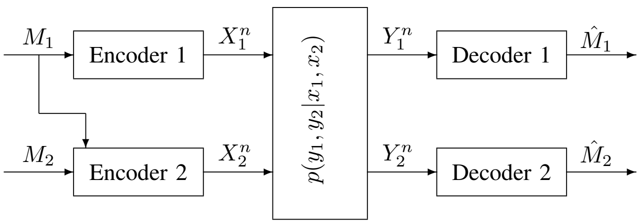

Fig. 1. The discrete memoryless cognitive interference channel (DM-CIC) with two transmitters and two receivers. M_1, M_2 are two messages, X_1, X_2 are the inputs, Y_1, Y_2 are the outputs, and $p(y_1, y_2|x_1, x_2)$ is the transition probability of channel.

II. Preliminaries and Definitions

The two-user DM-CIC is an interference channel that consists of two transmitter-receiver pairs, in which the *cognitive* transmitter non-causally knows the message of the *primary* user, in addition to its own message. In what follows, we formally define DM-CIC and two special classes of that.

A. Discrete memoryless cognitive interference channel

The DM-CIC is depicted in Fig. 1. Let M_1 and M_2 be two independent messages which are uniformly distributed on the sets of all messages of the first and second users, respectively. Transmitter i wishes to transmit message M_i to receiver i, in n channel use at rate R_i, and $i = 1, 2$. Message M_2 is available only at transmitter 2, while both transmitters know M_1. This channel is defined by a tuple $(\mathcal{X}_1, \mathcal{X}_2; p(y_1, y_2|x_1, x_2); \mathcal{Y}_1, \mathcal{Y}_2)$ where $\mathcal{X}_1, \mathcal{X}_2$ and $\mathcal{Y}_1, \mathcal{Y}_2$ are input and output alphabets, and $p(y_1, y_2|x_1, x_2)$ is channel transition probability density functions.

The capacity of the DM-CIC is known in the "cognitive less noisy" [10], "strong interference" [14], "weak interference" [1], and "better cognitive decoding" [9] regimes. These capacity results are listed in Table I, and labeled $\mathcal{C}_I$, $\mathcal{C}_{II}$, $\mathcal{C}_{III}$, and $\mathcal{C}'_{III}$, respectively. In all above cases, the cognitive receiver has a better condition (more information) than the primary one in some sense, as it can be understood from the corresponding conditions in Table I.

B. More Capable DM-CIC

Since the second transmitter has complete and non-causal knowledge of both messages, by sending the two messages, it can act like a *broadcast* transmitter. Particularly, in the absence of the first transmitter this channel becomes the well-known DM-BC [15]. In the presence of the primary transmitter, this channel is no longer a BC; however, similar to that in the DM-BC, one can define conditions for which one of the receivers is in a "better" condition than the other one in decoding the messages, e.g., one receiver is *less noisy* or *more capable* than the other [11].

In [12], the authors extended this notion to the DM-CIC, and studied the case where the primary receiver is more capable than the cognitive receiver. This led to the capacity of GCZIC at very strong interference. In what follows, we show that similar to the less noisy DM-CIC [10], and depending on which receiver is in the better condition than the other, two different more capable DM-CIC arises. These two are formally defined in the following.

Definition 1. The DM-CIC is said to be *primary-more-capable* if

$$I(X_1, X_2; Y_1) \geq I(X_1, X_2; Y_2), \tag{1}$$

for all $p(x_1, x_2)$.

Definition 2. The DM-CIC is said to be *cognitive-more-capable* if

$$I(X_1, X_2; Y_2) \geq I(X_1, X_2; Y_1), \tag{2}$$

for all $p(x_1, x_2)$.

It can be noted that in the first case the primary receiver has more information, about transmitted codewords, than the cognitive receiver whereas the reverse is true in the second case. Therefore, given the channel condition, a DM-CIC can be either in the *primary-more-capable* or in the *cognitive-more-capable* regimes. The former was studied in [12]. In this paper, we focus on the latter case.

III. Main Results

In this section, we first introduce a new outer bound on the capacity of the cognitive-more-capable DM-CIC. We then find an alternative representation of this outer bound; the new representation is the same as an achievable rate region for the DM-CIC which is based on superposition coding. Consequently, we establish the capacity region of the cognitive-more-capable DM-CIC in this section.

A. New Outer Bounds

The following provides an outer bound on the capacity of the cognitive-more-capable DM-CIC, defined in (2).

Theorem 1. *Define $\mathcal{R}_o$ as the set of all rate pairs (R_1, R_2) such that*

$$R_1 \leq I(U, X_1; Y_1), \tag{3a}$$
$$R_1 + R_2 \leq I(U, X_1; Y_1) + I(X_2; Y_2|U, X_1), \tag{3b}$$
$$R_1 + R_2 \leq I(X_1, X_2; Y_2), \tag{3c}$$

for the probability distribution $p(u, x_1, x_2)p(y_1, y_2|x_1, x_2)$. Then, for some $p(u, x_1, x_2)$, $\mathcal{R}_o$ provides an outer bound on the capacity region of the cognitive-more-capable DM-CIC, defined by (2).

Proof. The proof is presented in Section VI-A. □

Theorem 2. *Let $\mathcal{R}'_o$ be the set of all rate pairs (R_1, R_2) such that*

$$R_1 \leq I(U, X_1; Y_1), \quad (4a)$$
$$R_2 \leq I(X_2; Y_2|U, X_1), \quad (4b)$$
$$R_1 + R_2 \leq I(X_1, X_2; Y_2), \quad (4c)$$

for the probability distribution $p(u, x_1, x_2)p(y_1, y_2|x_1, x_2)$. Then $\mathcal{R}'_o \equiv \mathcal{R}_o$, that is, $\mathcal{R}'_o$ gives another representation of $\mathcal{R}_o$ and makes an outer bound on the capacity region of the cognitive-more-capable DM-CIC, for some $p(u, x_1, x_2)$.

Proof. To prove this we consider the following two cases:
case 1: when (3b) is redundant in $\mathcal{R}_o$, i.e.,

$$I(U, X_1; Y_1) + I(X_2; Y_2|U, X_1) \geq I(X_1, X_2; Y_2). \quad (5)$$

case 2: when (3c) is redundant in $\mathcal{R}_o$, i.e.,

$$I(U, X_1; Y_1) + I(X_2; Y_2|U, X_1) \leq I(X_1, X_2; Y_2). \quad (6)$$

In the first case, we can see that (4b) becomes redundant also. This is because the right-hand side of (4b) is greater than or equal to the difference between the right-hand sides of (4c) and (4a) if (5) holds. Consequently, the remaining constraints in $\mathcal{R}_o$ and $\mathcal{R}'_o$ are the same, and $\mathcal{R}'_o \equiv \mathcal{R}_o$.

In the second case, it is obvious that the third inequality is redundant both in (3) and (4). Therefore, the set of constraints in Theorem 1 reduces to

$$R_2 \leq I(U, X_2; Y_2), \quad (7a)$$
$$R_1 + R_2 \leq I(U, X_2; Y_2) + I(X_1; Y_1|U, X_2). \quad (7b)$$

Similarly, the set of constraints in Theorem 2 reduces to

$$R_1 \leq I(X_1; Y_1|U, X_2), \quad (8a)$$
$$R_2 \leq I(U, X_2; Y_2). \quad (8b)$$

Let $\mathcal{R}_{o1}$ denote the union of all rate pairs (R_1, R_2) that satisfy (7a)-(7b) and $\mathcal{R}'_{o1}$ be the union of all rate pairs (R_1, R_2) that satisfy (8a)-(8b); we show that $\mathcal{R}_{o1} \equiv \mathcal{R}'_{o1}$. Intuitively, the convex hull of these two regions is the same since the corner point of both regions, which remains after applying the convex hull operation, are exactly the same. More formally, using the same argument as El Gamal [16], we can see that any point on the boundary of $\mathcal{R}'_{o1}$ is also on the boundary of $\mathcal{R}_{o1}$. This is because R_2 can be thought of as the rate of common message that can be decoded at both receivers while R_1 is the rate of the private message. Now $(R_2, R_1) \in \mathcal{R}'_{o1}$ if and only if $(R_2 - t, R_1 + t) \in \mathcal{R}'_{o1}$ for any $0 \leq t \leq R_2$. In other words, the common rate R_2 can be partly or wholly private. Thus region $\mathcal{R}'_{o1}$ can be represented as $\mathcal{R}_{o1}$, i.e., $\mathcal{R}_{o1} \equiv \mathcal{R}'_{o1}$.[1]

Therefore, the proof of Theorem 2 is completed as in the both cases $\mathcal{R}'_o \equiv \mathcal{R}_o$. □

It should be indicated that the outer bounds in Theorem 1 and Theorem 2 are valid only for the cognitive-more-capable DM-CIC, defined in (2). However, if we remove the third inequalities (i.e., (3c) and (4c)) in those sets of inequalities, the remaining constraints in each set provide outer bounds for any DM-CIC. Therefore, as a corollary of Theorem 2 we have

Corollary 1. *The set of all rate pairs (R_1, R_2) such that*

$$R_1 \leq I(U, X_1; Y_1), \quad (9a)$$
$$R_2 \leq I(X_2; Y_2|U, X_1), \quad (9b)$$

for some $p(u, x_1, x_2)$ provides an outer bound on the capacity region of the DM-CIC.

Corollary 1 gives a simpler and more tractable representation of the outer bound introduced in [1, Theorem 3.2].

We next provide an achievable rate regions for the DM-CIC.

B. An Achievable Rate Region

The following theorem gives an achievable rate regions for the DM-CIC.

Theorem 3. *The union of rate regions given by*

$$R_1 \leq I(W, X_1; Y_1), \quad (10a)$$
$$R_2 \leq I(X_2; Y_2|W, X_1), \quad (10b)$$
$$R_1 + R_2 \leq I(X_1, X_2; Y_2), \quad (10c)$$

is achievable for the DM-CIC, where the union is over all probability distributions $p(w, x_1, x_2)$.

Proof. The proof of Theorem 3 uses the superposition coding idea in which Y_1 can only decode M_1 while Y_2 (the more capable receiver) is intended to decode both M_1 and M_2. Considering the space of all codewords, one can view the (W, X_1) as *cloud centers*, and the X_2 as *satellites* [17]. The decoding is based on joint typicality. The details of the proof can be found in [10]. □

C. The Capacity of the Cognitive-More-Capable DM-CIC

The capacity region of the cognitive-more-capable DM-CIC is established immediately in light of the outer bound in Theorem 2 and the inner bound in Theorem 3. That is, the region define by $\mathcal{R}'_o$, or equivalently the rate region characterized in Theorem 3, gives the capacity region of the DM-CIC when (2) holds.

Theorem 4. *For the cognitive-more-capable DM-CIC defined in* (2)*, the capacity region is given by the set of all rate pairs (R_1, R_2) such that*

$$R_1 \leq I(W, X_1; Y_1), \quad (11a)$$
$$R_2 \leq I(X_2; Y_2|W, X_1), \quad (11b)$$
$$R_1 + R_2 \leq I(X_1, X_2; Y_2), \quad (11c)$$

[1] Similarly, as stated in [11, Chapter 5] when proving the capacity of less noisy BC, the (convex hull of) region

$$R_2 \leq I(U; Y_2),$$
$$R_1 + R_2 \leq I(U; Y_2) + I(X_1; Y_1|U),$$

is an alternative characterization of

$$R_1 \leq I(X_1; Y_1|U),$$
$$R_2 \leq I(U; Y_2),$$

for some for some $p(u, x)$. By replacing U with (U, X_2) in these two regions we will get $\mathcal{R}_{o1}$ and $\mathcal{R}'_{o1}$, respectively.

for some $p(w, x_1, x_2)$.

Theorem 4 gives the largest capacity region for the DM-CIC channel to date. We prove this in the next section by showing that this capacity result contains all existing capacity results of the DM-CIC channel as its subsets.

IV. Comparison and Classification

In this section, we compare the capacity region obtained in Theorem 4 with all of the previously known capacity results for the DM-CIC. For ease of comparison, these results are summarized in Table I. We show that the capacity region of the cognitive-more-capable DM-CIC contains all other capacity regions listed in Table I, as subsets. We also clarify the relation between the other capacity results. More precisely, we prove that

$$\mathcal{C}_I \subseteq \mathcal{C}_{II} \subseteq \mathcal{C}_{III} \equiv \mathcal{C}'_{III} \subseteq \mathcal{C}_{IV}. \tag{12}$$

We first observe that

$$\mathcal{C}_I \subseteq \mathcal{C}'_{III} \subseteq \mathcal{C}_{IV}. \tag{13}$$

This is evident by conditions corresponding to $\mathcal{C}_I, \mathcal{C}'_{III}, \mathcal{C}_{IV}$ in Table I, because

$$\begin{aligned} & I(U;Y_1) \leq I(U;Y_2) \quad \forall\, p(u) \\ \Rightarrow\ & I(U,X_1;Y_1) \leq I(U,X_1;Y_2) \quad \forall\, p(u,x_1) \\ \Rightarrow\ & I(X_2,X_1;Y_1) \leq I(X_2,X_1;Y_2) \quad \forall\, p(u,x_1,x_2). \end{aligned}$$

We next prove that

$$\mathcal{C}_I \subseteq \mathcal{C}_{II}. \tag{14}$$

To show this we resort to a different representation of the "strong interference" condition, for which $\mathcal{C}_{II}$ hold. From [7, eq. (87)-(88)] we know that the DM-CIC is in the "strong interference" regime if

$$I(X_2;Y_2|X_1) \leq I(X_2;Y_1|X_1), \tag{15a}$$
$$I(X_1,X_2;Y_1) \leq I(X_1,X_2;Y_2), \tag{15b}$$

for all $p(x_1, x_2)$. Also, from [13, eq. (8a)-(8b)] we know that this set of conditions is equivalent to

$$I(U;Y_1|X_1) = I(U;Y_2|X_1), \tag{16a}$$
$$I(X_1;Y_1) \leq I(X_1;Y_2), \tag{16b}$$

for all $p(u, x_1, x_2)$.
Now, in light of (16), it is straightforward to show that the condition required for the cognitive-less-noisy regime, i.e.,

$$I(U;Y_1) \leq I(U;Y_2), \tag{17}$$

implies the conditions required for the strong interference regime. To prove this we show that (17) implies both (16a) and (16b). First, we see that if $I(U;Y_1) \leq I(U;Y_2)$ holds for any $p(u, x_1, x_2)$ then we obtain $I(U;Y_1|X_1) \leq I(U;Y_2|X_1)$ for all $p(u, x_1, x_2)$, thus (17) $\Rightarrow$ (16a). Also, for $U = X_2$ the condition in (17) reduces (16b), i.e., (17) $\Rightarrow$ (16b). Hence, the condition required for the cognitive-less-noisy regime implies that of the strong interference regime; this means that (14) holds. Finally, by virtue of (13) and (14) and the fact that $\mathcal{C}_{II} \subseteq \mathcal{C}_{III} \equiv \mathcal{C}'_{III}$ (see [13, Claim 1]) it is obvious that (12) is correct. In words, for a DM-CIC the followings are correct:

1) The better-cognitive-decoding and weak interference are equivalent (see [13]).
2) If a DM-CIC is in the strong interference regime then it is in the cognitive-more-capable regime, as well.
3) If a DM-CIC is in the cognitive-less-noisy regime then it is in the strong interference regime.
4) A better-cognitive-decoding DM-CIC is cognitive-more-capable, too.

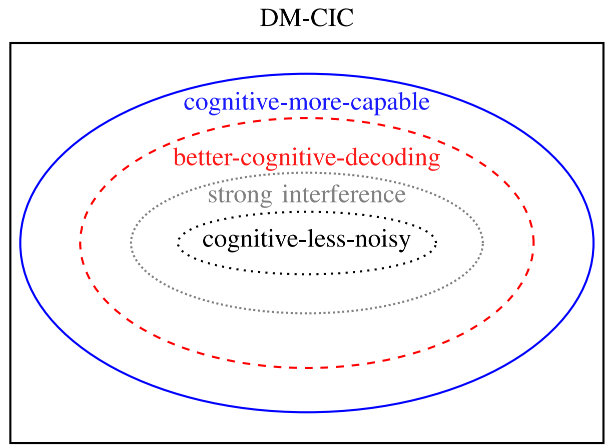

Fig. 2. The class of the discrete memoryless cognitive interference channels (DM-CIC). The cognitive receiver is *superior* than the primary receiver for the cognitive-more-capable and all its subclasses. The largest ellipse (blue, solid line) represents the cognitive-more-capable DM-CIC. The ellipse with red, dashed lines represents the better-cognitive-decoding DM-CIC; note that, this regime is equivalent to weak interference regime. The other two ellipses, i.e., dotted and densely dotted ellipses, respectively show the cognitive-less-noisy and strong interference DM-CIC.

Note that, the converse of statements 2, 3, and 4 does not hold in general. Figure 2 represents these relations, pictorially. In light of the above classifications, the constraints characterizing the capacity region of the cognitive-more-capable DM-CIC (i.e., $\mathcal{C}_{IV}$) can be used to represent the capacity region of all other classes of the DM-CIC listed in Table I.

Remark 1. The constraints in Theorem 4 provide the capacity region of the DM-CIC at the cognitive-less-noisy, strong interference, weak interference, better-cognitive-decoding, and cognitive-more-capable regimes. Furthermore, when the condition corresponding to each one of those subclasses holds, $\mathcal{C}_{IV}$ reduces to the corresponding capacity result.[2] For one thing, if a cognitive-more-capable DM-CIC further satisfies $I(U,X_1;Y_1) \leq I(U,X_1;Y_2)$ then it is easy to see that the third

[2]To better appreciate this, we may think of the two well-known "less noisy" and "more capable" BC and the relation between their capacity. We know that the condition required for a less noisy BC implies that of more capable BC [11]. That is, any less noisy BC is more capable also. Hence, the capacity of more capable BC reduces to that of less noisy BC once the condition required for less noisy BC is met. This is clear from the capacity regions of these two channels [11].

TABLE I

SUMMARY OF EXISTING AND NEW CAPACITY RESULTS FOR THE DM-CIC. THE SUBSCRIPTS 1 AND 2, RESPECTIVELY, DENOTE THE PRIMARY AND SECONDARY (COGNITIVE) USERS.*

Label	DM-CIC class	Condition	Capacity region	Reference
$\mathcal{C}_I$	cognitive-less-noisy	$I(U;Y_1) \leq I(U;Y_2)$	$R_1 \leq I(U;Y_1)$ $R_2 \leq I(X_2;Y_2\|U)$	[10]
$\mathcal{C}_{II}$	strong interference	$I(X_1,X_2;Y_1) \leq I(X_1,X_2;Y_2)$ $I(X_2;Y_2\|X_1) \leq I(X_2;Y_1\|X_1)$	$R_1 \leq I(X_1;Y_1)$ $R_2 \leq I(X_2;Y_2\|X_1)$	[7]
$\mathcal{C}_{III}$	weak interference	$I(X_1;Y_1) \leq I(X_1;Y_2)$ $I(U;Y_1\|X_1) \leq I(U;Y_2\|X_1)$	$R_1 \leq I(U,X_1;Y_1)$ $R_2 \leq I(X_2;Y_2\|U,X_1)$	[1]
$\mathcal{C}'_{III}$	better-cognitive-decoding	$I(U,X_1;Y_1) \leq I(U,X_1;Y_2)$	$R_1 \leq I(U,X_1;Y_1)$ $R_2 \leq I(X_2;Y_2\|X_1)$ $R_1+R_2 \leq I(U,X_1;Y_1)+I(X_2;Y_2\|U,X_1)$	[9]
$\mathcal{C}_{IV}$	cognitive-more-capable	$I(X_1,X_2;Y_1) \leq I(X_1,X_2;Y_2)$	$R_1 \leq I(U,X_1;Y_1)$ $R_2 \leq I(X_2;Y_2\|U,X_1)$ $R_1+R_2 \leq I(X_1,X_2;Y_2)$	Theorem 4

* It should be emphasized that $\mathcal{C}'_{III} \equiv \mathcal{C}_{III}$ [13] and $\mathcal{C}_I \subseteq \mathcal{C}_{II} \subseteq \mathcal{C}_{III} \subseteq \mathcal{C}_{IV}$.

constraint in $\mathcal{C}_{IV}$ becomes redundant, and the capacity region corresponding to the better-cognitive-decoding, is achieved. Hence $\mathcal{C}_{IV}$ unifies the representation of the capacity results for the DM-CIC in different regimes.

Remark 2. Superposition coding is optimal for several classes of the DM-CIC for which the cognitive receiver is superior than the primary one, as we detailed in this section. It is, however, not optimal in general since requiring the cognitive receiver to recover both messages (even though non-uniquely for the primary's message) can excessively constraint the achievable rate region.

Remark 3. All of the classes defined in Table I and depicted in Fig. 2, imply the superiority of the cognitive receiver the primary one. It is worth noting that, by swapping the indices 1 and 2 in the conditions, we can define similar classes in which the primary receiver is superior than the cognitive one. One may expect similar capacity results in the new cases by using superposition encoding in a different order. But it cannot come true because the factorization of probability distribution $p(u,x_1,x_2)$ is different since the primary encoder does not know x_2; thus, similar rate regions are not attainable over general distribution $p(u,x_1,x_2)$. This has been noted in [12].

V. CONCLUSIONS

We have established the capacity of a new class of DM-CIC, named the cognitive-more-capable DM-CIC, which gives the largest capacity region for the DM-CIC up to now. This is proved by showing that all previously known capacity regions, for the DM-CIC, are subsets of this new result, which is obtained by using superposition coding at the cognitive transmitter. The analysis of the other capacity results of the DM-CIC shows that superposition coding is the capacity-achieving techniques in those cases, too. Besides, we make a logical link between the different capacity results for this channel. This sheds more light on the existing capacity results and unifies all of them under the capacity region in the new regime.

REFERENCES

[1] W. Wu, S. Vishwanath, and A. Arapostathis, "Capacity of a class of cognitive radio channels: Interference channels with degraded message sets," *IEEE Transactions on Information Theory*, vol. 53, pp. 4391–4399, May 2007.

[2] A. Carleial, "Interference channels," *IEEE Transactions on Information Theory*, vol. 24, pp. 60–70, January 1978.

[3] G. Kramer, "Review of rate regions for interference channels," in *Proc. International Zurich Seminar on Communications*, pp. 162–165, February 2006.

[4] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 23, pp. 201–220, February 2005.

[5] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive channels," *IEEE Transactions on Information Theory*, vol. 52, pp. 1813–1827, May 2006.

[6] A. Jovicic and S. Vishwanath, "Cognitive radio: An information-theoretic perspective," *IEEE Transactions on Information Theory*, vol. 55, pp. 3945–3958, September 2009.

[7] I. Maric, R. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Transactions on Information Theory*, vol. 53, pp. 3536–3548, October 2007.

[8] I. Maric, A. Goldsmith, G. Kramer, and S. Shamai, "On the capacity of interference channels with one cooperating transmitter," *European Transactions Telecommunications*, vol. 19, pp. 405–420, April 2008.

[9] S. Rini, D. Tuninetti, and N. Devroye, "New inner and outer bounds for the memoryless cognitive interference channel and some new capacity results," *IEEE Transactions on Information Theory*, vol. 57, pp. 4087–4109, July 2011.

[10] M. Vaezi, "The capacity of less noisy cognitive interference channels," in *Proc. Fiftieth Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1769–1774, Oct. 2012.

[11] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[12] M. Vaezi and M. Vu, "On the capacity of the cognitive Z-interference channel," in *Proc. 12th Canadian Workshop on Information Theory (CWIT)*, pp. 30–33, May 2011.

[13] M. Vaezi, "Comments on new inner and outer bounds for the memoryless cognitive interference channel and some new capacity results," *IEEE Transactions on Information Theory*, vol. 59, pp. 4055–4056, June 2013.

[14] I. Maric, R. Yates, and G. Kramer, "The strong interference channel with unidirectional cooperation," in *Proc. Information Theory and Applications (ITA) Inaugural Workshop*, February 2006.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 2006.

[16] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Transactions on Information Theory*, vol. 25, pp. 166–169, March 1979.

[17] G. Kramer, *Topics in Multi-User Information Theory*. Now Publishers, 2007.

VI. Appendix

A. Proof of Theorem 1

The first two constraints in this outer bound (i.e., (3a) ,(3b)), which make an outer bound on the capacity of any DM-CIC, are proved in [1, Theorem 3.2]. Here, we prove that the last constraint (3c) holds for the cognitive-more-capable DM-CIC, i.e., a DM-CIC that satisfies (2). To do so, we can bound the rates $R_2 + R_2$ as

$$\begin{aligned} n(R_1 + R_2) &= H(M_1, M_2) \\ &= H(M_1|M_2) + H(M_2) \\ &= I(M_1; Y_1^n|M_2) + H(M_1|Y_1^n, M_2) \\ &\quad + I(M_2; Y_2^n) + H(M_2|Y_2^n) \\ &\leq I(M_1; Y_1^n|M_2) + I(M_2; Y_2^n) + n\epsilon_n \qquad (18) \end{aligned}$$

where (18) follows by Fano's inequality.

Next, we bound the mutual information terms on the right-hand side of the inequality in (18).

$$\begin{aligned} &I(M_1; Y_1^n|M_2) + I(M_2; Y_2^n) \\ &= \sum_{i=1}^n I(M_1; Y_{1i}|M_2, Y_1^{i-1}) + \sum_{i=1}^n I(M_2; Y_{2i}|Y_{2,i+1}^n) \qquad (19a) \\ &\leq \sum_{i=1}^n I(M_1, Y_{2,i+1}^n; Y_{1i}|M_2, Y_1^{i-1}) + \sum_{i=1}^n I(M_2, Y_{2,i+1}^n; Y_{2i}) \\ &= \sum_{i=1}^n I(M_1, Y_{2,i+1}^n; Y_{1i}|M_2, Y_1^{i-1}) \\ &\quad + \sum_{i=1}^n I(M_2, Y_{2,i+1}^n, Y_1^{i-1}; Y_{2i}) - \sum_{i=1}^n I(Y_1^{i-1}; Y_{2i}|M_2, Y_{2,i+1}^n) \\ &= \sum_{i=1}^n I(M_1; Y_{1i}|M_2, Y_1^{i-1}, Y_{2,i+1}^n) + \sum_{i=1}^n I(M_2, Y_{2,i+1}^n, Y_1^{i-1}; Y_{2i}) \\ &\quad - \sum_{i=1}^n I(Y_1^{i-1}; Y_{2i}, |M_2, Y_{2,i+1}^n) + \sum_{i=1}^n I(Y_{2,i+1}^n; Y_{1i}, |M_2, Y_1^{i-1}) \\ &= \sum_{i=1}^n I(M_1; Y_{1i}|V_i) + \sum_{i=1}^n I(V_i; Y_{2i}) \qquad (19b) \\ &\leq \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_{1i}|V_i) + \sum_{i=1}^n I(V_i; Y_{2i}) + n\epsilon_n \qquad (19c) \\ &\leq \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_{2i}|V_i) + \sum_{i=1}^n I(V_i; Y_{2i}) + n\epsilon_n \qquad (19d) \\ &= \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_{2i}) + n\epsilon_n \end{aligned}$$

in which (19a) follows by the chain rule; (19b) follows by the Csiszar sum identity and the auxiliary random variable $V_i \triangleq (M_2, Y_1^{i-1}, Y_{2,i+1}^n)$; (19c) follows from $M_1 \to (X_1, X_2) \to Y_1$; (19d) follows from the cognitive-more-capable condition in (2) that gives $I(X_1, X_2; Y_2) \geq I(X_1, X_2; Y_1)$, and implies that $I(X_1, X_2; Y_1|V) \leq I(X_1, X_2; Y_2|V)$.

Next, we define the time sharing random variable Q which is uniformly distributed over $[1 : n]$ and is independent of $(M_1, M_2, X_1^n, X_2^n, Y_1^n, Y_2^n)$. Also we define $X_1^n = X_{1Q}^n$, $X_2^n = X_{2Q}^n$, and $Y_2^n = Y_{2Q}^n$. Then we have

$$\begin{aligned} n(R_1 + R_2) &\leq \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_{2i}) + n\epsilon_n \\ &= nI(X_1, X_2; Y_2|Q) + n\epsilon_n \\ &\leq nI(X_1, X_2; Y_2) + n\epsilon_n. \end{aligned}$$

But $\epsilon_n \to 0$, as $n \to \infty$, because the probability of error is assumed to vanish. This completes the proof.