State of the cognitive interference channel: a new unified inner bound

Stefano Rini, Daniela Tuninetti and Natasha Devroye
University of Illinois at Chicago
Chicago, IL 60607, USA
Email: srini2, danielat, devroye@uic.edu

***Abstract*—The capacity region of the interference channel in which one transmitter non-causally knows the message of the other, termed the cognitive interference channel, has remained open since its inception in 2005. A number of subtly differing achievable rate regions and outer bounds have been derived, some of which are tight under specific conditions. In this work we present a new unified inner bound for the discrete memoryless cognitive interference channel. We show explicitly how it encompasses all known discrete memoryless achievable rate regions as special cases. The presented achievable region was recently used in deriving the capacity region of the linear high-SNR deterministic approximation of the Gaussian cognitive interference channel. The high-SNR deterministic approximation was then used to obtain the capacity of the Gaussian cognitive interference channel to within 1.87 bits.**

I. Introduction

The cognitive interference channel (CIFC)[1] is an interference channel in which one of the transmitters - dubbed the cognitive transmitter - has non-causal knowledge of the message of the other - dubbed the primary - transmitter. The study of this channel is motivated by cognitive radio technology which allows wireless devices to sense and adapt to their RF environment by changing their transmission parameters in software on the fly. One of the driving applications of cognitive radio technology is secondary spectrum sharing: currently licensed spectrum would be shared by primary (legacy) and secondary (usually cognitive) devices in the hope of improving spectral efficiency. The extra abilities of cognitive radios may be modeled information theoretically in a number of ways - see [6], [11] for surveys - one of which is through the assumption of non-causal primary message knowledge at the secondary, or cognitive, transmitter.

The two-dimensional capacity region of the CIFC has remained open in general since its inception in 2005 [7]. However, capacity is known in a number of classes of channels:

• **General deterministic CIFCs.** The capacity region of fully deterministic CIFCs in the flavor of the deterministic interference channel [1] has been obtained in [24]. A special case of the deterministic CIFC is the deterministic linear high-SNR approximation of the Gaussian CIFC, whose capacity region, in the spirit of [2], was obtained in [23].

• **Semi-deterministic CIFCs.** In [4] the capacity region for a class of channels in which the signal at the cognitive receiver is a deterministic function of the channel inputs is derived.

• **Discrete memoryless CIFCs.** First considered in [7], [8], its capacity region was obtained for very strong interference in [13] and for weak interference in [30]. Prior to this work and the recent work of [4], the largest known achievable rate regions were those of [8], [9], [15], [20]. The recent and independently derived region of [4] was shown to contain [15], [20], but was not conclusively shown to encompass [8] or the larger region of [9].

• **Gaussian CIFC.** The capacity region under weak interference was obtained in [16], [30], while that for very strong interference follows from [13]. Capacity for a class of Gaussion MIMO CIFCs is obtained in [28].

• **Z-CIFCs.** Inner and outer bounds when the cognitive-primary link is noiseless are obtained in [3], [19]. The Gaussian causal case is considered in [4], and is related to the general (non Z) causal CIFC explored in [26].

• **CIFCs with secrecy constraints.** Capacity of a CIFC in which the cognitive message is to be kept secret from the primary and the cognitive wishes to decode both messages is obtained in [18]. A cognitive multiple-access wiretap channel is considered in [27].

We focus on the discrete memoryless CIFC (DM-CIFC) and propose a new achievable rate region which encompasses all other known achievable rate regions. We will explicitly demonstrate how our new region encompasses and may be reduced to the other regions. The new unified achievable rate region has been shown to be useful as: 1) specific choices of random variables yield capacity in the deterministic CIFC [24] and hence also in the 2) linear high-SNR approximation of the Gaussian CIFC [23], 3) specific choices of Gaussian random variables have resulted in an achievable rate region which lies within 1.87 bits, regardless of channel parameters, of an outer bound [25]. Numerical simulations indicate the actual gap is smaller.

II. Channel Model

The Discrete Memoryless Cognitive InterFerence Channel (DM-CIFC), as shown in Fig. 1, consists of two transmitter-receiver pairs that exchange independent messages over a common channel. Transmitter i, $i \in \{1, 2\}$, has discrete input

[1] Other names for this channel include the cognitive radio channel [8], interference channel with degraded message sets [15], [30], the non-causal interference channel with one cognitive transmitter [4], the interference channel with one cooperating transmitter [20] and the interference channel with unidirectional cooperation [13], [21].

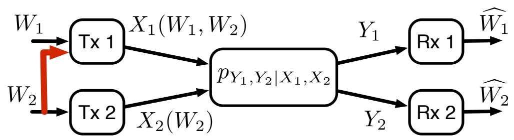

Fig. 1. The Cognitive Interference Channel.

alphabet $\mathcal{X}_i$ and its receiver has discrete output alphabet $\mathcal{Y}_i$. The channel is assumed to be memoryless with transition probability $p_{Y_1,Y_2|X_1,X_2}$. Encoder i, $i \in \{1,2\}$, wishes to communicate a message W_i uniformly distributed on $\mathcal{M}_i = [1 : 2^{NR_i}]$ to decoder i in N channel uses at rate R_i. Encoder 1 (i.e., the cognitive user) knows its own message W_1 and that of encoder 2 (the primary user), W_2. A rate pair (R_1, R_2) is achievable if there exist sequences of encoding functions

$$X_1^N = f_1^N(W_1, W_2), \quad f_1 : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{X}_1^N,$$
$$X_2^N = f_2^N(W_2), \quad f_2 : \mathcal{M}_2 \to \mathcal{X}_2^N,$$

with corresponding sequences of decoding functions

$$\widehat{W}_1 = g_1^N(Y_1^N), \quad g_1 : \mathcal{Y}_1^N \to \mathcal{M}_1,$$
$$\widehat{W}_2 = g_2^N(Y_2^N), \quad g_2 : \mathcal{Y}_2^N \to \mathcal{M}_2.$$

The capacity region is defined as the closure of the region of achievable (R_1, R_2) pairs [5]. Standard strong-typicality is assumed; properties may be found in [17].

III. A NEW UNIFIED ACHIEVABLE RATE REGION

As the DM-CIFC encompasses classical interference, multiple-access and broadcast channels, we expect to see a combination of their achievability proving techniques surface in any unified scheme for the CIFC:

- **Rate-splitting.** As in Han and Kobayashi [12] for the interference-channel and in the DM-CIFC regions of [8], [15], [20], rate-splitting is not necessary in the weak [30] and strong [13] interference regimes.
- **Superposition-coding.** Useful in multiple-access and broadcast channels [5], in the CIFC the superposition of private messages on top of common ones [15], [20] is proposed and is known to be capacity achieving in very strong interference [13].
- **Binning.** Gel'fand-Pinsker coding [10], often referred to as binning, allows a transmitter to "cancel" (portions of) the interference known to it at its intended receiver. Related binning techniques are used by Marton in deriving the largest known DM-broadcast channel achievable rate region [22].

We now present a new achievable region for the DM-CIFC which generalizes all best known achievable rate regions including [8], [15], [20], [30] as well as [4].

Theorem 1. *Region $\mathcal{R}_{RTD}$. A rate pair (R_1, R_2) such that*

$$R_1 = R_{1c} + R_{1pb}, \quad (1)$$
$$R_2 = R_{2c} + R_{2pa} + R_{2pb} \quad (2)$$

is achievable for a DM-CIFC if $(R'_{1c}, R'_{1pb}, R'_{2pb}, R_{1c}, R_{1pb}, R_{2c}, R_{2pa}, R_{2pb}) \in \mathbb{R}_+^8$ satisfies (3a)–(3k) *for some input distribution*

$$p_{U_{1c},U_{2c},U_{1pb},U_{2pb}} p_{X_1,X_2|U_{1c},U_{2c},U_{1pb},U_{2pb}} p_{Y_1,Y_2|X_1,X_2}.$$

The encoding scheme used in deriving this achievable rate region is shown in Fig.2. The key aspects of our scheme are the following, where we drop n for convenience:

- We **rate-split** the independent messages W_1 and W_2 uniformly distributed on $\mathcal{M}_1 = [1 : 2^{nR_1}]$ and $\mathcal{M}_2 = [1 : 2^{nR_2}]$ into the messages W_i, $i \in \{1c, 2c, 1pb, 2pb, 2pa\}$, all independent and uniformly distributed on $[1 : 2^{nR_i}]$, each encoded using the random variable U_i, such that

$$W_1 = (W_{1c}, W_{1pb}), \qquad R_1 = R_{1c} + R_{1pb},$$
$$W_2 = (W_{2c}, W_{2pb}, W_{2pa}), \quad R_2 = R_{2c} + R_{2pa} + R_{2pb}.$$

- **Tx2 (primary Tx):** Transmitter 2 sends X_2 that carries the private message W_{2pa} ("p" for private, "a" for alone) **superimposed** to the common message W_{2c} carried by U_{2c} ("c" for common).
- **Tx1 (cognitive Tx):** The common message of Tx1, encoded by U_{1c}, is **binned** against X_2 conditioned on U_{2c}. The private message of Tx2, W_{2pb}, encoded by U_{2pb} ("b" for broadcast) and a portion of the private message of Tx1, W_{1pb}, encoded as U_{1pb}, are **binned** against each other as in Marton's region [22] conditioned on U_{1c}, U_{2c} and U_{1c}, U_{2c}, X_2 respectively.

Tx1 sends X_1 over the channel. The incorporation of a Marton-like scheme at the cognitive transmitter was initially motivated by the fact that in certain regimes, this strategy was shown to be capacity achieving for the linear high-SNR deterministic CIFC [23].

The codebook generation, encoding and decoding as well as the error event analysis is provided in [24].

Remark:
- (3d) can be dropped when $R_{2c} = R_{2pa} = R_{2pb} = R'_{2pb} = 0$
- (3e) can be dropped when $R_{2pa} = R_{2pb} = R'_{2pb} = 0$
- (3g) can be dropped when $R_{2pb} = R'_{2pb} = 0$
- (3i) can be dropped when $R_{1c} = R'_{1c} = R_{1pb} = R'_{1pb} = 0$

IV. COMPARISON WITH EXISTING ACHIEVABLE REGIONS

We now show that the region of Theorem 1 contains all other known achievable rate regions for the DM-CIFC. We note that showing inclusion of the rate regions [4, Thm.2], [14], and [9] is sufficient to demonstrate the largest known DM-CIFC region, since the region of [4] is shown to contain those of [20, Th.1] and [15], and the region of [14] is claimed to contain all others. The region in [9] is explicitly shown, for the first time, to be included in another region.

A. Devroye et al.'s region [9, Thm. 1]

In the appendix we show that the region of [9, Thm. 1] $\mathcal{R}_{DMT}$, is contained in our new region $\mathcal{R}_{RTD}$ along the lines:
- We make a correspondence between the random variables and corresponding rates of $\mathcal{R}_{DMT}$ and $\mathcal{R}_{RTD}$.

$$
\begin{aligned}
{R'}_{1c} &\geq I(U_{1c}; X_2|U_{2c}) && \text{(3a)}\\
{R'}_{1c} + {R'}_{1pb} &\geq I(U_{1pb}, U_{1c}; X_2|U_{2c}) && \text{(3b)}\\
{R'}_{1c} + {R'}_{1pb} + {R'}_{2pb} &\geq I(U_{1pb}, U_{1c}; X_2|U_{2c}) + I(U_{2pb}; U_{1pb}|U_{1c}, U_{2c}, X_2) && \text{(3c)}\\
R_{2c} + R_{2pa} + (R_{1c} + {R'}_{1c}) + (R_{2pb} + {R'}_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}, X_2, U_{2c}) + I(U_{1c}; X_2|U_{2c}) && \text{(3d)}\\
R_{2pa} + (R_{1c} + {R'}_{1c}) + (R_{2pb} + {R'}_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}, X_2|U_{2c}) + I(U_{1c}; X_2|U_{2c}) && \text{(3e)}\\
R_{2pa} + (R_{2pb} + {R'}_{2pb}) &\leq I(Y_2; U_{2pb}, X_2|U_{1c}, U_{2c}) + I(U_{1c}; X_2|U_{2c}) && \text{(3f)}\\
(R_{1c} + {R'}_{1c}) + (R_{2pb} + {R'}_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}|X_2, U_{2c}) + I(U_{1c}; X_2|U_{2c}) && \text{(3g)}\\
(R_{2pb} + {R'}_{2pb}) &\leq I(Y_2; U_{2pb}|U_{1c}, X_2, U_{2c}) && \text{(3h)}\\
R_{2c} + (R_{1c} + {R'}_{1c}) + (R_{1pb} + {R'}_{1pb}) &\leq I(Y_1; U_{1pb}, U_{1c}, U_{2c}), && \text{(3i)}\\
(R_{1c} + {R'}_{1c}) + (R_{1pb} + {R'}_{1pb}) &\leq I(Y_1; U_{1pb}, U_{1c}|U_{2c}), && \text{(3j)}\\
(R_{1pb} + {R'}_{1pb}) &\leq I(Y_1; U_{1pb}|U_{1c}, U_{2c}), && \text{(3k)}
\end{aligned}
$$

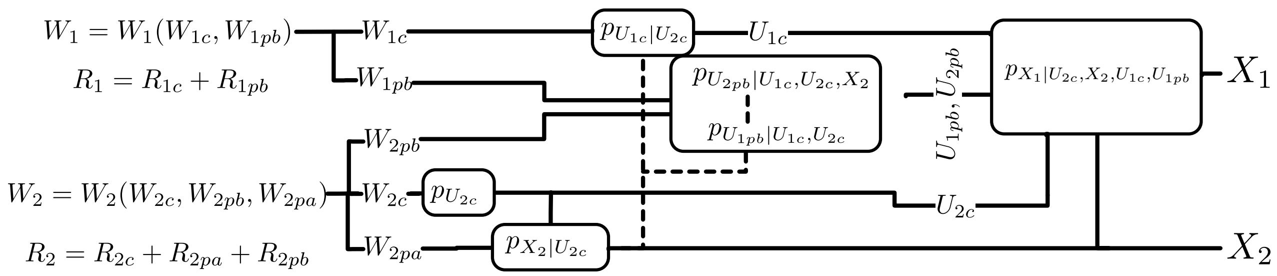

Fig. 2. The achievable encoding scheme of Thm 1. The ordering from left to right and the distributions demonstrate the codebook generation process. The dotted lines indicate binning. We see rate splits are used at both users, private messages $W_{1pb}, W_{2pa}, W_{2pb}$ are superimposed on common messages W_{1c}, W_{2c} and U_{1c} is binned against X_2 conditioned on U_{2c}, while U_{1pb} and U_{2pb} are binned against each and X_2 in a Marton-like fashion (conditioned on other subsets of random variables).

• We define new regions $\mathcal{R}_{DMT} \subseteq \mathcal{R}_{DMT}^{out}$ and $\mathcal{R}_{RTD}^{in} \subseteq \mathcal{R}_{RTD}$ which are easier to compare: they have identical input distribution decompositions and similar rate equations.
• For any fixed input distribution, an equation-by-equation comparison leads to $\mathcal{R}_{DMT} \subseteq \mathcal{R}_{DMT}^{out} \subseteq \mathcal{R}_{RTD}^{in} \subseteq \mathcal{R}_{RTD}$.

B. Cao and Chen's region [4, Thm. 2]

The independently derived region in [4, Thm. 2] uses a similar encoding structure as that of $\mathcal{R}_{RTD}$ with two exceptions: a) the binning is done sequentially rather than jointly as in $\mathcal{R}_{RTD}$ leading to binning constraints (43)–(45) in [4, Thm. 2] as opposed to (3a)–(3c) in Thm.1. Notable is that both schemes have adopted a Marton-like binning scheme at the cognitive transmitter, as first introduced in the context of the CIFC in [3]. b) While the cognitive messages are rate-split in identical fashions, the primary message is split into 2 parts in [4, Thm. 2] ($R_1 = R_{11} + R_{10}$, note the reversal of indices) while we explicitly split the primary message into three parts $R_2 = R_{2c} + R_{2pa} + R_{2pb}$. In the Appendix we show that the region of [4, Thm.2], denoted as $\mathcal{R}_{CC} \subseteq \mathcal{R}_{RTD}$ in two steps:
• We first show that we may WLOG set $U_{11} = \emptyset$ in [4, Thm.2], creating a new region R'_{CC}.
• We next make a correspondence between our random variables and those of [4, Thm.2] and obtain identical regions.

C. Jiang et al.'s region [14, Thm. 4.1]

The scheme originally designed for the more general broadcast channel with cognitive relays (or interference-chanel with a cognitive relay) may be tailored/reduced to derive a region for the cognitive interference channel. This scheme also incorporates a broadcasting strategy. However, the common messages are created independently instead of having the common message from transmitter 1 being superposed to the common message from transmitter 2. The former choice introduces more rate constraints than the latter and allows us to show inclusion in $\mathcal{R}_{RTD}$ after equating random variables.

V. Conclusion

A new achievable rate region for the DM-CIFC has been derived and shown to encompass all known achievable rate regions. Of note is the inclusion of a Marton-like broadcasting scheme at the cognitive transmitter. Specific choices of this region have been shown to achieve capacity for the linear high-SNR approximation of the Gaussian CIFC [23], [24], and the deterministic CIFC in general [24]. This region has furthermore been shown to achieve within 1.87 bits of an outer bound, regardless of channel parameters in [24], [25]. Numerical evaluation of the region under Gaussian input distributions for the Gaussian CIFC is currently underway, while extensions of the CIFC to multiple users will be investigated in the longer term.

REFERENCES

[1] A. El Gamal and M.H.M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 343–346, Mar. 1982.

[2] A. Avestimehr, S. Diggavi, and D. Tse, "A deterministic model for wireless relay networks an its capacity," in *Information Theory for Wireless Networks, 2007 IEEE Information Theory Workshop on*, July 2007, pp. 1–6.

[3] Y. Cao and B. Chen, "Interference channel with one cognitive transmitter," in *Asilomar Conference on Signals, Systems, and Computers*, Oct. 2008.

[4] ——, "Interference Channels with One Cognitive Transmitter," *Arxiv preprint arXiv:09010.0899v1*, 2009.

[5] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[6] N. Devroye, P. Mitran, M. Sharif, S. S. Ghassemzadeh, and V. Tarokh, "Information theoretic analysis of cognitive radio systems," in *Cognitive Wireless Communication Networks*, V. Bhargava and E. Hossain, Eds. Springer, 2007.

[7] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," in *39th Annual Conf. on Information Sciences and Systems (CISS)*, Mar. 2005.

[8] ——, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[9] N. Devroye, "Information theoretic limits of cognition and cooperation in wireless networks," Ph.D. dissertation, Harvard University, 2007.

[10] S. Gel'fand and M. Pinsker, "Coding for channel with random parameters," *Problems of control and information theory*, 1980.

[11] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, 2009.

[12] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *Information Theory, IEEE Transactions on*, vol. 27, no. 1, pp. 49–60, Jan 1981.

[13] R. D. Yates. I. Maric and G. Kramer, "The strong interference channel with unidirectional cooperation," *The Information Theory and Applications (ITA) Inaugural Workshop*, Feb 2006, uCSD La Jolla, CA,.

[14] J. Jiang, I. Maric, A. Goldsmith, and S. Cui, Achievable rate regions for broadcast channels with cognitive relays, in *Proc. IEEE Information Theory Workshop (ITW 2009)*, Taormina, Italy, Oct. 1116, 2009.

[15] J. Jiang and Y. Xin, "On the achievable rate regions for interference channels with degraded message sets," *Information Theory, IEEE Transactions on*, vol. 54, no. 10, pp. 4707–4712, Oct. 2008.

[16] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *Proc. IEEE Int. Symp. Inf. Theory*, pp. 2413–2417, July 2006.

[17] G. Kramer, *Topics in Multi-User Information Theory*, ser. Foundations and Trends in Communications and Information Theory. Vol. 4: No 45, pp 265-444, 2008.

[18] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai, and S. Verdú, "Capacity of cognitive interference channels with and without secrecy," *IEEE Trans. on Inf. Theory*, vol. 55, no. 2, pp. 604–619, Feb. 2009.

[19] N. Liu, I. Maric, A. Goldsmith, and S. Shamai, "The capacity region of the cognitive z-interference channel with one noiseless component," *http://www.scientificcommons.org/38908274*, 2008. [Online]. Available: http://arxiv.org/abs/0812.0617

[20] I. Maric, A. Goldsmith, G. Kramer, and S. Shamai, "On the capacity of interference channels with a cognitive transmitter," *European Transactions on Telecommunications*, vol. 19, pp. 405–420, Apr. 2008.

[21] I. Maric, R. Yates, and G. Kramer, "The capacity region of the strong interference channel with common information," in *Signals, Systems and Computers, 2005. Conference Record of the Thirty-Ninth Asilomar Conference on*, 2005, pp. 1737–1741.

[22] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *Information Theory, IEEE Transactions on*, vol. 25, no. 3, pp. 306–311, May 1979.

[23] S. Rini, D. Tuninetti, and N. Devroye, "The capacity region of gaussian cognitive radio channels at high snr," *Proc. IEEE ITW Taormina, Italy*, vol. Oct., 2009.

[24] S. Rini, "On the role of cognition and cooperation in wireless networks: an information theoretic perspective - a preliminary thesis," http://sites.google.com/site/rinistefano/my-thesis-proposal.

[25] S. Rini, D. Tuninetti, and N. Devroye, "The capacity region of gaussian cognitive radio channels to within 1.87 bits," *Proc. IEEE ITW Cairo, Egypt*, 2010, http://www.ece.uic.edu/∼devroye/conferences.html.

[26] S. H. Seyedmehdi, Y. Xin, J. Jiang, and X. Wang, "An improved achievable rate region for the causal cognitive radio," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2009.

[27] O. Simeone and A. Yener, "The cognitive multiple access wire-tap channel," in *Proc. Conf. on Information Sciences and Systems (CISS)*, Mar. 2009.

[28] S. Sridharan and S. Vishwanath, "On the capacity of a class of mimo cognitive radios," in *Information Theory Workshop, 2007. ITW '07. IEEE*, Sept. 2007, pp. 384–389.

[29] Willems, F. and Van der Meulen, E., *IEEE Transactions on Information Theory*, no.3-pp 313-327,1985

[30] W. Wu, S. Vishwanath, and A. Arapostathis, "Capacity of a class of cognitive radio channels: Interference channels with degraded message sets," *Information Theory, IEEE Transactions on*, vol. 53, no. 11, pp. 4391–4399, Nov. 2007.

APPENDIX

A. Proof that $X_{2a} = \emptyset$ WLOG in [20, Th.1]

In their notation, after the Fourier-Motzkin elimination of [20, Th.1] we obtain the achievable rate region

$$
\begin{aligned}
R_1 &\leq I(U_{1a}; Y_1|U_{1c}, Q) - I(U_{1a}; X_{2a}, X_{2b}|U_{1c}, Q) \\
&\quad + I(X_{2b}, U_{1c}; Y_2|X_{2a}, Q) \qquad (4)\\
R_1 &\leq I(U_{1a}, U_{1c}; Y_1|Q) - I(U_{1a}, U_{1c}; X_{2a}, X_{2b}|Q) \\
R_2 &\leq I(X_2, U_{1c}; Y_2|Q) \\
R_2 &\leq I(X_2; Y_2, U_{1c}|Q) \\
R_1 + R_2 &\leq I(U_{1a}; Y_1|U_{1c}, Q) \\
&\quad - I(U_{1a}; X_{2a}, X_{2b}|U_{1c}, Q) + I(X_2, U_{1c}; Y_2|Q)
\end{aligned}
$$

for any distribution $p_{X_1,X_2,X_{2a},X_{2b},U_{1c},U_{1a},Q}$. For a given $p_{X_{2a},X_{2b},U_{1c},U_{1a},Q}$ of [20, Th.1] consider a related distribution $p_{X'_{2a},X'_{2b},U'_{1c},U'_{1a},Q'}$ such that

$$
\begin{aligned}
&(U'_{1c}, U'_{1a}, Q') = (U_{1c}, U_{1a}, Q) \\
&X'_{2b} = (X_{2a}, X_{2b}), \quad X'_{2a} = \emptyset
\end{aligned}
$$

All rate constraints but (4) are the same under both distributions. Comparing (4) under the two distributions:

$$
\begin{aligned}
&(4)|_{p_{X'_{2a},X'_{2b},U'_{1c},U'_{1a},Q'}} \\
&= I(U'_{1a}; Y_1|U'_{1c}, Q') - I(U'_{1a}; X'_{2a}, X'_{2b}|U'_{1c}, Q') + I(X'_{2b}, U'_{1c}; Y_2|X'_{2a}, Q') \\
&= I(U_{1a}; Y_1|U_{1c}, Q) - I(U_{1a}; X_{2a}, X_{2b}|U_{1c}, Q) + I(X_{2a}, X_{2b}, U_{1c}; Y_2|Q) \\
&= I(U_{1a}; Y_1|U_{1c}, Q) - I(U_{1a}; X_{2a}, X_{2b}|U_{1c}, Q) + I(X_{2a}; Y_2|Q) + I(X_{2b}, U_{1c}; Y_2|X_{2a}, Q) \\
&= I(X_{2a}; Y_2|Q) + (4)|_{p_{X_{2a},X_{2b},U_{1c},U_{1a},Q}} \\
&\geq (4)|_{p_{X_{2a},X_{2b},U_{1c},U_{1a},Q}}.
\end{aligned}
$$

B. Containment of [9, Thm. 1] in $\mathcal{R}_{RTD}$

We show this inclusion with the following steps:

- We enlarge the region $\mathcal{R}_{DMT}$ by removing some rate constraints.
- We further enlarge the region by enlarging the set of possible input distributions. This allows us to remove the V_{11} and Q from the inner bound. We refer to this region as $\mathcal{R}^{out}_{DMT}$ since is enlarges the original achievable region.
- We make a correspondence between the random variables and corresponding rates of $\mathcal{R}^{out}_{DMT}$ and $\mathcal{R}_{RTD}$.
- We choose a particular subset of $\mathcal{R}_{RTD}$, $\mathcal{R}^{in}_{RTD}$, for which we can more easily show $\mathcal{R}_{DMT} \subseteq \mathcal{R}^{out}_{DMT} \subset \mathcal{R}^{in}_{RTD} \subseteq \mathcal{R}_{RTD}$, since $\mathcal{R}^{out}_{DMT}$ and $\mathcal{R}^{in}_{RTD}$ have identical input distribution decompositions and similar rate bound equations.

Enlarge the region $\mathcal{R}_{DMT}$

We first enlarge the rate region of [9, Thm. 1], $\mathcal{R}_{DMT}$ by removing a number of constraints (specifically, we remove equations (2.6, 2.8, 2.10, 2.13, 2.14, 2.16 2.17) of [9, Thm. 1]) to obtain the region $\mathcal{R}^{out}_{DMT}$ defined as the set of all rate pairs satisfying:

$$R'_{21} = I(V_{21}; V_{11}, V_{12}|W) \tag{5a}$$
$$R'_{22} = I(V_{22}; V_{11}, V_{12}|W) \tag{5b}$$

$$R_{11} \leq I(Y_1, V_{12}, V_{21}; V_{11}|W) \tag{5c}$$
$$R_{21} + R'_{21} \leq I(Y_1, V_{11}, V_{12}; V_{21}|W) \tag{5d}$$
$$R_{11} + R_{21} + R'_{21} \leq I(Y_1, V_{12}; V_{11}, V_{21}|W) + I(V_{11}; V_{21}|W) \tag{5e}$$
$$R_{11} + R_{21} + R'_{21} + R_{12} \leq I(Y_1; V_{11}, V_{21}, V_{12}|W) + I(V_{11}, V_{12}; V_{21}|W) \tag{5f}$$

$$R_{22} + R'_{22} \leq I(Y_2, V_{12}, V_{21}; V_{22}|W) \tag{5g}$$
$$R_{22} + R'_{22} + R_{21} + R'_{21} \leq I(Y_2, V_{12}; V_{22}, V_{21}|W) + I(V_{22}; V_{21}|W) \tag{5h}$$
$$R_{22} + R'_{22} + R_{21} + R'_{21} + R_{12} \leq I(Y_2; V_{22}, V_{21}, V_{12}|W) + I(V_{22}, V_{21}; V_{12}|W). \tag{5i}$$

(5j)

taken over the union of distributions

$$p_W p_{V_{11}} p_{V_{12}} p_{X_1|V_{11},V_{12}} p_{V_{21}|V_{11}V_{12}} p_{V_{22}|V_{11},V_{12}} p_{X_2|V_{11},V_{12},V_{21},V_{22}}.$$

Following the line of thoughts in [29, Appendix D] it is possible to show that without loss of generality we can set X_1 to be a deterministic function of V_{11} and V_{12}, allowing us insert X_1 next to V_{11}, V_{12} as follows:

$$\begin{aligned}
R'_{21} &= I(V_{21}; X_1, V_{11}, V_{12}|W) && \text{(6a)}\\
R'_{22} &= I(V_{22}; X_1, V_{11}, V_{12}|W) && \text{(6b)}\\
\\
R_{11} &\leq I(Y_1, V_{12}, V_{21}; V_{11}|W) && \text{(6c)}\\
R_{21} + R'_{21} &\leq I(Y_1, X_1, V_{11}, V_{12}; V_{21}|W) && \text{(6d)}\\
R_{11} + R_{21} + R'_{21} &\leq I(Y_1, V_{12}; V_{11}, V_{21}|W) + I(V_{11}; V_{21}|W) && \text{(6e)}\\
R_{11} + R_{21} + R'_{21} + R_{12} &\leq I(Y_1; X_1, V_{11}, V_{12}, V_{21}|W) + I(X_1, V_{11}, V_{12}; V_{21}|W) && \text{(6f)}\\
\\
R_{22} + R'_{22} &\leq I(Y_2, V_{12}, V_{21}; V_{22}|W) && \text{(6g)}\\
R_{22} + R'_{22} + R_{21} + R'_{21} &\leq I(Y_2, V_{12}; V_{22}, V_{21}|W) + I(V_{22}; V_{21}|W) && \text{(6h)}\\
R_{22} + R'_{22} + R_{21} + R'_{21} + R_{12} &\leq I(Y_2; V_{22}, V_{21}, V_{12}|W) + I(V_{22}, V_{21}; V_{12}|W) && \text{(6i)}
\end{aligned}$$

Using the factorization of the auxiliary RV's, we may insert X_1 next to V_{11} in equation (6f).

For equation (6c):

$$\begin{aligned}
R_{11} &\leq I(Y_1, V_{12}, V_{21}; V_{11}|W)\\
&= I(Y_1, V_{21}; V_{11}|V_{12}, W) + I(V_{12}; V_{11}|W)\\
&= I(Y_1, V_{21}; V_{11}|V_{12}, W)\\
&= I(Y_1, V_{21}; X_1, V_{11}|V_{12}, W)\\
&= I(Y_1; X_1, V_{11}|V_{12}, V_{21}, W) + I(V_{21}; X_1, V_{11}|V_{12}, W).
\end{aligned}$$

For equation (6d) we have:

$$\begin{aligned}
R_{11} + R_{21} + R'_{21} &\leq I(Y_1, V_{12}; V_{11}, V_{21}|W) + I(V_{11}; V_{21}|W)\\
&= I(Y_1; V_{11}, V_{21}|V_{12}, W) + I(V_{12}; V_{11}, V_{21}|W) + I(V_{11}; V_{21}|W)\\
&= I(Y_1; V_{11}, V_{21}|V_{12}, W) + I(V_{12}; V_{21}|V_{11}, W) + I(V_{11}; V_{21}|W)\\
&= I(Y_1; V_{11}, V_{21}|V_{12}, W) + I(V_{11}, V_{12}; V_{21}|W)\\
&= I(Y_1; X_1, V_{11}, V_{21}|V_{12}, W) + I(X_1, V_{11}, V_{12}; V_{21}|W)
\end{aligned}$$

The original region is thus equivalent to

$$\begin{aligned}
R'_{21} &= I(V_{21}; X_1, V_{11}, V_{12}|W) && \text{(7a)}\\
R'_{22} &= I(V_{22}; X_1, V_{11}, V_{12}|W) && \text{(7b)}\\
\\
R_{11} &\leq I(Y_1; X_1, V_{11}|V_{12}, V_{21}|W) + I(V_{21}; X_1|V_{12}, W) && \text{(7c)}\\
R_{21} + R'_{21} &\leq I(Y_1, X_1, V_{11}, V_{12}; V_{21}|W) && \text{(7d)}\\
R_{11} + R_{21} + R'_{21} &\leq I(Y_1; X_1, V_{11}, V_{21}|V_{12}, W) + I(X_1; V_{21}|W) && \text{(7e)}\\
R_{11} + R_{21} + R'_{21} + R_{12} &\leq I(Y_1; X_1, V_{11}, V_{21}, V_{12}|W) + I(X_1, V_{11}, V_{12}; V_{21}|W) && \text{(7f)}\\
\\
R_{22} + R'_{22} &\leq I(Y_2, V_{12}, V_{21}; V_{22}|W) && \text{(7g)}\\
R_{22} + R'_{22} + R_{21} + R'_{21} &\leq I(Y_2, V_{12}; V_{22}, V_{21}|W) + I(V_{22}; V_{21}|W) && \text{(7h)}\\
R_{22} + R'_{22} + R_{21} + R'_{21} + R_{12} &\leq I(Y_2; V_{22}, V_{21}, V_{12}|W) + I(V_{22}, V_{21}; V_{12}|W) && \text{(7i)}\\
& && \text{(7j)}
\end{aligned}$$

taken over the union over all distributions

$$p_W p_{V_{11}} p_{V_{12}} p_{X_1|V_{11},V_{12}} p_{V_{21}|X_1,V_{11}V_{12}} p_{V_{22}|X_1,V_{11},V_{12}} p_{X_2|X_1,V_{11},V_{12},V_{21},V_{22}}.$$

RV, rate of Theorem 1	RV, rate of [9, Thm. 1]	Comments
U_{2c}, R_{2c}	V_{12}, R_{12}	TX 2 $\rightarrow$ RX 1, RX 2
U_{1c}, R_{1c}	V_{21}, R_{21}	TX 1 $\rightarrow$ RX 1, RX 2
U_{1pb}, R_{1pb}	V_{22}, R_{22}	TX 1 $\rightarrow$ RX 1
X_2, R_{2pa}	X_1', R_{11}	TX 2 $\rightarrow$ RX 2
$U_{2pb} = \emptyset, R'_{2pb} = 0$	–	TX 1 $\rightarrow$ RX 2
$R'_{1c} = I(U_{1c}; X_2 \vert U_{2c})$	$L_{21} - R_{21} = I(V_{21}; V_{11}, V_{12})$	Binning rate
$R'_{1pb} = I(U_{1pb}; X_2 \vert U_{1c}, U_{2c})$	$L_{22} - R_{22} = I(V_{22}; V_{11}, V_{12})$	Binning rates
X_1	X_2	

TABLE I
ASSIGNMENT OF RV'S OF APPENDIX B

Enlarge the input distribution and eliminate V_{11} and W
Now increase the set of possible input distribution of the input by letting V_{11} to have any joint distribution with V_{12}. This is done by substituting $p_{V_{11}}$ with $p_{V_{11}|V_{12}}$ in the expression of the input distribution. With this substitution we have:

$$\begin{aligned}
& p_W p_{V_{11}|V_{12}} p_{V_{12}} p_{X_1|V_{11},V_{12}} p_{V_{21}|X_1,V_{11}V_{12}} p_{V_{22}|X_1,V_{11},V_{12}} p_{X_2|X_1,V_{11},V_{12},V_{21},V_{22}} \\
& \subseteq \; p_W p_{V_{12}} p_{V_{11},X_1|V_{12}} p_{V_{21}|X_1,V_{11}V_{12}} p_{V_{22}|X_1,V_{11},V_{12}} p_{X_2|X_1,V_{11},V_{12},V_{21},V_{22}} \\
& \subseteq \; p_W p_{V_{12}} p_{X_1'|V_{12}} p_{V_{21}|X_1',V_{12}} p_{V_{22}|X_1',V_{12}} p_{X_2|X_1',V_{12},V_{21},V_{22}}
\end{aligned}$$

with $X_1' = (X_1, V_{11})$. Since V_{12} is decoded at both decoders, the time sharing random W may be incorporated with V_{12} without loss of generality and thus can be dropped. The region described in (7) is convex and time sharing does not increase the achievable region since the region is already convex. With these simplifications, the region $\mathcal{R}^{out}_{DMT}$ is now defined as

$$\begin{aligned}
R'_{21} &= I(V_{21}; X_1', V_{12}) && \text{(8a)} \\
R'_{22} &= I(V_{22}; X_1', V_{12}) && \text{(8b)}
\end{aligned}$$

$$\begin{aligned}
R_{11} &\leq I(Y_1; X_1'|V_{12}, V_{21}) + I(V_{21}; X_1|V_{12}) && \text{(8c)} \\
R_{21} + R'_{21} &\leq I(Y_1, X_1', V_{12}; V_{21}) && \text{(8d)} \\
R_{11} + R_{21} + R'_{21} &\leq I(Y_1; X_1', V_{21}|V_{12}) + I(X_1; V_{21}) && \text{(8e)} \\
R_{11} + R_{21} + R'_{21} + R_{12} &\leq I(Y_1; X_1', V_{21}, V_{12}) + I(X_1', V_{12}; V_{21}) && \text{(8f)}
\end{aligned}$$

$$\begin{aligned}
R_{22} + R'_{22} &\leq I(Y_2, V_{12}, V_{21}; V_{22}) && \text{(8g)} \\
R_{22} + R'_{22} + R_{21} + R'_{21} &\leq I(Y_2, V_{12}; V_{22}, V_{21}) + I(V_{22}; V_{21}) && \text{(8h)} \\
R_{22} + R'_{22} + R_{21} + R'_{21} + R_{12} &\leq I(Y_2; V_{22}, V_{21}, V_{12}) + I(V_{22}, V_{21}; V_{12}) && \text{(8i)}
\end{aligned}$$

union over all the distributions

$$p_{V_{12}} p_{X_1'|V_{12}} p_{V_{21}|X_1',V_{12}} p_{V_{22}|X_1',V_{12}} p_{X_2|X_1',V_{12},V_{21},V_{22}}$$

Correspondence between the random variables and rates. When referring to [9] please note that the index of the primary and cognitive user are reversed with respect to our notation (i.e $1 \rightarrow 2$ and vice-versa). Consider the correspondences between the variables of [9, Thm. 1] and those of Theorem 1 in Table I to obtain the region $\mathcal{R}^{out}_{DMT}$ defined as the set of rate pairs satisfying

$$\begin{aligned}
R'_{1c} &= I(U_{1c}; X_2, U_{2c}) && \text{(9a)} \\
R'_{1pb} &= I(U_{1pb}; X_2, U_{2c}) && \text{(9b)} \\
R_{2pa} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_2; U_{1c}, U_{2c}, X_2) + I(X_2, U_{2c}; U_{1c}) && \text{(9c)} \\
R_{2pa} + R_{1c} + R'_{1c} &\leq I(Y_2; X_2, U_{1c}|U_{2c}) + I(X_2; U_{1c}) && \text{(9d)} \\
R_{1c} + R'_{1c} &\leq I(Y_2, X_2, U_{2c}; U_{1c}) && \text{(9e)} \\
R_{2pa} &\leq I(Y_2; X_2|U_{2c}, U_{1c}) + I(U_{1c}; X_2|U_{2c}) && \text{(9f)} \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_1; U_{1pb}, U_{1c}, U_{2c}) + I(U_{1pb}, U_{1c}; U_{2c}) && \text{(9g)} \\
R_{1c} + R_{1pb} + R'_{1c} + R'_{1pb} &\leq I(Y_1, U_{2c}; U_{1pb}, U_{1c}) + I(U_{1pb}; U_{1c}) && \text{(9h)} \\
R_{1pb} + R'_{1pb} &\leq I(Y_1, U_{2c}, U_{1c}; U_{1pb}) && \text{(9i)}
\end{aligned}$$

taken over the union of all distributions

$$p_{U_{2c}} p_{X_2|U_{2C}} p_{U_{1c}|X_2} p_{U_{1pb}|X_2} p_{X_1|X_2,U_{1c},U_{1pb}}. \tag{10}$$

Next, we using the correspondences of the table and restrict the fully general input distribution of Theorem 1 to match the more constrained factorization of (10), obtaining a region $\mathcal{R}_{RTD}^{in} \subseteq \mathcal{R}_{RTD}$ defined as the set of rate tuples satisfying

$$\begin{aligned}
R'_{1c} &= I(U_{1c}; X_2|U_{2c}) &\text{(11a)}\\
R'_{1c} + R'_{1pb} &= I(X_2; U_{1c}, U_{1pb}|U_{2c}) &\text{(11b)}\\
R_{2c} + R_{1c} + R_{2pa} + R'_{1c} &\leq I(Y_2; U_{2c}, U_{1c}, X_2) + I(U_{1c}; X_2|U_{2c}) &\text{(11c)}\\
R_{2pa} + R_{1c} + R'_{1c} &\leq I(Y_2; U_{1c}, X_2|U_{2c}) + I(U_{1c}; X_2|U_{2c}) &\text{(11d)}\\
R_{1c} + R'_{1c} &\leq I(Y_2; U_{1c}|U_{2c}, X_2) + I(U_{1c}; X_2|U_{2c}) &\text{(11e)}\\
R_{2pa} &\leq I(Y_2; X_2|U_{2c}, U_{1c}) + I(U_{1c}; X_2|U_{2c}) &\text{(11f)}\\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_1; U_{2c}, U_{1c}, U_{1pb}) &\text{(11g)}\\
R_{1c} + R_{1pb} + R'_{1c} + R'_{1pb} &\leq I(Y_1; U_{1c}, U_{1pb}|U_{2c}) &\text{(11h)}\\
R_{1pb} + R'_{1pb} &\leq I(Y_1; U_{1pb}|U_{2c}, U_{1c}) &\text{(11i)}
\end{aligned}$$

taken over the union of all distributions that factor as

$$p_{U_{2c},X_2} p_{U_{1c}|X_2} p_{U_{1pb}|X_2} p_{X_1|X_2,U_{1c},U_{1pb}}.$$

Equation-by-equation comparison. We now show that $\mathcal{R}_{DMT}^{out} \subseteq \mathcal{R}_{RTD}^{in}$ by fixing an input distribution (which are the same for these two regions) and comparing the rate regions equation by equation. We refer to the equation numbers directly, and look at the difference between the corresponding equations in the two new regions.

- (11c)-(11a) vs (9c)-(9a): Noting the cancelation / interplay between the binning rates, we see that

$$((\text{11c}) - (\text{11a})) - ((\text{9d}) - (\text{9a})) = 0.$$

- (11d)-(11a) vs. (9d)-(9a):

$$\begin{aligned}
&((\text{11d}) - (\text{11a})) - ((\text{9d}) - (\text{9a}))\\
&\quad = -I(X_2; U_{1c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = I(U_{2c}; U_{1c}|X_2)\\
&\quad = 0
\end{aligned}$$

- (11e)-(11a) vs. (9e)-(9a): again noting the cancelations,

$$((\text{11e}) - (\text{11a})) - ((\text{9e}) - (\text{9a})) = 0$$

- (11f) vs. (9f):

$$(\text{11f}) - (\text{9f}) = 0$$

- (11g)-(11b) vs. (9g)-(9b)-(9a)

$$\begin{aligned}
&((\text{11g}) - (\text{11b})) - ((\text{9g}) - (\text{9b}) - (\text{9a}))\\
&\quad = -I(X_2; U_{1c}, U_{1pb}|U_{2c})\\
&\qquad -I(U_{1pb}, U_{1c}; U_{2c}) + I(U_{1c}; U_{2c}, X_2) + I(U_{1pb}; U_{2c}, X_2)\\
&\quad = -I(U_{1pb}, U_{1c}; X_2, U_{2c}) + I(U_{1c}; U_{2c}, X_2) + I(U_{1pb}; U_{2c}, X_2)\\
&\quad = -I(U_{1pb}; X_2, U_{2c}) - I(U_{1c}; X_2, U_{2c}|U_{1pb}) + I(U_{1c}; U_{2c}, X_2) + I(U_{1pb}; U_{2c}, X_2)\\
&\quad = -I(U_{1c}; X_2, U_{2c}|U_{1pb}) + I(U_{1c}; U_{2c}, X_2)\\
&\quad = -H(U_{1c}|U_{1pb}) + H(U_{1c}|X_2, U_{2c}, U_{1pb}) + H(U_{1c}) - H(U_{1c}|X_2, U_{2c})\\
&\quad = I(U_{1c}; U_{1pb}) > 0
\end{aligned}$$

  where we have used the fact that U_{1c} and U_{1pb} are conditionally independent given (U_{2c}, X_2).
- (11h) – (11b) vs. (9h) – (9b) – (9a):

$$\begin{aligned}
&((\text{11h}) - (\text{11b})) - ((\text{9h}) - (\text{9b}) - (\text{9a}))\\
&\quad = -I(X_2; U_{1c}, U_{1pb}|U_{2c}) - I(U_{2c}; U_{1c}, U_{1pb}) + I(U_{1pb}; U_{2c}, X_2) - I(U_{1pb}; U_{1c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = -I(X_2, U_{2c}; U_{1c}, U_{1pb}) + I(U_{1pb}; U_{2c}, X_2) - I(U_{1pb}; U_{1c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = -I(X_2, U_{2c}; U_{1pb}) - I(U_{1c}; X_2, U_{2c}|U_{1pb}) + I(U_{1pb}; U_{2c}, X_2) - I(U_{1pb}; U_{1c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = -I(U_{1c}; X_2, U_{2c}, U_{1pb}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = -I(U_{1c}; X_2, U_{2c}) - I(U_{1c}; U_{1pb}|X_2, U_{2c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = 0
\end{aligned}$$

where we have used the fact that U_{1c} and U_{1pb} are conditionally independent given (U_{2c}, X_2).

- (11i) $-$ (11b) $+$ (11a) vs. (9i) $-$ (9b):

$$\begin{aligned}
&((11\mathrm{i}) - (11\mathrm{b}) + (11\mathrm{a})) - ((9\mathrm{i}) - (9\mathrm{b}))\\
&\quad = -I(U_{1pb}; X_2|U_{2c}, U_{1c}) - I(U_{1pb}; U_{2c}, U_{1c}) + I(U_{1pb}; X_2, U_{2c})\\
&\quad = -I(U_{1pb}; X_2, U_{2c}, U_{1c}) + I(U_{1pb}; U_{2c}, X_2)\\
&\quad = -I(U_{1pb}; U_{1c}|U_{2c}, X_2)\\
&\quad = 0
\end{aligned}$$

C. Containment of [4, Thm. 2] in $\mathcal{R}_{RTD}$

The independently derived region in [3, Thm. 2] uses a similar encoding structure as that of $\mathcal{R}_{RTD}$ with two exceptions: a) the binning is done sequentially rather than jointly as in $\mathcal{R}_{RTD}$ leading to binning constraints (43)–(45) in [3, Thm. 2] as opposed to (3a)–(3c) in Thm.1. Notable is that both schemes have adopted a Marton-like binning scheme at the cognitive transmitter, as first introduced in the context of the CIFC in [3]. b) While the cognitive messages are rate-split in identical fashions, the primary message is split into 2 parts in [3, Thm. 2] ($R_1 = R_{11} + R_{10}$, note the reversal of indices) while we explicitly split the primary message into three parts $R_2 = R_{2c} + R_{2pa} + R_{2pb}$. We show that the region of [3, Thm.2], denoted as $\mathcal{R}_{CC} \subseteq \mathcal{R}_{RTD}$ in two steps:

- We first show that we may WLOG set $U_{11} = \emptyset$ in [3, Thm.2], creating a new region R'_{CC}.
- We next make a correspondence between our RV's and those of [3, Thm.2] and obtain identical regions.

We note that the primary and cognitive indices are permuted in [3].

We first show that U_{11} in [3, Thm. 2] may be dropped WLOG. Consider the region $\mathcal{R}_{CC}$ of [3, Thm. 2], defined as the union over all distributions $p_{U_{10},U_{11},V_{11},V_{20},V_{22},X_1,X_2} p_{Y_1,Y_2|X_1,X_2}$ of all rate tuples satisfying:

$$R_1 \le I(Y_1; V_{11}, U_{11}, V_{20}, U_{10}) \tag{12}$$
$$R_2 \le I(Y_2; V_{20}, V_{22}|U_{10}) - I(V_{22}, V_{20}; U_{11}|U_{10}) \tag{13}$$
$$R_1 + R_2 \le I(Y_1; V_{11}, U_{11}|V_{20}, U_{10}) + I(Y_2; V_{22}, V_{20}, U_{10}) - I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) \tag{14}$$
$$R_1 + R_2 \le I(Y_1; V_{11}, U_{11}, V_{20}, U_{10}) + I(Y_2; V_{22}|V_{20}, U_{10}) - I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) \tag{15}$$
$$\begin{aligned}2R_2 + R_1 \le\ & I(Y_1; V_{11}, U_{11}, V_{20}|U_{10}) + I(Y_2; V_{22}|V_{20}, U_{10}) + I(Y_2; V_{20}, V_{22}, U_{10})\\ & - I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) - I(V_{22}, V_{20}; U_{11}|U_{10})\end{aligned} \tag{16}$$

Now let $\mathcal{R}'_{CC}$ be the region obtained by setting $U'_{11} = \emptyset$ and $V'_{11} = (V_{11}, U_{11})$ while keeping all remaining RV's identical. Then $\mathcal{R}'_{CC}$ is the union over all distributions $p_{U_{10},V'_{11},V_{20},V_{22},X_1,X_2} p_{Y_1,Y_2|X_1,X_2}$, with $V'_{11} = (V_{11}, U_{11})$ in $\mathcal{R}_{CC}$, of all rate tuples satisfying:

$$R_1 \le I(Y_1; V_{11}, U_{11}, V_{20}, U_{10}) \tag{17}$$
$$R_2 \le I(Y_2; V_{20}, V_{22}|U_{10}) \tag{18}$$
$$R_1 + R_2 \le I(Y_1; V_{11}, U_{11}|V_{20}, U_{10}) + I(Y_2; V_{22}, V_{20}, U_{10}) - I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) \tag{19}$$
$$R_1 + R_2 \le I(Y_1; V_{11}, U_{11}, V_{20}, U_{10}) + I(Y_2; V_{22}|V_{20}, U_{10}) - I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) \tag{20}$$
$$\begin{aligned}2R_2 + R_1 \le\ & I(Y_1; V_{11}, U_{11}, V_{20}|U_{10}) + I(Y_2; V_{22}|V_{20}, U_{10}) + I(Y_2; V_{20}, V_{22}, U_{10})\\ & - I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10})\end{aligned} \tag{21}$$

Comparing the two regions equation by equation, we see that

- (12)= (17)
- (13) $<$ (18) as this choice of RV's sets the generally positive mutual information to 0
- (14)=(19)
- (15)=(20)
- (16) $<$ (21) as this choice of RV's sets the generally positive mutual information to 0

From the previous, we may set $U_{11} = \emptyset$ in the region $\mathcal{R}_{CC}$ of [3, Thm. 2] without loss of generality, obtaining the region $\mathcal{R}'_{CC}$ defined in (17) – (21). We show that $\mathcal{R}'_{CC}$ may be obtained from the region $\mathcal{R}_{RTD}$ with the assigment of RV's, rates and binning rates in Table II.

RV, rate of Theorem 1	RV, rate of [9, Thm. 1]	Comments
U_{2c}, R_{2c}	U_{10}, R_{10}	TX 2 $\to$ RX 1, RX 2
$X_2 = U_{2c}$, $R_{2pa} = 0$	$U_{11} = \emptyset$, $R_{11} = 0$	TX 2 $\to$ RX 2
U_{1c}, R_{1c}	V_{20}, R_{20}	TX 1 $\to$ RX 1, RX 2
U_{1pb}, R_{1pb}	V_{22}, R_{22}	TX 1 $\to$ RX 1
U_{2pb}, R_{2pb}	V_{11}	TX 1 $\to$ RX 2
R'_{1c}	$L_{20} - R_{20}$	
R'_{1pb}	$L_{22} - R_{22}$	
R'_{2pb}	$L_{11} - R_{11}$	
X_1	X_2	
X_2	X_1	

TABLE II
ASSIGNMENT OF RV'S OF SECTION C

Evaluating $\mathcal{R}'_{CC}$ defined by (17) – (21) with the above assignment, translating all RV's into the notation used here, we obtain the region:

$$\begin{array}{lcl}
R'_{1c} & \geq & 0 \\
R'_{1pb} + R'_{2pb} & \geq & I(U_{1pb}; U_{2pb}|U_{2c}, U_{1c}) \\
R_{2pb} + R'_{2pb} & \leq & I(Y_2; U_{2pb}|U_{2c}, U_{1c}) \\
R_{2pb} + R'_{2pb} + R_{1c} + R'_{1c} & \leq & I(Y_2; U_{1c}, U_{2pb}|U_{2c}) \\
R_{2pb} + R'_{2pb} + R_{1c} + R'_{1c} + R_{2c} & \leq & I(Y_2; U_{1c}, U_{2c}, U_{2pb}) \\
R_{1pb} + R'_{1pb} & \leq & I(Y_1; U_{1pb}|U_{2c}, U_{1c}) \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} & \leq & I(Y_1; U_{1pb}, U_{1c}|U_{2c}) \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} + R_{2c} & \leq & I(Y_1; U_{1pb}, U_{1c}, U_{2c})
\end{array}$$

Note that we may take binning rate equations $R'_{1c} \geq 0$ and $R'_{1pb} + R'_{2pb} \geq I(U_{1pb}; U_{2pb}|U_{2c}, U_{1c})$ to be equality without loss of generality - the largest region will take $R'_{1c}, R'_{1pb}, R'_{2pb}$ as small as possible. The region $\mathcal{R}_{RTD}$ with $R_{2pa} = 0$

$$\begin{array}{lcl}
R'_{1c} & \geq & 0 \\
R'_{1c} + R'_{1pb} & \geq & 0 \\
R'_{1c} + R'_{1pb} + R'_{2pb} & \geq & I(U_{1pb}; U_{2pb}|U_{2c}, U_{1c}) \\
R_{2pb} + R'_{2pb} & \leq & I(Y_2; U_{2pb}|U_{2c}, U_{1c}) \\
R_{2pb} + R'_{2pb} + R_{1c} + R'_{1c} & \leq & I(Y_2; U_{1c}, U_{2pb}|U_{2c}) \\
R_{2pb} + R'_{2pb} + R_{1c} + R'_{1c} + R_{2c} & \leq & I(Y_2; U_{1c}, U_{2c}, U_{2pb}) \\
R_{1pb} + R'_{1pb} & \leq & I(Y_1; U_{1pb}|U_{2c}, U_{1c}) \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} & \leq & I(Y_1; U_{1pb}, U_{1c}|U_{2c}) \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} + R_{2c} & \leq & I(Y_1; U_{1pb}, U_{1c}, U_{2c})
\end{array}$$

For $R'_{1c} = 0$ these two regions are identical, showing that $\mathcal{R}_{RTD}$ is surely no smaller than $\mathcal{R}_{CC}$. For $R'_{1c} > 0$, $\mathcal{R}_{RTD}$, the binning rates of the region $\mathcal{R}_{RTD}$ are looser than the ones in $\mathcal{R}_{CC}$. This is probably due to the fact that the first one uses joint binning and latter one sequential binning. Therefore $\mathcal{R}_{RTD}$ may produce rates larger than $\mathcal{R}_{CC}$. However, in general, no strict inclusion of $\mathcal{R}_{CC}$ in $\mathcal{R}_{RTD}$ has been shown.

D. Containment of [14, Thm. 4.1] in $\mathcal{R}_{RTD}$:

In this scheme the common messages are created independently instead of having the common message from transmitter 1 being superposed to the common message from transmitter 2. The former choice introduces more rate constraints than the latter and allows us to show inclusion in $\mathcal{R}_{RTD}$.

The region of [14] is expressed as the set of all rate tuples satisfying

$$\begin{aligned}
R'_{22} &\geq I(W_2;V_1|U_1,U_2) && \text{(22a)}\\
R'_{11}+R'_{22} &\geq I(W_2;W_1,V_1|U_1,U_2) && \text{(22b)}\\
R_{11}+R'_{11} &\leq I(V_1,W_1;Y_1|U_1,U_2) && \text{(22c)}\\
R_{12}+R_{11}+R'_{11} &\leq I(U_1,V_1,W_1;Y_1|U_2) && \text{(22d)}\\
R_{21}+R_{11}+R'_{11} &\leq I(U_2,V_1,W_1;Y_1|U_1) && \text{(22e)}\\
R_{12}+R_{21}+R_{11}+R'_{11} &\leq I(U_1,V_1,W_1,U_2;Y_1) && \text{(22f)}\\
R_{22}+R'_{22} &\leq I(W_2;Y_2|U_1,U_2) && \text{(22g)}\\
R_{21}+R_{22}+R'_{22} &\leq I(U_2,W_2;Y_2|U_1) && \text{(22h)}\\
R_{12}+R_{22}+R'_{22} &\leq I(U_1,W_2;Y_2|U_2) && \text{(22i)}\\
R_{12}+R_{21}+R_{22}+R'_{22} &\leq I(U_1,U_2,W_2;Y_2) && \text{(22j)}
\end{aligned}$$

taken over the union over of distributions

$$p_{u_1}p_{v_1|u_1}p_{x_1|v_1,u_1}p_{u_2}p_{w_1,w_2|v_1,u_1,u_2}p_{x_0|w_1,w_2,v_1,u_1,u_2}p_{y_1,y_2|x_1,x_0}$$

for $(R'_{11},R'_{22},R_{11},R_{12},R_{21},R_{22})\in\mathbb{R}_+^6$.

Following the argument of [29, Appendix D] we can show that WLG we can take X_1 and X_2 to be deterministic functions, so that we can write

$$\begin{aligned}
R'_{22} &\geq I(W_2;V_1,X_1|U_1,U_2) && \text{(23a)}\\
R'_{11}+R'_{22} &\geq I(W_2;W_1,V_1,X_1|U_1,U_2) && \text{(23b)}\\
R_{11}+R'_{11} &\leq I(V_1,X_1,W_1;Y_1|U_1,U_2) && \text{(23c)}\\
R_{12}+R_{11}+R'_{11} &\leq I(U_1,V_1,X_1,W_1;Y_1|U_2) && \text{(23d)}\\
R_{21}+R_{11}+R'_{11} &\leq I(U_2,V_1,X_1,W_1;Y_1|U_1) && \text{(23e)}\\
R_{12}+R_{21}+R_{11}+R'_{11} &\leq I(U_1,V_1,X_1W_1,U_2;Y_1) && \text{(23f)}\\
R_{22}+R'_{22} &\leq I(W_2;Y_2|U_1,U_2) && \text{(23g)}\\
R_{21}+R_{22}+R'_{22} &\leq I(U_2,W_2;Y_2|U_1) && \text{(23h)}\\
R_{12}+R_{22}+R'_{22} &\leq I(U_1,W_2;Y_2|U_2) && \text{(23i)}\\
R_{12}+R_{21}+R_{22}+R'_{22} &\leq I(U_1,U_2,W_2;Y_2). && \text{(23j)}
\end{aligned}$$

We can now eliminate one random variable by noticing that

$$p_{u_1}p_{v_1|u_1}p_{x_1|v_1,u_1}p_{u_2}p_{w_1,w_2|v_1,u_1,u_2}p_{x_0|w_1,w_2,v_1,u_1,u_2}p_{y_1,y_2|x_1,x_0}=p_{u_1}p_{v_1,x_1|u_1}p_{u_2}p_{w_1,w_2|v_1,u_1,x_1,u_2}p_{x_0|w_1,w_2,v_1,u_1,x_1,u_2}p_{y_1,y_2|x_1,x_0}$$

, and setting $V'_1=V_1,X_1$, to obtain the region

$$\begin{aligned}
R'_{22} &\geq I(W_2;V'_1|U_1,U_2) && \text{(24a)}\\
R'_{11}+R'_{22} &\geq I(W_2;W_1,V'_1|U_1,U_2) && \text{(24b)}\\
R_{11}+R'_{11} &\leq I(V'_1,W_1;Y_1|U_1,U_2) && \text{(24c)}\\
R_{12}+R_{11}+R'_{11} &\leq I(U_1,V'_1,W_1;Y_1|U_2) && \text{(24d)}\\
R_{21}+R_{11}+R'_{11} &\leq I(U_2,V'_1,W_1;Y_1|U_1) && \text{(24e)}\\
R_{12}+R_{21}+R_{11}+R'_{11} &\leq I(U_1,V'_1W_1,U_2;Y_1) && \text{(24f)}\\
R_{22}+R'_{22} &\leq I(W_2;Y_2|U_1,U_2) && \text{(24g)}\\
R_{21}+R_{22}+R'_{22} &\leq I(U_2,W_2;Y_2|U_1) && \text{(24h)}\\
R_{12}+R_{22}+R'_{22} &\leq I(U_1,W_2;Y_2|U_2) && \text{(24i)}\\
R_{12}+R_{21}+R_{22}+R'_{22} &\leq I(U_1,U_2,W_2;Y_2) && \text{(24j)}
\end{aligned}$$

taken over the union of all distributions of the form

$$p_{u_1}p_{v'_1|u_1}p_{u_2}p_{w_1,w_2|v'_1,u_1,u_2}p_{x_0|w_1,w_2,v'_1,u_1,u_2}p_{y_1,y_2|v'_1,x_0}$$

RV, rate of Theorem 1	RV, rate of [9, Thm. 1]	Comments
U_{2c}, R_{2c}	U_1, R_{12}	TX 2 → RX 1, RX 2
X_2, R_{2pa}	V_1', R_{11}'	TX 2 → RX 2
U_{1c}, R_{1c}	U_2, R_{21}	TX 1 → RX 1, RX 2
U_{1pb}, R_{1pb}	W_2, R_{22}	TX 1 → RX 1
$U_{2pb}, R_{2pb} = 0$	W_1	TX 1 → RX 2
R_{1c}'	$L_{20} - R_{20}$	
R_{1pb}'	$L_{11} - R_{11}$	
R_{2pb}'	$L_{22} - R_{22}$	
X_1	X_0	
X_2	X_1	

TABLE III
ASSIGNMENT OF RV'S OF SECTION D

We equate the RV's in the region of [14] with the RV's in Theorem 1 as in Table III.
With the substitution in the achievable rate region of (24), we obtain the region

$$
\begin{aligned}
R_{1pb}' &\geq I(U_{1pb}; X_2|U_{2c}, U_{1c}) && \text{(25a)}\\
R_{1pb}' + R_{2pb}' &\geq I(U_{1pb}; U_{2pb}, X_2|U_{2c}, U_{1c}) && \text{(25b)}\\
R_{2pa} + R_{2pb}' &\leq I(X_2, U_{2pb}; Y_2|U_{2c}, U_{1c}) && \text{(25c)}\\
R_{2c} + R_{2pa} + R_{2pb}' &\leq I(U_{2c}, X_2, U_{2pb}; Y_2|U_{1c}) && \text{(25d)}\\
R_{1c} + R_{2pa} + R_{2pb}' &\leq I(U_{1c}, X_2, U_{2pb}; Y_2|U_{2c}) && \text{(25e)}\\
R_{2c} + R_{1c} + R_{2pa} + R_{2pb}' &\leq I(U_{2c}, X_2, U_{1c}, U_{1pb}; Y_2) && \text{(25f)}\\
R_{1pb} + R_{1pb}' &\leq I(U_{1pb}; Y_1|U_{2c}, U_{1c}) && \text{(25g)}\\
R_{1c} + R_{1pb} + R_{1pb}' &\leq I(U_{1c}, U_{1pb}; Y_1|U_{2c}) && \text{(25h)}\\
R_{2c} + R_{1pb} + R_{1pb}' &\leq I(U_{2c}, U_{1pb}; Y_1|U_{1c}) && \text{(25i)}\\
R_{2c} + R_{1c} + R_{1pb} + R_{1pb}' &\leq I(U_{2c}, U_{1c}, U_{1pb}; Y_1) && \text{(25j)}
\end{aligned}
$$

taken over the union of all distributions of the form

$$p_{U_{1c}} p_{U_{2c}} p_{X_2|U_{2c}} p_{U_{1pb}, U_{2pb}|U_{1c}, U_{2c}, X_2} p_{X_1|U_{2c}, U_{1c}, U_{1pb}, U_{2pb}}.$$

Set $R_{2pb} = 0$ and $R_{1c}' = I(U_{1c}; X_2|U_{2c})$ in the achievable scheme of Theorem 1 and consider the factorization of the remaining RV's as in the scheme of (25), that is, according to

$$p_{U_{1c}} p_{U_{2c}} p_{X_2|U_{2c}} p_{U_{1pb}|U_{1c}, U_{2c}, X_2} p_{U_{2pb}|U_{1c}, U_{2c}, X_2} p_{X_1|U_{2c}, X_2, U_{1c}, U_{1pb}, U_{2pb}}.$$

With this factorization of the distributions, we obtain the achievable region

$$
\begin{aligned}
R_{1c}' &= I(U_{1c}; X_2|U_{2c}) && \text{(26a)}\\
R_{1pb}' &\geq I(U_{1pb}; X_2|U_{2c}, U_{1c}) && \text{(26b)}\\
R_{1pb}' + R_{2pb}' &\geq I(U_{1pb}; X_2, U_{2pb}|U_{2c}, U_{1c}) && \text{(26c)}\\
R_{2pa} + R_{2pb}' &\leq I(Y_2; X_2, U_{2pb}|U_{2c}, U_{1c}) + I(U_{1c}; X_2|U_{2c}) && \text{(26d)}\\
R_{1c} + R_{2pa} + R_{2pb}' &\leq I(Y_2; U_{1c}, X_2, U_{2pb}|U_{2c}) && \text{(26e)}\\
R_{2c} + R_{1c} + R_{2pa} + R_{2pb}' &\leq I(Y_2; U_{2pb}, U_{1c}, U_{2c}, X_2) && \text{(26f)}\\
R_{1pb} + R_{1pb}' &\leq I(Y_1; U_{1pb}|U_{2c}, U_{1c}) && \text{(26g)}\\
R_{1c} + R_{1pb} + R_{1pb}' &\leq I(Y_1; U_{1c}, U_{1pb}|U_{2c}) && \text{(26h)}\\
R_{2c} + R_{1c} + R_{1pb} + R_{1pb}' &\leq I(Y_1; U_{2c}, U_{1c}, U_{1pb}) && \text{(26i)}
\end{aligned}
$$

Note that with this particular factorization we have that $I(U_{1c}; X_2|U_{2c}) = 0$, since X_2 is conditionally independent on U_{1c} given U_{2c}.

We now compare the region of (25) and (26) for a fixed input distribution, equation by equation:

$$\begin{aligned}
&(26b) = (25a)\\
&(26c) = (25b)\\
&(26d) = (25c)\\
&(26e) = (25e)\\
&(26f) = (25f)\\
&(26g) = (25g)\\
&(26h) = (25h)\\
&(26i) = (25j)
\end{aligned}$$

clearly (25d) and (25i) are extra bounds that further restrict the region in [14] to be smaller than the region of Theorem 1.