Mismatched Multi-letter Successive Decoding for the Multiple-Access Channel

Jonathan Scarlett, Alfonso Martinez and Albert Guillén i Fàbregas

***Abstract*—This paper studies channel coding for the discrete memoryless multiple-access channel with a given (possibly suboptimal) decoding rule. A multi-letter successive decoding rule depending on an arbitrary non-negative decoding metric is considered, and achievable rate regions and error exponents are derived both for the standard MAC (independent codebooks), and for the cognitive MAC (one user knows both messages) with superposition coding. In the cognitive case, the rate region and error exponent are shown to be tight with respect to the ensemble average. The rate regions are compared with those of the commonly-considered decoder that chooses the message pair maximizing the decoding metric, and numerical examples are given for which successive decoding yields a strictly higher sum rate for a given pair of input distributions.**

I. Introduction

The mismatched decoding problem [1]–[3] seeks to characterize the performance of channel coding when the decoding rule is fixed and possibly suboptimal (e.g., due to channel uncertainty or implementation constraints). Extensions of this problem to multiuser settings are not only of interest in their own right, but can also provide valuable insight into the single-user setting [3]–[5]. In particular, significant attention has been paid to the mismatched multiple-access channel (MAC), described as follows. User $\nu = 1, 2$ transmits a codeword $\boldsymbol{x}_\nu$ from a codebook $\mathcal{C}_\nu = \{\boldsymbol{x}_\nu^{(1)}, \cdots, \boldsymbol{x}_\nu^{(M_\nu)}\}$, and the output sequence $\boldsymbol{y}$ is generated according to $W^n(\boldsymbol{y}|\boldsymbol{x}_1, \boldsymbol{x}_2) \triangleq \prod_{i=1}^n W(y_i|x_{1,i}, x_{2,i})$ for some transition law $W(y|x_1, x_2)$. The *mismatched* decoder estimates the message pair as

$$(\hat{\mathsf{m}}_1, \hat{\mathsf{m}}_2) = \arg\max_{(i,j)} q^n(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y}), \tag{1}$$

where $q^n(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) \triangleq \prod_{i=1}^n q(x_{1,i}, x_{2,i}, y_i)$ for some non-negative decoding metric $q(x_1, x_2, y)$. The metric $q(x_1, x_2, y) = W(y|x_1, x_2)$ corresponds to optimal maximum-likelihood (ML) decoding, whereas the introduction of mismatch can significantly increase the error probability and lead to smaller achievable rate regions [1], [3]. Even in the single-user case, characterizing the capacity with mismatch is a long-standing open problem.

Given that the decoder only knows the metric $q^n(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y})$ corresponding to each codeword pair, one may question whether there exists a decoding rule that provides better performance than the maximum-metric rule in (1), and that is well-motivated from a practical perspective. The second of these requirements is not redundant; for instance, if the values $\{\log q(x_1, x_2, y)\}$ are rationally independent (i.e., no values can be written as linear combinations of the others with rational coefficients), then one could consider a highly artificial and impractical decoder that uses these values to infer the joint empirical distribution of $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})$, and in turn uses that to implement the maximum-likelihood (ML) rule. While such a decoder is a function of $\{q^n(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y})\}_{i,j}$ and clearly outperforms the maximum-metric rule, it does not bear any practical interest.

There are a variety of well-motivated decoding rules that are of interest beyond maximum-metric, including threshold decoding [6], [7], likelihood decoding [8], [9], and successive decoding [10], [11]. In this paper, we focus on the latter, and consider the following two-step decoding rule:

$$\hat{\mathsf{m}}_1 = \arg\max_i \sum_j q^n(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y}), \tag{2}$$

$$\hat{\mathsf{m}}_2 = \arg\max_j q^n(\boldsymbol{x}_1^{(\hat{\mathsf{m}}_1)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y}). \tag{3}$$

The study of this decoder is of interest for several reasons:

- The decoder depends on the exact same quantities as the maximum-metric decoder (1) (namely, $q^n(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y})$ for each (i, j)), meaning a comparison of the two rules is in a sense fair. We will see the successive rule can sometimes achieve random-coding rates that are not achieved by the maximum-metric rule, which is the first result of this kind for the mismatched MAC.
- The first decoding step (2) can be considered a mismatched version of the optimal decoding rule for (one user of) the interference channel. Hence, as well as giving an achievable rate region for the MAC with mismatched successive decoding, our results directly quantify the loss due to mismatch for the interference channel.
- More broadly, successive decoding is of significant practical interest for multiple-access scenarios, since it permits the use of single-user codes, as well as linear decoding

J. Scarlett is with the Department of Computer Science and Department of Mathematics, National University of Singapore, 117417. (e-mail: scarlett@comp.nus.edu.sg).

A. Martinez is with the Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08018 Barcelona, Spain (e-mail: alfonso.martinez@ieee.org).

A. Guillén i Fàbregas is with the Institució Catalana de Recerca i Estudis Avançats (ICREA), the Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08018 Barcelona, Spain, and also with the Department of Engineering, University of Cambridge, Cambridge, CB2 1PZ, U.K. (e-mail: guillen@ieee.org).

This work has been funded in part by the European Research Council under ERC grant agreement 259663, by the European Union's 7th Framework Programme under grant agreement 303633 and by the Spanish Ministry of Economy and Competitiveness under grants RYC-2011-08150 and TEC2012-38800-C03-03.

This work was presented in part at the 2014 IEEE International Symposium on Information Theory, Honolulu, HI.

complexity in the number of users [11]. While the specific successive decoder that we consider does not enjoy these practical benefits, it may still serve as an interesting point of comparison for such variants.

The rule in (2) is *multi-letter*, in the sense that the objective function does not factorize into a product of n symbols on $(\mathcal{X}_1, \mathcal{Y})$. Single-letter successive decoders [10, Sec. 4.5.1] could also potentially be studied from a mismatched decoding perspective by introducing a second decoding metric $q_2(x_1, y)$, but we focus on the above rule depending only on a *single* metric $q(x_1, x_2, y)$.

Under the above definitions of W, q, W^n and q^n, and assuming the corresponding alphabets $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{Y}$ to be finite, we consider two distinct classes of MACs:

1) For the *standard MAC* [3], encoder $\nu = 1, 2$ takes as input m_ν equiprobable on $\{1, \cdots, M_\nu\}$, and transmits the corresponding codeword $\boldsymbol{x}_\nu^{(\mathsf{m}_\nu)}$ from a codebook $\mathcal{C}_\nu$.
2) For the *cognitive MAC* [4] (or *MAC with degraded message sets* [10, Ex. 5.18]), the messages m_ν are still equiprobable on $\{1, \cdots, M_\nu\}$, but user 2 has access to both messages, while user 1 only knows m_1. Thus, $\mathcal{C}_1$ contains codewords indexed as $\boldsymbol{x}_1^{(i)}$, and $\mathcal{C}_2$ contains codewords indexed as $\boldsymbol{x}_2^{(i,j)}$.

For each of these, we say that a rate pair (R_1, R_2) is achievable if, for all $\delta > 0$, there exist sequences of codebooks $\mathcal{C}_{1,n}$ and $\mathcal{C}_{2,n}$ with $M_1 \geq e^{n(R_1 - \delta)}$ and $M_2 \geq e^{n(R_2 - \delta)}$ respectively, such that the error probability

$$p_{\mathrm{e}} \triangleq \mathbb{P}[(\hat{\mathsf{m}}_1, \hat{\mathsf{m}}_2) \neq (\mathsf{m}_1, \mathsf{m}_2)] \quad (4)$$

tends to zero under the decoding rule described by (2)–(3). Our results will not depend on the method for breaking ties, so for concreteness, we assume that ties are broken as errors.

For fixed rates R_1 and R_2, an error exponent $E(R_1, R_2)$ is said to be achievable if there exists a sequence of codebooks $\mathcal{C}_{1,n}$ and $\mathcal{C}_{2,n}$ with $M_1 \geq \exp(nR_1)$ and $M_2 \geq \exp(nR_2)$ codewords of length n such that

$$\liminf_{n\to\infty} -\frac{1}{n} \log p_{\mathrm{e}} \geq E(R_1, R_2). \quad (5)$$

Letting $\mathcal{E}_\nu \triangleq \{\hat{\mathsf{m}}_\nu \neq \mathsf{m}_\nu\}$ for $\nu = 1, 2$, we observe that if $q(x_1, x_2, y) = W(y|x_1, x_2)$, then (2) is the decision rule that minimizes $\mathbb{P}[\mathcal{E}_1]$. Using this observation, we show in Appendix A that the successive decoder with $q = W$ is guaranteed to achieve the same rate region and error exponent as that of optimal non-successive maximum-likelihood decoding.

A. Previous Work and Contributions

The vast majority of previous works on mismatched decoding have focused on achievability results via random coding, and the only general converse results are written in terms of non-computable information-spectrum type quantities [7]. For the point-to-point setting with mismatch, the asymptotics of random codes with independent codewords are well-understood for the i.i.d. [12], constant-composition [1], [13]–[15] and cost-constrained [2], [16] ensembles. Dual expressions and continuous alphabets were studied in [2].

The mismatched MAC was introduced by Lapidoth [3], who showed that (R_1, R_2) is achievable provided that

$$R_1 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y} \,:\, \widetilde{P}_{X_1} = Q_1, \widetilde{P}_{X_2Y} = P_{X_2Y}, \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} I_{\widetilde{P}}(X_1; X_2, Y), \quad (6)$$

$$R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y} \,:\, \widetilde{P}_{X_2} = Q_2, \widetilde{P}_{X_1Y} = P_{X_1Y}, \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} I_{\widetilde{P}}(X_2; X_1, Y), \quad (7)$$

$$R_1 + R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y} \,:\, \widetilde{P}_{X_1} = Q_1, \widetilde{P}_{X_2} = Q_2, \widetilde{P}_Y = P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)], \\ I_{\widetilde{P}}(X_1;Y) \leq R_1,\, I_{\widetilde{P}}(X_2;Y) \leq R_2}} D(\widetilde{P}_{X_1X_2Y} \| Q_1 \times Q_2 \times \widetilde{P}_Y), \quad (8)$$

where Q_1 and Q_2 are arbitrary input distributions, and $P_{X_1X_2Y} \triangleq Q_1 \times Q_2 \times W$. The corresponding ensemble-tight error exponent was given by the present authors in [5], along with equivalent dual expressions and generalizations to continuous alphabets. Error exponents were also presented for the MAC with general decoding rules in [17], but the results therein are primarily targeted to optimal or universal metrics; in particular, when applied to the mismatched setting, the exponents are not ensemble-tight.

The mismatched cognitive MAC was introduced by Somekh-Baruch [4], who used superposition coding to show that (R_1, R_2) is achievable provided that

$$R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y} \,:\, \widetilde{P}_{X_1X_2} = Q_{X_1X_2}, \widetilde{P}_{X_1Y} = P_{X_1Y}, \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} I_{\widetilde{P}}(X_2; Y | X_1), \quad (9)$$

$$R_1 + R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y} \,:\, \widetilde{P}_{X_1X_2} = Q_{X_1X_2}, \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)], \\ I_{\widetilde{P}}(X_1,;Y) \leq R_1}} I_{\widetilde{P}}(X_1, X_2; Y), \quad (10)$$

where $Q_{X_1X_2}$ is an arbitrary joint input distribution, and $P_{X_1X_2Y} \triangleq Q_{X_1X_2} \times W$. The ensemble-tight error exponent was also given therein. Various forms of superposition coding were also studied by the present authors in [5], but with a focus on the single-user channel rather than the cognitive MAC.

Both of the above regions are known to be tight with respect to the ensemble average for constant-composition random coding, meaning that any looseness is due to the random-coding ensemble itself, rather than the bounding techniques used in the analysis [3], [4]. This notion of tightness was first explored in the single-user setting in [15]. We also note that the above regions lead to improved achievability bounds for the single-user setting [3], [4].

The main contributions of this paper are achievable rate regions for both the standard MAC (Section II-A) and cognitive MAC (Section II-B) under the successive decoding rule in (2)–(3). For the cognitive case, we also provide an ensemble tightness result. Both regions are numerically compared to their counterparts for maximum-metric decoding, and in each case, it is observed that the successive rule can provide a strictly higher sum rate, though neither the successive nor maximum-metric region is included in the other in general.

A by-product of our analysis is achievable error exponents corresponding to the rate regions. Our exponent for the standard MAC is related to that of Etkin *et al.* [18] for the interference channel, as both use parallel coding. Similarly, our exponent for the cognitive MAC is related to that of Kaspi and Merhav [19], since both use superposition coding. Like these works, we make use of type class enumerators; however, a key difference is that we avoid applying a Gallager-type bound in the initial step, and we instead proceed immediately with type-based methods.

In a work that developed independently of ours, the interference channel perspective was pursued in depth in the *matched* case in [20], with a focus on error exponents. The error exponent of [20] is similar to that derived in the present paper, but also contains an extra maximization term that, at least in principle, could improve the exponent. Currently, no examples are known where such an improvement is obtained. Moreover, while the analysis techniques of [20] extend to the mismatched case, doing so leads to the same achievable rate region as ours; the only potential improvement is in the exponent. Finally, we note that while our focus is solely on codebooks with independent codewords, error exponents were also given for the Han-Kobayashi construction in [20].

Another line of related work studied the achievable rates of polar coding with mismatch [21]–[24], using a computationally efficient successive decoding rule. A single-letter achievable rate was given, and it was shown that for a given mismatched transition law (i.e., a conditional probability distribution incorrectly used as if it were the true channel), this decoder can sometimes outperform the maximum-metric decoder. As mentioned above, we make analogous observations in the present paper, albeit for a multiple-access scenario with a very different type of successive decoding.

B. Notation

Bold symbols are used for vectors (e.g., $\boldsymbol{x}$), and the corresponding i-th entry is written using a subscript (e.g., x_i). Subscripts are used to denote the distributions corresponding to expectations and mutual information quantities (e.g., $\mathbb{E}_P[\cdot]$, $I_P(X;Y)$). The marginals of a joint distribution P_{XY} are denoted by P_X and P_Y. We write $P_X = \widetilde{P}_X$ to denote element-wise equality between two probability distributions on the same alphabet. The set of all sequences of length n with a given empirical distribution P_X (i.e., type [25, Ch. 2]) is denoted by $T^n(P_X)$, and similarly for joint types. We write $f(n) \doteq g(n)$ if $\lim_{n\to\infty} \frac{1}{n}\log\frac{f(n)}{g(n)} = 0$, and similarly for $\dot{\le}$ and $\dot{\ge}$. We write $[\alpha]^+ = \max(0,\alpha)$, and denote the indicator function by $\mathbb{1}\{\cdot\}$

II. Main Results

A. Standard MAC

Before presenting our main result for the standard MAC, we state the random-coding distribution that is used in its proof. For $\nu = 1,2$, we fix an input distribution $Q_\nu \in \mathcal{P}(\mathcal{X}_\nu)$, and let $Q_{\nu,n}$ be a type with the same support as Q_ν such that $\max_{x_\nu} |Q_{\nu,n}(x_\nu) - Q_\nu(x_\nu)| \le \frac{1}{n}$. We set

$$P_{\boldsymbol{X}_\nu}(\boldsymbol{x}_\nu) = \frac{1}{|T^n(Q_{\nu,n})|}\mathbb{1}\big\{\boldsymbol{x}_\nu \in T^n(Q_{\nu,n})\big\}, \qquad (11)$$

and consider codewords $\{\boldsymbol{X}_\nu^{(i)}\}_{i=1}^{M_\nu}$ that are independently distributed according to $P_{\boldsymbol{X}_\nu}$. Thus,

$$\Big(\{\boldsymbol{X}_1^{(i)}\}_{i=1}^{M_1}, \{\boldsymbol{X}_2^{(j)}\}_{i=1}^{M_2}\Big) \sim \prod_{i=1}^{M_1} P_{\boldsymbol{X}_1}(\boldsymbol{x}_1^{(i)}) \prod_{j=1}^{M_2} P_{\boldsymbol{X}_2}(\boldsymbol{x}_2^{(j)}). \qquad (12)$$

Our achievable rate region is written in terms of the following functions:

$$\overline{F}(\widetilde{P}_{X_1X_2Y}, \widetilde{P}'_{X_1X_2Y}, R_2) \triangleq \max\Big\{\mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)], \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)] + \big[R_2 - I_{\widetilde{P}'}(X_2;X_1,Y)\big]^+\Big\}, \qquad (13)$$

$$\underline{F}(P_{X_1X_2Y}, R_2) \triangleq \max\Big\{\mathbb{E}_P[\log q(X_1,X_2,Y)], \max_{P'_{X_1X_2Y}\in\mathcal{T}'_1(P_{X_1X_2Y},R_2)} \mathbb{E}_{P'}[\log q(X_1,X_2,Y)] + R_2 - I_{P'}(X_2;X_1,Y)\Big\}, \qquad (14)$$

where

$$\mathcal{T}'_1(P_{X_1X_2Y}, R_2) \triangleq \Big\{P'_{X_1X_2Y} \,:\, P'_{X_1Y} = P_{X_1Y}, P'_{X_2} = P_{X_2}, I_{P'}(X_2;X_1,Y) \le R_2\Big\}. \qquad (15)$$

We will see in our analysis that $P_{X_1X_2Y}$ corresponds to the joint type of the transmitted codewords and the output sequence, and $\widetilde{P}_{X_1X_2Y}$ corresponds to the joint type of some incorrect codeword of user 1, the transmitted codeword of user 2, and the output sequence. Moreover, $P'_{X_1X_2Y}$ and $\widetilde{P}'_{X_1X_2Y}$ similarly correspond to joint types, the difference being that the X_2 marginal is associated with exponentially many sequences in the summation in (2).

Theorem 1. *For any input distributions Q_1 and Q_2, the pair (R_1,R_2) is achievable for the standard MAC with the mismatched successive decoding rule in* (2)–(3) *provided that*

$$R_1 \le \min_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y})\in\mathcal{T}_1(Q_1\times Q_2\times W,R_2)} I_{\widetilde{P}}(X_1;X_2,Y) + \big[I_{\widetilde{P}'}(X_2;X_1,Y) - R_2\big]^+, \qquad (16)$$

$$R_2 \le \min_{\widetilde{P}_{X_1X_2Y}\in\mathcal{T}_2(Q_1\times Q_2\times W)} I_{\widetilde{P}}(X_2;X_1,Y), \qquad (17)$$

where

$$\mathcal{T}_1(P_{X_1X_2Y}, R_2) \triangleq \Big\{(\widetilde{P}_{X_1X_2Y}, \widetilde{P}'_{X_1X_2Y}) \,:\, \widetilde{P}_{X_2Y} = P_{X_2Y}, \widetilde{P}_{X_1} = P_{X_1}, \widetilde{P}'_{X_1Y} = \widetilde{P}_{X_1Y}, \widetilde{P}'_{X_2} = P_{X_2}, \overline{F}(\widetilde{P}_{X_1X_2Y}, \widetilde{P}'_{X_1X_2Y}, R_2) \ge \underline{F}(P_{X_1X_2Y}, R_2)\Big\}, \qquad (18)$$

$$\mathcal{T}_2(P_{X_1X_2Y}) \triangleq \Big\{\widetilde{P}_{X_1X_2Y} \,:\, \widetilde{P}_{X_2} = P_{X_2}, \widetilde{P}_{X_1Y} = P_{X_1Y}, \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \ge \mathbb{E}_P[\log q(X_1,X_2,Y)]\Big\}. \qquad (19)$$

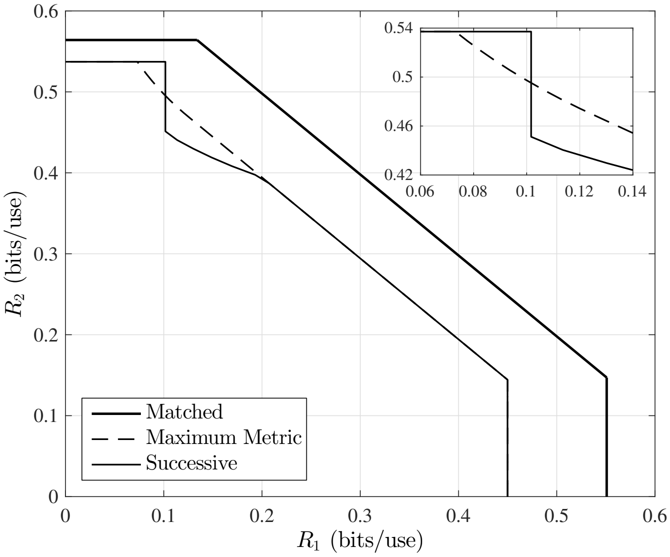

Figure 1. Achievable rate regions for the standard MAC given in (21) with mismatched successive decoding and mismatched maximum-metric decoding.

Proof. See Section III. □

Although the minimization in (16) is a non-convex optimization problem, it can be cast in terms of convex optimization problems, thus facilitating its computation. The details are provided in Appendix B.

While our focus is on achievable rates, the proof of Theorem 1 also provides error exponents. The exponent corresponding to (17) is precisely that corresponding to the error event for user 2 with maximum-metric decoding in [5, Sec. III], and the exponent corresponding to (16) is given by

$$\min_{P_{X_1X_2Y}\,:\,P_{X_1}=Q_1,P_{X_2}=Q_2} D(P_{X_1X_2Y}\|Q_1\times Q_2\times W) + \left[\tilde{I}_1(P_{X_1X_2Y},R_2)-R_1\right]^+, \quad (20)$$

where $\tilde{I}_1(P_{X_1X_2Y},R_2)$ denotes the right-hand side of (16) with an arbitrary distribution $P_{X_1X_2Y}$ in place of $Q_1\times Q_2\times W$. As discussed in Section I-A, this exponent is closely related to a parallel work on the error exponent of the interference channel [20].

Numerical Example: We consider the MAC with $\mathcal{X}_1=\mathcal{X}_2=\{0,1\}$, $\mathcal{Y}=\{0,1,2\}$, and

$$W(y|x_1,x_2)=\begin{cases}1-2\delta_{x_1x_2} & y=x_1+x_2\\ \delta_{x_1x_2} & \text{otherwise,}\end{cases} \quad (21)$$

where $\{\delta_{x_1x_2}\}$ are constants. The mismatched decoder uses $q(x_1,x_2,y)$ of a similar form, but with a fixed value $\delta\in\left(0,\frac{1}{3}\right)$ in place of $\{\delta_{x_1x_2}\}$. All such choices of δ are equivalent for maximum-metric decoding, but not for successive decoding.

We set $\delta_{00}=0.01$, $\delta_{01}=0.1$, $\delta_{10}=0.01$, $\delta_{11}=0.3$, $\delta=0.15$, and $Q_1=Q_2=(0.5,0.5)$. Figure 1 plots the achievable rates regions of successive decoding (Theorem 1), maximum-metric decoding ((6)–(8)), and matched decoding (giving the same region for successive and maximum-metric).

Interestingly, neither of the mismatched rate regions is included in the other, thus suggesting that the two decoding rules are fundamentally different. For the given input distribution, the sum rate for successive decoding exceeds that of maximum-metric decoding. Furthermore, upon taking the convex hull (which is justified by a time sharing argument), the region for successive decoding is strictly larger. While we observed similar behaviors for other choices of Q_1 and Q_2, it remains unclear as to whether this is always the case. Furthermore, while the rate region for maximum-metric decoding is ensemble-tight, it is unclear whether the same is true of the region given in Theorem 1.

To gain insight into the shape of the achievable rate region for successive decoding, it is instructive to consider the various parts of the region. When doing so, the reader may wish to note that the condition in (16) can equivalently be expressed as three related conditions holding simultaneously; see Appendix B, leading to the conditions (131), (133), and (134). We have the following:

- The horizontal line at $R_2\approx 0.54$ corresponds to the requirement on R_2 in (17), which is identical to the condition in (7) for maximum-metric decoding.
- The vertical line at $R_1\approx 0.45$ also coincides with a condition for maximum-metric decoding, namely, (6). It is unsurprising that the two rate regions coincide at $R_2=0$, since if user 2 only has one message then the two decoding rules are identical. For small but positive R_2, the rate region boundaries still coincide even though the decoding rules differ, and the successive decoding curve is dominated by condition (134) in Appendix B.
- The straight diagonal part of the achievable rate region also matches that of maximum-metric decoding. In this case, the successive decoding curve is dominated by condition (133) in Appendix B; the term $\max\{0,I_{\widetilde{P}'}(X_2;X_1,Y)-R_2\}$ expressed by the $[\cdot]^+$ function is dominated by $I_{\widetilde{P}'}(X_2;X_1,Y)-R_2$, and the overall condition becomes a sum-rate bound, i.e., an upper bound on R_1+R_2.
- In the remaining part of the curve, as R_1 decreases, the rate region boundary bends downwards, and then becomes vertical. In this part, the successive decoding curve is dominated by (131) in Appendix B, with R_2 being large enough for the term $\max\{0,I_{\widetilde{P}'}(X_2;X_1,Y)-R_2\}$ to equal zero. The step-like behavior at $R_1\approx 0.1$ corresponds to a change in the dominant term of $\underline{F}$ (see (14)); in the non-vertical part, the dominant term is $\mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)]$, whereas in the vertical part, R_2 is large enough for the other term to dominate.

It is worth noting that under optimal decoding for the interference channel (taking the form (2)), it is known that for R_1 below a certain threshold, R_2 can be arbitrarily large while still ensuring that user 1's message is estimated correctly [26]. This is in analogy with the step-like behavior in Figure 1.

Finally, we note that the mismatched maximum-metric decoding region also has a non-pentagonal and non-convex shape (see the zoomed part of Figure 1), though its deviation from the usual pentagonal shape is milder than the successive decoder in this example.

B. Cognitive MAC

In this section, we consider the analog of Theorem 1 for the cognitive MAC. Besides being of interest in its own right, this will provide a case where ensemble-tightness can be established, and with the numerical results still exhibiting similar phenomena to those shown in Figure 1.

We again begin by introducing the random coding ensemble. We fix a joint distribution $Q_{X_1X_2} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2)$, let $Q_{X_1X_2,n}$ be the corresponding closest joint type in the same way as the previous subsection, and write the resulting marginals as Q_{X_1}, $Q_{X_1,n}$, $Q_{X_2|X_1}$, $Q_{X_2|X_1,n}$, and so on. We consider superposition coding, treating user 1's messages as the "cloud centers", and user 2's messages as the "satellite codewords". More precisely, defining

$$P_{\boldsymbol{X}_1}(\boldsymbol{x}_1) = \frac{1}{|T^n(Q_{X_1,n})|}\mathbb{1}\{\boldsymbol{x}_1 \in T^n(Q_{X_1,n})\}, \qquad (22)$$

$$P_{\boldsymbol{X}_2|\boldsymbol{X}_1}(\boldsymbol{x}_2|\boldsymbol{x}_1) = \frac{1}{|T^n_{\boldsymbol{x}_1}(Q_{X_2|X_1,n})|}\mathbb{1}\{\boldsymbol{x}_2 \in T^n_{\boldsymbol{x}_1}(Q_{X_2|X_1,n})\}, \qquad (23)$$

the codewords are distributed as follows:

$$\left\{\left(\boldsymbol{X}_1^{(i)}, \{\boldsymbol{X}_2^{(i,j)}\}_{j=1}^{M_2}\right)\right\}_{i=1}^{M_1} \sim \prod_{i=1}^{M_1}\left(P_{\boldsymbol{X}_1}(\boldsymbol{x}_1^{(i)})\prod_{j=1}^{M_2}P_{\boldsymbol{X}_2|\boldsymbol{X}_1}(\boldsymbol{x}_2^{(i,j)}|\boldsymbol{x}_1^{(i)})\right). \qquad (24)$$

For the remaining definitions, we use similar notation to the standard MAC, with an additional subscript to avoid confusion. The analogous quantities to (13)–(15) are

$$\overline{F}_{\mathrm{c}}(\widetilde{P}'_{X_1X_2Y}, R_2) \triangleq \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)] + \left[R_2 - I_{\widetilde{P}'}(X_2;Y|X_1)\right]^+, \qquad (25)$$

$$\underline{F}_{\mathrm{c}}(P_{X_1X_2Y}, R_2) \triangleq \max\Bigg\{\mathbb{E}_P[\log q(X_1,X_2,Y)], \max_{P'_{X_1X_2Y} \in \mathcal{T}'_{1\mathrm{c}}(P_{X_1X_2Y},R_2)} \mathbb{E}_{P'}[\log q(X_1,X_2,Y)] + R_2 - I_{P'}(X_2;Y|X_1)\Bigg\}, \qquad (26)$$

where

$$\mathcal{T}'_{1\mathrm{c}}(P_{X_1X_2Y}, R_2) \triangleq \Big\{P'_{X_1X_2Y} \,:\, P'_{X_1Y} = P_{X_1Y}, P'_{X_1X_2} = P_{X_1X_2}, I_{P'}(X_2;Y|X_1) \le R_2\Big\}. \qquad (27)$$

Our main result for the cognitive MAC is as follows.

Theorem 2. *For any input distribution $Q_{X_1X_2}$, the pair (R_1, R_2) is achievable for the cognitive MAC with the mismatched successive decoding rule in* (2)–(3) *provided that*

$$R_1 \le \min_{\widetilde{P}'_{X_1X_2Y} \in \mathcal{T}_{1\mathrm{c}}(Q_{X_1X_2}\times W, R_2)} I_{\widetilde{P}'}(X_1;Y) + \left[I_{\widetilde{P}'}(X_2;Y|X_1) - R_2\right]^+, \qquad (28)$$

$$R_2 \le \min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{2\mathrm{c}}(Q_{X_1X_2}\times W)} I_{\widetilde{P}}(X_2;Y|X_1), \qquad (29)$$

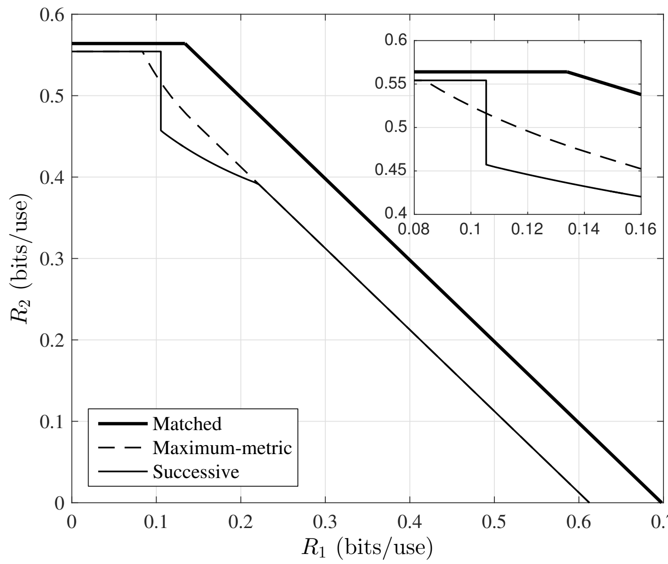

Figure 2. Achievable rate regions for the cognitive MAC given in (21) with mismatched successive decoding and mismatched maximum-metric decoding.

where

$$\mathcal{T}_{1\mathrm{c}}(P_{X_1X_2Y}, R_2) \triangleq \Big\{\widetilde{P}'_{X_1X_2Y} \,:\, P'_{X_1X_2} = P_{X_1X_2}, \widetilde{P}'_Y = P_Y, \overline{F}_{\mathrm{c}}(\widetilde{P}'_{X_1X_2Y}, R_2) \ge \underline{F}_{\mathrm{c}}(P_{X_1X_2Y}, R_2)\Big\}, \qquad (30)$$

$$\mathcal{T}_{2\mathrm{c}}(P_{X_1X_2Y}) \triangleq \Big\{\widetilde{P}_{X_1X_2Y} \,:\, \widetilde{P}_{X_1X_2} = P_{X_1X_2}, \widetilde{P}_{X_1Y} = P_{X_1Y}, \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \ge \mathbb{E}_P[\log q(X_1,X_2,Y)]\Big\}. \qquad (31)$$

Conversely, for any rate pair (R_1, R_2) failing to meet both of (28)–(29)*, the random-coding error probability resulting from* (22)–(24) *tends to one as $n \to \infty$.*

Proof. See Section IV. □

In Appendix B, we cast (28) in terms of convex optimization problems. Similarly to the previous subsection, the exponent corresponding to (29) is precisely that corresponding to the second user in [4, Thm. 1], and the exponent corresponding to (28) is given by

$$\min_{P_{X_1X_2Y} \,:\, P_{X_1X_2} = Q_{X_1X_2}} D(P_{X_1X_2Y} \| Q_{X_1X_2} \times W) + \left[I_{0\mathrm{c}}(P_{X_1X_2Y}, R_2) - R_1\right]^+, \qquad (32)$$

where $I_{0\mathrm{c}}(P_{X_1X_2Y}, R_2)$ denotes the right-hand side of (28) with an arbitrary distribution $P_{X_1X_2Y}$ in place of $Q_{X_1X_2} \times W$. Similarly to the rate region, the proof of Theorem 2 shows that these exponents are tight with respect to the ensemble average (sometimes called *exact random-coding exponents* [27]).

Numerical Example: We consider again consider the transition law (and the corresponding decoding metric with a single value of δ) given in (21) with $\delta_{00} = 0.01$, $\delta_{01} = 0.1$, $\delta_{10} = 0.01$, $\delta_{11} = 0.3$, $\delta = 0.15$, and $Q_{X_1X_2} = Q_1 \times Q_2$ with $Q_1 = Q_2 = (0.5, 0.5)$. Figure 2 plots the achievable rates regions of successive decoding (Theorem 2), maximum-metric decoding ((9)–(10)), and matched decoding (again yielding the same region whether successive or maximum-metric, *cf.* Appendix A).

We see that the behavior of the decoders is analogous to the non-cognitive case observed in Figure 1. The key difference here is that we know that all three regions are tight with respect to the ensemble average. Thus, we may conclude that the somewhat unusual shape of the region for successive decoding is not merely an artifact of our analysis, but it is indeed inherent to the random-coding ensemble and the decoder.

III. Proof of Theorem 1

The proof of Theorem 1 is based on the method of type class enumeration (e.g. see [26]–[28]), and is perhaps most similar to that of Somekh-Baruch and Merhav [27].

Step 1: Initial bound

We assume without loss of generality that $\mathsf{m}_1 = \mathsf{m}_2 = 1$, and we write $\boldsymbol{X}_\nu = \boldsymbol{X}_\nu^{(1)}$ and let $\overline{\boldsymbol{X}}_\nu$ denote an arbitrary codeword $\boldsymbol{X}_\nu^{(j)}$ with $j \ne 1$. Thus,

$$(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}, \overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2) \sim P_{\boldsymbol{X}_1}(\boldsymbol{x}_1) P_{\boldsymbol{X}_2}(\boldsymbol{x}_2) \times W^n(\boldsymbol{y}|\boldsymbol{x}_1, \boldsymbol{x}_2) P_{\boldsymbol{X}_1}(\overline{\boldsymbol{x}}_1) P_{\boldsymbol{X}_2}(\overline{\boldsymbol{x}}_2). \quad (33)$$

We define the following error events:

(Type 1) $\sum_j q^n(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y}) \ge \sum_j q^n(\boldsymbol{X}_1, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y})$ for some $i \ne 1$;

(Type 2) $q^n(\boldsymbol{X}_1, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y}) \ge q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$ for some $j \ne 1$.

Denoting the probabilities of these events by $\overline{p}_{\mathrm{e},1}$ and $\overline{p}_{\mathrm{e},2}$ respectively, it follows that the overall random-coding error probability $\overline{p}_{\mathrm{e}}$ is upper bounded by $\overline{p}_{\mathrm{e},1} + \overline{p}_{\mathrm{e},2}$.

The analysis of the type-2 error event is precisely that of one of the three error types for maximum-metric decoding [3], [5], yielding the rate condition in (17). We thus focus on the type-1 event. We let $\overline{p}_{\mathrm{e},1}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})$ denote the probability of the type-1 event conditioned on $(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}, \boldsymbol{Y}) = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})$, and we denote the joint type of $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})$ by $P_{X_1X_2Y}$. We write the objective function in (2) as

$$\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1) \triangleq q^n(\overline{\boldsymbol{x}}_1, \boldsymbol{x}_2, \boldsymbol{y}) + \sum_{j \ne 1} q^n(\overline{\boldsymbol{x}}_1, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}). \quad (34)$$

This quantity is random due to the randomness of $\{\boldsymbol{X}_2^{(j)}\}$. The starting point of our analysis is the union bound:

$$\overline{p}_{\mathrm{e},1}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) \le (M_1 - 1) \mathbb{P}\big[\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{X}}_1) \ge \Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)\big]. \quad (35)$$

The difficulty in analyzing (35) is that for two different codewords $\boldsymbol{x}_1$ and $\overline{\boldsymbol{x}}_1$, $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)$ and $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1)$ are not independent, and their joint statistics are complicated. We will circumvent this issue by conditioning on high probability events under which these random quantities can be bounded by deterministic values.

Step 2: An auxiliary lemma

We introduce some additional notation. For a given realization $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})$ of $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$, we let $\widetilde{P}_{X_1X_2Y}$ denote its joint type and we write $q^n(\widetilde{P}'_{X_1X_2Y}) \triangleq q^n(\overline{\boldsymbol{x}}_1, \overline{\boldsymbol{x}}_2, \boldsymbol{y})$. In addition, for a general sequence $\overline{\boldsymbol{x}}_1$, we define the *type enumerator*

$$N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) = \sum_{j \ne 1} \mathbb{1}\Big\{(\overline{\boldsymbol{x}}_1, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in T^n(\widetilde{P}'_{X_1X_2Y})\Big\}, \quad (36)$$

which represents the random number of $\boldsymbol{X}_2^{(j)}$ $(j \ne 1)$ such that $(\overline{\boldsymbol{x}}_1, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in T^n(\widetilde{P}'_{X_1X_2Y})$. As we will see below, when $\overline{\boldsymbol{X}}_1 = \overline{\boldsymbol{x}}_1$, the quantity $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1)$ can be re-written in terms of $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\cdot)$, and $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)$ can similarly be re-written in terms of $N_{\boldsymbol{x}_1\boldsymbol{y}}(\cdot)$.

The key to replacing random quantities by deterministic ones is to condition on events that hold with probability one approaching faster than exponentially, thus not affecting the exponential behavior of interest. The following lemma will be used for this purpose, characterizing the behavior of $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})$ for various choices of R_2 and $\widetilde{P}'_{X_1X_2Y}$. The proof can be found in [26], [27], and is based on the fact that

$$\mathbb{P}\big[(\overline{\boldsymbol{x}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in T^n(\widetilde{P}'_{X_1X_2Y})\big] \doteq e^{-nI_{\widetilde{P}'}(X_2; X_1, Y)}, \quad (37)$$

which is a standard property of types [25, Ch. 2].

Lemma 1. [26], [27] *Fix the pair* $(\overline{\boldsymbol{x}}_1, \boldsymbol{y}) \in T^n(\widetilde{P}_{X_1Y})$*, a constant* $\delta > 0$*, and a type* $\widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1,n}(Q_{2,n}, \widetilde{P}_{X_1Y})$*.*

1) *If* $R_2 \ge I_{\widetilde{P}'}(X_2; X_1, Y) + \delta$*, then*

$$M_2 e^{-n(I_{\widetilde{P}'}(X_2; X_1, Y) + \delta)} \le N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) \le M_2 e^{-n(I_{\widetilde{P}'}(X_2; X_1, Y) - \delta)} \quad (38)$$

with probability tending to one faster than exponentially.

2) *If* $R_2 < I_{\widetilde{P}'}(X_2; X_1, Y) + \delta$*, then*

$$N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) \le e^{n\,2\delta} \quad (39)$$

with probability tending to one faster than exponentially.

Roughly speaking, Lemma 1 states that if $R_2 > I_{\widetilde{P}'}(X_2; X_1, Y)$ then the type enumerator is highly concentrated about its mean, whereas if $R_2 < I_{\widetilde{P}'}(X_2; X_1, Y)$ then the type enumerator takes a subexponential value (possibly zero) with overwhelming probability.

Given a joint type $\widetilde{P}_{X_1Y}$, define the event

$$\mathcal{A}_\delta(\widetilde{P}_{X_1Y}) = \Big\{\text{(38) holds for all } \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1,n}(Q_{2,n}, \widetilde{P}_{X_1Y}) \text{ with } R_2 \ge I_{\widetilde{P}'}(X_2; X_1, Y) + \delta\Big\} \cap \Big\{\text{(39) holds for all } \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1,n}(Q_{2,n}, \widetilde{P}_{X_1Y}) \text{ with } R_2 < I_{\widetilde{P}'}(X_2; X_1, Y) + \delta\Big\}, \quad (40)$$

where

$$\mathcal{S}'_{1,n}(Q_{2,n}, \widetilde{P}_{X_1Y}) \triangleq \Big\{\widetilde{P}'_{X_1X_2Y} \in \mathcal{P}_n(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) : \widetilde{P}'_{X_1Y} = \widetilde{P}_{X_1Y}, \widetilde{P}'_{X_2} = Q_{2,n}\Big\}, \quad (41)$$

and where we recall the definition of $Q_{2,n}$ at the start of Section II-A. By Lemma 1 and the union bound, $\mathbb{P}[\mathcal{A}_\delta(\widetilde{P}_{X_1Y})] \to 1$ faster than exponentially. and hence we can safely condition any event on $\mathcal{A}_\delta(\widetilde{P}_{X_1Y})$ without changing the exponential behavior of the corresponding probability. This can be seen by writing the following for any event $\mathcal{E}$:

$$\mathbb{P}[\mathcal{E}] = \mathbb{P}[\mathcal{E}\cap\mathcal{A}] + \mathbb{P}[\mathcal{E}\cap\mathcal{A}^c] \tag{42}$$
$$\le \mathbb{P}[\mathcal{E}\,|\,\mathcal{A}] + \mathbb{P}[\mathcal{A}^c], \tag{43}$$

$$\mathbb{P}[\mathcal{E}] \ge \mathbb{P}[\mathcal{E}\cap\mathcal{A}] \tag{44}$$
$$= (1-\mathbb{P}[\mathcal{A}^c])\mathbb{P}[\mathcal{E}\,|\,\mathcal{A}] \tag{45}$$
$$\ge \mathbb{P}[\mathcal{E}\,|\,\mathcal{A}] - \mathbb{P}[\mathcal{A}^c]. \tag{46}$$

Using these observations, we will condition on $\mathcal{A}_\delta$ several times throughout the remainder of the proof.

Step 3: Bound $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)$ by a deterministic value

From (34), we have

$$\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1) = q^n(\widetilde{P}_{X_1X_2Y}) + \sum_{\widetilde{P}'_{X_1X_2Y}} N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})q^n(\widetilde{P}'_{X_1X_2Y}). \tag{47}$$

Since the codewords are generated independently, $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})$ is binomially distributed with M_2-1 trials and success probability $\mathbb{P}\big[(\overline{\boldsymbol{x}}_1,\overline{\boldsymbol{X}}_2,\boldsymbol{y}) \in T^n(\widetilde{P}'_{X_1X_2Y})\big]$. By construction, we have $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) = 0$ unless $\widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1,n}(Q_{2,n},\widetilde{P}_{X_1Y})$, where $\mathcal{S}'_{1,n}$ is defined in (41).

Conditioned on $\mathcal{A}_\delta(P_{X_1Y})$, we have the following:

$$\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1) = q^n(P_{X_1X_2Y}) + \sum_{P'_{X_1X_2Y}} N_{\boldsymbol{x}_1\boldsymbol{y}}(P'_{X_1X_2Y})q^n(P'_{X_1X_2Y}) \tag{48}$$
$$\ge q^n(P_{X_1X_2Y}) + \max_{\substack{P'_{X_1X_2Y}\in\mathcal{S}'_{1,n}(Q_{2,n},P_{X_1Y})\\ R_2\ge I_{P'}(X_2;X_1,Y)+\delta}} N_{\boldsymbol{x}_1\boldsymbol{y}}(P'_{X_1X_2Y})q^n(P'_{X_1X_2Y}) \tag{49}$$
$$\ge q^n(P_{X_1X_2Y}) + M_2 \times \max_{\substack{P'_{X_1X_2Y}\in\mathcal{S}'_{1,n}(Q_{2,n},P_{X_1Y})\\ R_2\ge I_{P'}(X_2;X_1,Y)+\delta}} e^{-n(I_{P'}(X_2;X_1,Y)+\delta)}q^n(P'_{X_1X_2Y}) \tag{50}$$
$$\triangleq \underline{G}_{\delta,n}(P_{X_1X_2Y}), \tag{51}$$

where (50) follows from (38). Unlike $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)$, the quantity $\underline{G}_{\delta,n}(P_{X_1X_2Y})$ is deterministic. Substituting (51) into (35) and using the fact that $\mathbb{P}\big[\mathcal{A}_\delta(\widetilde{P}_{X_1Y})\big] \to 1$ faster than exponentially, we obtain

$$\bar{p}_{\mathrm{e},1}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}) \,\dot{\le}\, M_1\mathbb{P}\big[\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{X}}_1) \ge \underline{G}_{\delta,n}(P_{X_1X_2Y})\big]. \tag{52}$$

Step 4: An expansion based on types

Since the statistics of $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1)$ depend on $\overline{\boldsymbol{x}}_1$ only through the joint type of $(\overline{\boldsymbol{x}}_1,\boldsymbol{x}_2,\boldsymbol{y})$, we can write (52) as follows:

$$\bar{p}_{\mathrm{e},1}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}) \,\dot{\le}\, M_1 \sum_{\widetilde{P}_{X_1X_2Y}} \mathbb{P}\big[(\overline{\boldsymbol{X}}_1,\boldsymbol{x}_2,\boldsymbol{y})\in T^n(\widetilde{P}_{X_1X_2Y})\big] \times \mathbb{P}\big[\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1) \ge \underline{G}_{\delta,n}(P_{X_1X_2Y})\big] \tag{53}$$
$$\doteq M_1 \max_{\widetilde{P}_{X_1X_2Y}\in\mathcal{S}_{1,n}(Q_{1,n},P_{X_2Y})} e^{-nI_{\widetilde{P}}(X_1;X_2,Y)} \times \mathbb{P}\big[\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1) \ge \underline{G}_{\delta,n}(P_{X_1X_2Y})\big], \tag{54}$$

where $\overline{\boldsymbol{x}}_1$ denotes an arbitrary sequence such that $(\overline{\boldsymbol{x}}_1,\boldsymbol{x}_2,\boldsymbol{y}) \in T^n(\widetilde{P}_{X_1X_2Y})$, and

$$\mathcal{S}_{1,n}(Q_{1,n},P_{X_2Y}) \triangleq \Big\{\widetilde{P}_{X_1X_2Y}\in\mathcal{P}_n(\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Y}) : \widetilde{P}_{X_1} = Q_{1,n}, \widetilde{P}_{X_2Y} = P_{X_2Y}\Big\}. \tag{55}$$

In (54), we have used an analogous property to (37) and the fact that by construction, the joint type of $(\overline{\boldsymbol{X}}_1,\boldsymbol{x}_2,\boldsymbol{y})$ is in $\mathcal{S}_{1,n}(Q_{1,n},P_{X_2Y})$ with probability one.

Step 5: Bound $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1)$ by a deterministic value

Next, we again use Lemma 1 in order to replace $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1)$ in (54) by a deterministic quantity. We have from (47) that

$$\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1) \le q^n(\widetilde{P}_{X_1X_2Y}) + p_0(n)\max_{\widetilde{P}'_{X_1X_2Y}} N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})q^n(\widetilde{P}'_{X_1X_2Y}), \tag{56}$$

where $p_0(n)$ is a polynomial corresponding to the total number of joint types. Substituting (56) into (54), we obtain

$$\bar{p}_{\mathrm{e},1}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}) \,\dot{\le}\, M_1 \max_{\widetilde{P}_{X_1X_2Y}\in\mathcal{S}_{1,n}(Q_{1,n},P_{X_2Y})} \max_{\widetilde{P}'_{X_1X_2Y}\in\mathcal{S}'_1(Q_{2,n},\widetilde{P}_{X_1Y})} e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}\,\mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\big], \tag{57}$$

where

$$\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y}) \triangleq \Big\{q^n(\widetilde{P}_{X_1X_2Y}) + p_0(n)N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})q^n(\widetilde{P}'_{X_1X_2Y}) \ge \underline{G}_{\delta,n}(P_{X_1X_2Y})\Big\}, \tag{58}$$

and we have used the union bound to take the maximum over $\widetilde{P}'_{X_1X_2Y}$ outside the probability in (57). Continuing, we have for any $\widetilde{P}_{X_1X_2Y}\in\mathcal{S}_{1,n}(Q_{1,n},P_{X_2Y})$ that

$$\max_{\widetilde{P}'_{X_1X_2Y}\in\mathcal{S}'_{1,n}(Q_{2,n},\widetilde{P}_{X_1Y})} \mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\big] = \max\Bigg\{\max_{\substack{\widetilde{P}'_{X_1X_2Y}\in\mathcal{S}'_{1,n}(Q_{2,n},\widetilde{P}_{X_1Y})\\ R_2\ge I_{\widetilde{P}'}(X_2;X_1,Y)+\delta}} \mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\big], \max_{\substack{\widetilde{P}'_{X_1X_2Y}\in\mathcal{S}'_{1,n}(Q_{2,n},\widetilde{P}_{X_1Y})\\ R_2< I_{\widetilde{P}'}(X_2;X_1,Y)+\delta}} \mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\big]\Bigg\}. \tag{59}$$

Step 5a – Simplify the First Term: For the first term on the right-hand side of (59), observe that conditioned on $\mathcal{A}_\delta(\widetilde{P}_{X_1Y})$ in (40), we have for $\widetilde{P}'_{X_1X_2Y}$ satisfying $R_2 \geq I_{\widetilde{P}'}(X_2;X_1,Y)+\delta$ that

$$N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})q^n(\widetilde{P}'_{X_1X_2Y}) \leq M_2 e^{-n(I_{\widetilde{P}'}(X_2;X_1,Y)-\delta)}q^n(\widetilde{P}'_{X_1X_2Y}), \quad (60)$$

where we have used (38). Hence, and since $\mathbb{P}\big[\mathcal{A}_\delta(\widetilde{P}_{X_1Y})\big] \to 1$ faster than exponentially, we have

$$\mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\big] \dot{\leq} \mathbb{1}\Big\{q^n(\widetilde{P}_{X_1X_2Y}) + M_2 p_0(n) e^{-n(I_{\widetilde{P}'}(X_2;X_1,Y)-\delta)}q^n(\widetilde{P}'_{X_1X_2Y}) \geq \underline{G}_{\delta,n}(P_{X_1X_2Y})\Big\}. \quad (61)$$

Step 5b – Simplify the Second Term: For the second term on the right-hand side of (59), we define the event $\mathcal{B} \triangleq \big\{N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})>0\big\}$, yielding

$$\mathbb{P}[\mathcal{B}] \dot{\leq} M_2 e^{-nI_{\widetilde{P}'}(X_2;X_1,Y)}, \quad (62)$$

which follows from the union bound and (37). Whenever $R_2 < I_{\widetilde{P}'}(X_2;X_1,Y)+\delta$, we have

$$\begin{aligned}
&\mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\big] \\
&\leq \mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\,\big|\,\mathcal{B}^c\big] + \mathbb{P}[\mathcal{B}]\mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\,\big|\,\mathcal{B}\big] && (63)\\
&\dot{\leq} \mathbb{1}\big\{q^n(\widetilde{P}_{X_1X_2Y}) \geq \underline{G}_{\delta,n}(P_{X_1X_2Y})\big\} \\
&\quad + M_2 e^{-nI_{\widetilde{P}'}(X_2;X_1,Y)}\mathbb{P}\big[\mathcal{E}_{P,\widetilde{P}}(\widetilde{P}'_{X_1X_2Y})\,\big|\,\mathcal{B}\big], && (64)\\
&\dot{\leq} \mathbb{1}\big\{q^n(\widetilde{P}_{X_1X_2Y}) \geq \underline{G}_{\delta,n}(P_{X_1X_2Y})\big\} \\
&\quad + M_2 e^{-nI_{\widetilde{P}'}(X_2;X_1,Y)}\mathbb{1}\big\{q^n(\widetilde{P}_{X_1X_2Y}) \\
&\quad + p_0(n)e^{n\,2\delta}q^n(\widetilde{P}'_{X_1X_2Y}) \geq \underline{G}_{\delta,n}(P_{X_1X_2Y})\big\}, && (65)
\end{aligned}$$

where (64) follows using (62) and (58) along with the fact that $\mathcal{B}^c$ implies $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})=0$, and (65) follows by conditioning on $\mathcal{A}_\delta(\widetilde{P}_{X_1Y})$ and using (39).

Step 6: Deduce the exponent for fixed $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})$

Observe that $\underline{F}(P_{X_1X_2Y},R_2)$ in (14) equals the exponent of $\underline{G}_{\delta,n}$ in (51) in the limit as $\delta\to 0$ and $n\to\infty$. Similarly, the exponents corresponding to the other quantities appearing in the indicator functions in (61) and (65) tend to

$$\overline{F}_1(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y},R_2) \triangleq \max\Big\{\mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)], \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)] + R_2 - I_{\widetilde{P}'}(X_2;X_1,Y)\Big\}, \quad (66)$$

$$\overline{F}_2(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) \triangleq \max\Big\{\mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)], \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)]\Big\}. \quad (67)$$

We claim that combining these expressions with (57), (59), (61) and (65) and taking $\delta\to 0$ (e.g., analogously to [4, p. 737], we may set $\delta = n^{-1/2}$), gives the following:

$$\begin{aligned}
&\bar{p}_{\mathrm{e},1}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}) \dot{\leq} \\
&\max\Bigg\{\max_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y})\in\mathcal{T}_1^{(1)}(P_{X_1X_2Y},R_2)} M_1 e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}, \\
&\max_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y})\in\mathcal{T}_1^{(2)}(P_{X_1X_2Y},R_2)} M_1 e^{-nI_{\widetilde{P}}(X_1;X_2,Y)} \times M_2 e^{-nI_{\widetilde{P}'}(X_2;X_1,Y)}\Bigg\}, \quad (68)
\end{aligned}$$

where[1]

$$\begin{aligned}
\mathcal{T}_1^{(1)}(P_{X_1X_2Y},R_2) \triangleq \Big\{&(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) : \\
&\widetilde{P}_{X_1X_2Y}\in\mathcal{S}_1(Q_1,P_{X_2Y}), \\
&\widetilde{P}'_{X_1X_2Y}\in\mathcal{S}'_1(Q_2,\widetilde{P}_{X_1Y}), \\
&I_{\widetilde{P}'}(X_2;X_1,Y)\leq R_2, \\
&\overline{F}_1(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y},R_2)\geq \underline{F}(P_{X_1X_2Y},R_2)\Big\}, \quad (69)
\end{aligned}$$

$$\begin{aligned}
\mathcal{T}_1^{(2)}(P_{X_1X_2Y},R_2) \triangleq \Big\{&(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) : \\
&\widetilde{P}_{X_1X_2Y}\in\mathcal{S}_1(Q_1,P_{X_2Y}), \\
&\widetilde{P}'_{X_1X_2Y}\in\mathcal{S}'_1(Q_2,\widetilde{P}_{X_1Y}), \\
&I_{\widetilde{P}'}(X_2;X_1,Y)\geq R_2, \\
&\overline{F}_2(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y})\geq \underline{F}(P_{X_1X_2Y},R_2)\Big\}, \quad (70)
\end{aligned}$$

and

$$\mathcal{S}_1(Q_1,P_{X_2Y}) \triangleq \Big\{\widetilde{P}_{X_1X_2Y}\in\mathcal{P}(\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Y}) : \widetilde{P}_{X_1}=Q_1, \widetilde{P}_{X_2Y}=P_{X_2Y}\Big\}, \quad (71)$$

$$\mathcal{S}'_1(Q_2,\widetilde{P}_{X_1Y}) \triangleq \Big\{\widetilde{P}'_{X_1X_2Y}\in\mathcal{P}(\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Y}) : \widetilde{P}'_{X_1Y}=\widetilde{P}_{X_1Y}, \widetilde{P}'_{X_2}=Q_2\Big\}. \quad (72)$$

To see that this is true, we note the following:

- For the first term on the right-hand side of (68), the objective function follows from (56), and the additional constraint $\overline{F}_1(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y},R_2)\geq\underline{F}(P_{X_1X_2Y},R_2)$ in (69) follows since the left-hand side in (61) has exponent $\overline{F}_1$ and the right-hand side has exponent $\underline{F}$ by the definition of $\underline{G}_{\delta,n}$ in (51).
- For the second term on the right-hand side of (68), the objective function follows from (56) and the second term in (65), and the latter (along with $\underline{G}_{\delta,n}$ in (51)) also leads to the final constraint in (70).
- The first term in (65) is upper bounded by the right-hand side of (61), and we already analyzed the latter in order to obtain the first term in (68). Hence, this term can safely be ignored.

[1]Strictly speaking, these sets depend on (Q_1,Q_2), but this dependence need not be explicit, since we have $P_{X_1}=Q_1$ and $P_{X_2}=Q_2$.

Step 7: Deduce the achievable rate region

By a standard property of types [25, Ch. 2], $\mathbb{P}\big[(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{Y})\in T^n(P_{X_1X_2Y})\big]$ decays to zero exponentially fast when $P_{X_1X_2Y}$ is bounded away from $Q_1\times Q_2\times W$. Therefore, we can safely substitute $P_{X_1X_2Y}=Q_1\times Q_2\times W$ to obtain the following rate conditions for the first decoding step:

$$R_1\le\min_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y})\in\mathcal{T}_1^{(1)}(Q_1\times Q_2\times W,R_2)}I_{\widetilde{P}}(X_1;X_2,Y),\tag{73}$$

$$R_1+R_2\le\min_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y})\in\mathcal{T}_1^{(2)}(Q_1\times Q_2\times W,R_2)}I_{\widetilde{P}}(X_1;X_2,Y)+I_{\widetilde{P}'}(X_2;X_1,Y).\tag{74}$$

Finally, we claim that (73)–(74) can be united to obtain (16). To see this, we consider two cases:

- If $R_2>I_{\widetilde{P}'}(X_2;X_1,Y)$, then the $[\cdot]^+$ term in (16) equals zero, yielding the objective in (73). Similarly, in this case, the term $\overline{F}$ in (13) simplifies to $\overline{F}_1$ in (66).
- If $R_2\le I_{\widetilde{P}'}(X_2;X_1,Y)$, then the $[\cdot]^+$ term in (16) equals $I_{\widetilde{P}'}(X_2;X_1,Y)-R_2$, yielding the objective in (73). In this case, the term $\overline{F}$ in (13) simplifies to $\overline{F}_2$ in (67).

IV. Proof of Theorem 2

The achievability and ensemble tightness proofs for Theorem 2 follow similar steps; to avoid repetition, we focus on the ensemble tightness part. The achievability part is obtained using exactly the same high-level steps, while occasionally replacing upper bounds by lower bounds as needed via the techniques presented in Section III.

Step 1: Initial bound

We consider the two error events introduced at the beginning of Section III, and observe that $\overline{p}_{\mathrm{e}}\ge\frac{1}{2}\max\{\overline{p}_{\mathrm{e},1},\overline{p}_{\mathrm{e},2}\}$. The analysis of $\overline{p}_{\mathrm{e},2}$ is precisely that given in [4, Thm. 1], so we focus on $\overline{p}_{\mathrm{e},1}$.

We assume without loss of generality that $\mathsf{m}_1=\mathsf{m}_2=1$, and we write $\boldsymbol{X}_\nu=\boldsymbol{X}_\nu^{(1)}$ ($\nu=1,2$), let $\boldsymbol{X}_2^{(j)}$ denote $\boldsymbol{X}_2^{(1,j)}$, let $\overline{\boldsymbol{X}}_2^{(j)}$ denote $\boldsymbol{X}^{(i,j)}$ for some fixed $i\ne1$, and let $(\overline{\boldsymbol{X}}_1,\overline{\boldsymbol{X}}_2)$ denote $(\boldsymbol{X}_1^{(i)},\boldsymbol{X}_2^{(i,j)})$ for some fixed (i,j) with $i\ne1$. Thus,

$$(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{Y},\overline{\boldsymbol{X}}_1,\overline{\boldsymbol{X}}_2)\sim P_{\boldsymbol{X}_1}(\boldsymbol{x}_1)P_{\boldsymbol{X}_2|\boldsymbol{X}_1}(\boldsymbol{x}_2|\boldsymbol{x}_1)\times W^n(\boldsymbol{y}|\boldsymbol{x}_1,\boldsymbol{x}_2)P_{\boldsymbol{X}_1}(\overline{\boldsymbol{x}}_1)P_{\boldsymbol{X}_2|\boldsymbol{X}_1}(\overline{\boldsymbol{x}}_2|\overline{\boldsymbol{x}}_1).\tag{75}$$

Moreover, analogously to (34), we define

$$\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)\triangleq q^n(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})+\sum_{j\ne1}q^n(\boldsymbol{x}_1,\boldsymbol{X}_2^{(1,j)},\boldsymbol{y})\tag{76}$$

$$\tilde{\Xi}_{\boldsymbol{y}}(\boldsymbol{x}_1^{(i)})\triangleq\sum_j q^n(\boldsymbol{x}_1^{(i)},\boldsymbol{X}_2^{(i,j)},\boldsymbol{y}).\tag{77}$$

Note that here we use separate definitions corresponding to $\boldsymbol{x}_1$ and $\boldsymbol{x}_1^{(i)}$ ($i\ne1$) since in the cognitive MAC, each user-1 sequence is associated with a different set of user-2 sequences.

Fix a joint type $P_{X_1X_2Y}$ and a triplet $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})\in T^n(P_{X_1X_2Y})$, and let $\overline{p}_{\mathrm{e},1}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})$ be the type-1 error probability conditioned on $(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1,1)},\boldsymbol{Y})=(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})$; here we assume without loss of generality that $\mathsf{m}_1=\mathsf{m}_2=1$. We have

$$\overline{p}_{\mathrm{e},1}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})=\mathbb{P}\Bigg[\bigcup_{i=2}^{M_1}\Big\{\tilde{\Xi}_{\boldsymbol{y}}(\boldsymbol{X}_1^{(i)})\ge\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)\Big\}\Bigg]\tag{78}$$

$$\ge\frac{1}{2}\min\Big\{1,(M_1-1)\mathbb{P}\big[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{X}}_1)\ge\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)\big]\Big\},\tag{79}$$

where (79) follows since the truncated union bound is tight to within a factor of $\frac{1}{2}$ for independent events [29, Lemma A.2]. Note that this argument fails for the standard MAC; there, the independence requirement does not hold, so it is unclear whether (35) is tight upon taking the minimum with 1.

We now bound the inner probability in (79), which we denote by $\Phi_1(P_{X_1X_2Y})$. By similarly defining

$$\Phi_2(P_{X_1X_2Y},\widetilde{P}_{X_1Y})\triangleq\mathbb{P}\big[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{X}}_1)\ge\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)\,|\,(\overline{\boldsymbol{X}}_1,\boldsymbol{y})\in T^n(\widetilde{P}_{X_1Y})\big],\tag{80}$$

we obtain

$$\Phi_1(P_{X_1X_2Y})\ge\max_{\widetilde{P}_{X_1Y}}\mathbb{P}\big[(\overline{\boldsymbol{X}}_1,\boldsymbol{y})\in T^n(\widetilde{P}_{X_1Y})\big]\Phi_2(P_{X_1X_2Y},\widetilde{P}_{X_1Y})\tag{81}$$

$$\doteq\max_{\widetilde{P}_{X_1Y}\,:\,\widetilde{P}_{X_1}=Q_{X_1},\widetilde{P}_Y=P_Y}e^{-nI_{\widetilde{P}}(X_1;Y)}\Phi_2(P_{X_1X_2Y},\widetilde{P}_{X_1Y}),\tag{82}$$

where (82) is a standard property of types [25, Ch. 2]. We proceed by bounding Φ_2; to do so, we let $\overline{\boldsymbol{x}}_1$ be an arbitrary sequence such that $(\overline{\boldsymbol{x}}_1,\boldsymbol{y})\in T^n(\widetilde{P}_{X_1Y})$. By symmetry, any such sequence may be considered.

Step 2: Type class enumerators

We write each metric $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}$ in terms of type class enumerators. Specifically, again writing $q^n(P_{X_1X_2Y})$ to denote the n-fold product metric for a given joint type, we note the following analogs of (47):

$$\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\boldsymbol{x}_1)=q^n(P_{X_1X_2Y})+\sum_{P'_{X_1X_2Y}}\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})\tag{83}$$

$$\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1)=\sum_{\widetilde{P}'_{X_1X_2Y}}\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1,\widetilde{P}'_{X_1X_2Y}),\tag{84}$$

where

$$\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})\triangleq N_{\boldsymbol{x}_1\boldsymbol{y}}(P'_{X_1X_2Y})q^n(P'_{X_1X_2Y}),\tag{85}$$

$$\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1,\widetilde{P}'_{X_1X_2Y})\triangleq\tilde{N}_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})q^n(\widetilde{P}'_{X_1X_2Y}),\tag{86}$$

and

$$N_{\boldsymbol{x}_1\boldsymbol{y}}(P'_{X_1X_2Y})\triangleq\sum_{j\ne1}\mathbb{1}\Big\{(\boldsymbol{x}_1,\boldsymbol{X}_2^{(j)},\boldsymbol{y})\in T^n(P'_{X_1X_2Y})\Big\},\tag{87}$$

$$\tilde{N}_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})\triangleq\sum_j\mathbb{1}\Big\{(\overline{\boldsymbol{x}}_1,\overline{\boldsymbol{X}}_2^{(j)},\boldsymbol{y})\in T^n(\widetilde{P}'_{X_1X_2Y})\Big\}.\tag{88}$$

Note the minor differences in these definitions compared to those for the standard MAC, resulting from the differing codebook structure associated with superposition coding. Using these definitions, we can bound (80) as follows:

$$\Phi_2(P_{X_1X_2Y},\widetilde{P}_{X_1Y})$$
$$=\mathbb{P}\bigg[\sum_{\widetilde{P}'_{X_1X_2Y}}\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1,\widetilde{P}'_{X_1X_2Y})$$
$$\geq q^n(P_{X_1X_2Y})+\sum_{P'_{X_1X_2Y}}\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})\bigg] \tag{89}$$
$$\geq\mathbb{P}\bigg[\max_{\widetilde{P}'_{X_1X_2Y}}\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1,\widetilde{P}'_{X_1X_2Y})$$
$$\geq q^n(P_{X_1X_2Y})+p_0(n)\max_{P'_{X_1X_2Y}}\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})\bigg] \tag{90}$$
$$\geq\max_{\widetilde{P}'_{X_1X_2Y}}\mathbb{P}\bigg[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1,\widetilde{P}'_{X_1X_2Y})$$
$$\geq q^n(P_{X_1X_2Y})+p_0(n)\max_{P'_{X_1X_2Y}}\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})\bigg] \tag{91}$$
$$\triangleq\max_{\widetilde{P}'_{X_1X_2Y}}\Phi_3(P_{X_1X_2Y},\widetilde{P}_{X_1Y},\widetilde{P}'_{X_1X_2Y}), \tag{92}$$

where $p_0(n)$ is a polynomial corresponding to the number of joint types.

Step 3: An auxiliary lemma

We define the sets

$$\mathcal{S}_{1\mathrm{c},n}(Q_{X_1,n},P_Y)\triangleq\Big\{\widetilde{P}_{X_1Y}\in\mathcal{P}_n(\mathcal{X}_1\times\mathcal{Y}):\ \widetilde{P}_{X_1}=Q_{X_1,n},\widetilde{P}_Y=P_Y\Big\}, \tag{93}$$

$$\mathcal{S}'_{1\mathrm{c},n}(Q_{X_1X_2,n},\widetilde{P}_{X_1Y})\triangleq\Big\{\widetilde{P}'_{X_1X_2Y}\in\mathcal{P}_n(\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Y}):\ \widetilde{P}'_{X_1Y}=\widetilde{P}_{X_1Y},\widetilde{P}_{X_1X_2}=Q_{X_1X_2,n}\Big\}. \tag{94}$$

The following lemma provides analogous properties to Lemma 1 for joint types within $\mathcal{S}'_{1\mathrm{c},n}$, with suitable modifications to handle the fact that we are proving ensemble tightness rather than achievability. It is based on the fact that $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})$ has a binomial distribution with success probability $\mathbb{P}[(\overline{\boldsymbol{x}}_1,\overline{\boldsymbol{X}}_2,\boldsymbol{y})\in T^n(\widetilde{P}'_{X_1X_2Y})\,|\,\overline{\boldsymbol{X}}_1=\overline{\boldsymbol{x}}_1]\doteq e^{-nI_{\widetilde{P}'}(X_2;Y|X_1)}$ by (23).

Lemma 2. *Fix a joint type $\widetilde{P}_{X_1Y}$ and a pair $(\overline{\boldsymbol{x}}_1,\boldsymbol{y})\in T^n(\widetilde{P}_{X_1Y})$. For any joint type $\widetilde{P}'_{X_1X_2Y}\in\mathcal{S}'_{1,n}(Q_{X_1X_2,n},\widetilde{P}_{X_2Y})$ and constant $\delta>0$, the type enumerator $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})$ satisfies the following:*

1) *If $R_2\geq I_{\widetilde{P}'}(X_2;Y|X_1)-\delta$, then $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})\leq M_2e^{-n(I_{\widetilde{P}'}(X_2;Y|X_1)-2\delta)}$ with probability approaching one faster than exponentially.*
2) *If $R_2\geq I_{\widetilde{P}'}(X_2;Y|X_1)+\delta$, then $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})\geq M_2e^{-n(I_{\widetilde{P}'}(X_2;Y|X_1)+\delta)}$ with probability approaching one faster than exponentially.*
3) *If $R_2\leq I_{\widetilde{P}'}(X_2;Y|X_1)-\delta$, then*
   a) *$N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})\leq e^{n\delta}$ with probability approaching one faster than exponentially;*
   b) *$\mathbb{P}\big[N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})>0\big]\doteq M_2e^{-nI_{\widetilde{P}'}(X_2;Y|X_1)}$.*

Moreover, the analogous properties hold for the type enumerator $N_{\boldsymbol{x}_1\boldsymbol{y}}(P'_{X_1X_2Y})$ and any joint types P_{X_1Y} (with $P_{X_1}=Q_{X_1,n}$) and $P'_{X_1X_2Y}\in\mathcal{S}'_{1,n}(Q_{X_1X_2,n},P_{X_1Y})$.

Proof. Parts 1, 2 and 3a are proved in the same way as Lemma 1; we omit the details to avoid repetition with [26], [27]. Part 3b follows by writing the probability that $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}>0$ as a union of the M_1-1 events in (87) holding, and using the fact that the truncated union bound is tight to within a factor of $\frac{1}{2}$ [29, Lemma A.2]. The truncation need not explicitly be included, since the assumption of part 3 implies that $M_2e^{-nI_{\widetilde{P}'}(X_2;Y|X_1)}\to 0$. □

Given a joint type P_{X_2Y} (respectively, $\widetilde{P}_{X_1Y}$), let $\mathcal{A}_\delta(\widetilde{P}_{X_1Y})$ (respectively, $\tilde{\mathcal{A}}_\delta(\widetilde{P}_{X_1Y})$) denote the union of the high-probability events in Lemma 2 (in parts 1, 2 and 3a) taken over all $P'_{X_1X_2Y}\in\mathcal{S}_{1,n}(Q_{X_1X_2},P_{X_2Y})$ (respectively, $\widetilde{P}'_{X_1X_2Y}\in\mathcal{S}'_{1,n}(Q_{X_1X_2},\widetilde{P}_{X_1Y})$). By the union bound, the probability of these events tends to one faster than exponentially, and hence we can safely condition any event accordingly without changing the exponential behavior of the corresponding probability (see (42)–(46)).

Step 4: Bound $\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})$ by a deterministic value

We first deal with $\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})$ in (91). Defining the event

$$\mathcal{B}_\delta\triangleq\Big\{N_{\boldsymbol{x}_1\boldsymbol{y}}(P'_{X_1X_2Y})=0\text{ for all }P'_{X_1X_2Y}\ \text{such that }R_2\leq I_{\widetilde{P}'}(X_2;Y|X_1)-\delta\Big\}, \tag{95}$$

we immediately obtain from Property 3b in Lemma 2 that $\mathbb{P}\big[\mathcal{B}_\delta^c\big]\,\dot{\leq}\,e^{-n\delta}\to 0$, and hence

$$\Phi_3(P_{X_1X_2Y},\widetilde{P}_{X_1Y},\widetilde{P}'_{X_1X_2Y})$$
$$\geq\mathbb{P}\bigg[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1,\widetilde{P}'_{X_1X_2Y})\geq q^n(P_{X_1X_2Y})+p_0(n)\max_{P'_{X_1X_2Y}}\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})\cap\mathcal{B}_\delta\bigg] \tag{96}$$
$$\doteq\mathbb{P}\bigg[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1,\widetilde{P}'_{X_1X_2Y})\geq q^n(P_{X_1X_2Y})+p_0(n)\max_{P'_{X_1X_2Y}}\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})\,\Big|\,\mathcal{B}_\delta\bigg]. \tag{97}$$

Next, conditioned on $\mathcal{B}_\delta$ and the events in Lemma 2, we have

$$q^n(P_{X_1X_2Y})+p_0(n)\max_{P'_{X_1X_2Y}}\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y})$$
$$=q^n(P_{X_1X_2Y})+p_0(n)\max_{\substack{P'_{X_1X_2Y}\in\mathcal{S}'_{1\mathrm{c},n}(Q_{X_1X_2,n},\widetilde{P}_{X_1Y}):\\ R_2\geq I_{P'}(X_2;Y|X_1)-\delta}}\Xi_{\boldsymbol{y}}(\boldsymbol{x}_1,P'_{X_1X_2Y}) \tag{98}$$
$$\leq q^n(P_{X_1X_2Y})+p_0(n)\times\max_{\substack{P'_{X_1X_2Y}\in\mathcal{S}'_{1\mathrm{c},n}(Q_{X_1X_2,n},\widetilde{P}_{X_1Y}):\\ R_2\geq I_{P'}(X_2;Y|X_1)-\delta}}M_2e^{-n(I_{\widetilde{P}'}(X_2;Y|X_1)-2\delta)}\times q^n(P'_{X_1X_2Y}) \tag{99}$$

$$\triangleq \overline{G}_{\delta,n}(P_{X_1X_2Y}), \tag{100}$$

where in (99) we used part 1 of Lemma 2. It follows that

$$\begin{aligned}\Phi_3(P_{X_1X_2Y}, \widetilde{P}_{X_1Y}, \widetilde{P}'_{X_1X_2Y}) \\ \dot{\geq} \mathbb{P}\big[\Xi_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y})\big],\end{aligned} \tag{101}$$

where the conditioning on $\mathcal{B}_\delta$ has been removed since it is independent of the statistics of $\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y})$.

Step 5: Bound $\Xi_{\boldsymbol{x}_2\boldsymbol{y}}(\overline{\boldsymbol{x}}_1)$ by a deterministic value

We now deal with the quantity $\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y})$. Substituting (101) into (92) and constraining the maximization in two different ways, we obtain

$$\begin{aligned}\Phi_2(P_{X_1X_2Y}, \widetilde{P}_{X_1Y}) \,\dot{\geq}\, \max\Bigg\{ \\ \max_{\substack{\widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1\mathrm{c},n}(Q_{X_1X_2,n}, \widetilde{P}_{X_1Y}) : \\ R_2 \geq I_{\widetilde{P}'}(X_2;Y|X_1) + \delta}} \mathbb{P}\Big[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y})\Big], \\ \max_{\substack{\widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1\mathrm{c},n}(Q_{X_1X_2,n}, \widetilde{P}_{X_1Y}) : \\ R_2 \leq I_{\widetilde{P}'}(X_2;Y|X_1) - \delta}} \mathbb{P}\Big[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y})\Big]\Bigg\}.\end{aligned} \tag{102}$$

For $R_2 \geq I_{\widetilde{P}'}(X_2;Y|X_1) + \delta$, we have from part 2 of Lemma 2 that, conditioned on $\tilde{\mathcal{A}}_\delta(\widetilde{P}_{X_1Y})$,

$$\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y}) \geq M_2 e^{-n(I_{\widetilde{P}'}(X_2;Y|X_1)+\delta)} q^n(\widetilde{P}'_{X_1X_2Y}). \tag{103}$$

On the other hand, for $R_2 \leq I_{\widetilde{P}'}(X_2;Y|X_1) - \delta$, we have

$$\begin{aligned}&\mathbb{P}\Big[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y})\Big] &(104)\\ &= \mathbb{P}\Big[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y}) \cap N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) > 0\Big] &(105)\\ &= \mathbb{P}\Big[N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) > 0\Big]\mathbb{P}\Big[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y}) \,\Big|\, N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) > 0\Big] &(106)\\ &\doteq M_2 e^{-nI_{\widetilde{P}'}(X_2;Y|X_1)}\mathbb{P}\Big[\tilde{\Xi}_{\boldsymbol{y}}(\overline{\boldsymbol{x}}_1, \widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y}) \,\Big|\, N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) > 0\Big] &(107)\\ &\dot{\geq} \mathbb{1}\Big\{q^n(\widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y})\Big\} M_2 e^{-nI_{\widetilde{P}'}(X_2;Y|X_1)}, &(108)\end{aligned}$$

where (105) follows since the event under consideration is zero unless $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y}) > 0$, (107) follows from part 3b of Lemma 2, and (108) follows since when $N_{\overline{\boldsymbol{x}}_1\boldsymbol{y}}(\widetilde{P}'_{X_1X_2Y})$ is positive it must be at least one.

Step 6: Deduce the exponent for fixed $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})$

We have now handled both cases in (102). Combining them, and substituting the result into (82), we obtain

$$\begin{aligned}\Phi_1(P_{X_1X_2Y}) \,\dot{\geq} \max_{\widetilde{P}_{X_1Y} \in \mathcal{S}_{1\mathrm{c},n}(Q_{X_1,n}, P_Y)} e^{-nI_{\widetilde{P}}(X_1;Y)} \\ \times \max\Bigg\{ \max_{\substack{\widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1\mathrm{c},n}(Q_{X_1X_2,n}, \widetilde{P}_{X_1Y}) : \\ R_2 \geq I_{\widetilde{P}'}(X_2;Y|X_1) + \delta}} \mathbb{1}\Big\{M_2 e^{-n(I_{\widetilde{P}'}(X_2;Y|X_1)+\delta)} \times q^n(\widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y})\Big\}, \\ \max_{\substack{\widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1\mathrm{c},n}(Q_{X_1X_2,n}, \widetilde{P}_{X_1Y}) : \\ R_2 \leq I_{\widetilde{P}'}(X_2;Y|X_1) - \delta}} M_2 e^{-nI_{\widetilde{P}'}(X_2;Y|X_1)} \times \mathbb{1}\Big\{q^n(\widetilde{P}'_{X_1X_2Y}) \geq \overline{G}_{\delta,n}(P_{X_1X_2Y})\Big\}\Bigg\}.\end{aligned} \tag{109}$$

Observe that $\overline{F}_{\mathrm{c}}(P_{X_1X_2Y})$ in (14) equals the exponent of $\overline{G}_{\delta,n}$ in (100) in the limit as $\delta \to 0$ and $n \to \infty$. Similarly, the exponent corresponding to the quantity in the first indicator function in (109) tends to

$$\begin{aligned}\overline{F}_{1\mathrm{c}}(\widetilde{P}'_{X_1X_2Y}, R_2) \\ \triangleq \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)] + R_2 - I_{\widetilde{P}'}(X_2;Y|X_1).\end{aligned} \tag{110}$$

Recalling that Φ_1 is the inner probability in (79), we obtain the following by taking $\delta \to 0$ sufficiently slowly and using the continuity of the underlying terms in the optimizations:

$$\begin{aligned}\overline{p}_{\mathrm{e},1}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) \,\dot{\geq}\, \max\Bigg\{ \max_{(\widetilde{P}_{X_1Y}, \widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_{1\mathrm{c}}^{(1)}(P_{X_1X_2Y}, R_2)} M_1 e^{-nI_{\widetilde{P}}(X_1;Y)}, \\ \max_{(\widetilde{P}_{X_1Y}, \widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_{1\mathrm{c}}^{(2)}(P_{X_1X_2Y}, R_2)} M_1 e^{-nI_{\widetilde{P}}(X_1;Y)} \times M_2 e^{-nI_{\widetilde{P}'}(X_2;Y|X_1)}\Bigg\},\end{aligned} \tag{111}$$

where

$$\begin{aligned}\mathcal{T}_{1\mathrm{c}}^{(1)}(P_{X_1X_2Y}, R_2) \triangleq \Big\{(\widetilde{P}_{X_1Y}, \widetilde{P}'_{X_1X_2Y}) : \\ \widetilde{P}_{X_1Y} \in \mathcal{S}_{1\mathrm{c}}(Q_{X_1}, P_Y), \\ \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1\mathrm{c}}(Q_{X_1X_2}, \widetilde{P}_{X_1Y}), \\ I_{\widetilde{P}'}(X_2;Y|X_1) \leq R_2, \\ \overline{F}_{1\mathrm{c}}(\widetilde{P}'_{X_1X_2Y}, R_2) \geq \underline{F}_{\mathrm{c}}(P_{X_1X_2Y}, R_2)\Big\},\end{aligned} \tag{112}$$

$$\begin{aligned}\mathcal{T}_{1\mathrm{c}}^{(2)}(P_{X_1X_2Y}, R_2) \triangleq \Big\{(\widetilde{P}_{X_1Y}, \widetilde{P}'_{X_1X_2Y}) : \\ \widetilde{P}_{X_1Y} \in \mathcal{S}_{1\mathrm{c}}(Q_{X_1}, P_Y), \\ \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1\mathrm{c}}(Q_{X_1X_2}, \\ \widetilde{P}_{X_1Y}), I_{\widetilde{P}'}(X_2;Y|X_1) \geq R_2, \\ \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)] \geq \underline{F}(P_{X_1X_2Y}, R_2)\Big\},\end{aligned} \tag{113}$$

and

$$\mathcal{S}_{1c}(Q_{X_1}, P_Y) \triangleq \Big\{ \widetilde{P}_{X_1X_2Y} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) : \widetilde{P}_{X_1} = Q_{X_1}, \widetilde{P}_Y = P_Y \Big\}, \quad (114)$$

$$\mathcal{S}'_{1c}(Q_{X_1X_2}, \widetilde{P}_{X_1Y}) \triangleq \Big\{ \widetilde{P}'_{X_1X_2Y} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) : \widetilde{P}'_{X_1Y} = \widetilde{P}_{X_1Y}, \widetilde{P}'_{X_1X_2} = Q_{X_1X_2} \Big\}. \quad (115)$$

More specifically, this follows from the same argument as Step 6 in Section III.

Step 7: Deduce the achievable rate region

Similarly to Section III, the fact that the joint type of $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$ approaches $Q_{X_1X_2} \times W$ with probability approaching one means that we can substitute $P_{X_1X_2Y} = Q_{X_1X_2} \times W$ to obtain the following rate conditions:

$$R_1 \le \min_{(\widetilde{P}_{X_1Y}, \widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_{1c}^{(1)}(Q_{X_1X_2} \times W, R_2)} I_{\widetilde{P}}(X_1;Y), \quad (116)$$

$$R_1 + R_2 \le \min_{(\widetilde{P}_{X_1Y}, \widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_{1c}^{(2)}(Q_{X_1X_2} \times W, R_2)} I_{\widetilde{P}}(X_1;Y) + I_{\widetilde{P}'}(X_2;Y|X_1). \quad (117)$$

The proof of (28) is now concluded via the same argument as Step 7 in Section III, using the definitions of $\overline{F}_c$, $\overline{F}_{1c}$, $\mathcal{S}_{1c}$, $\mathcal{S}'_{1c}$, $\mathcal{T}_{1c}^{(1)}$ and $\mathcal{T}_{1c}^{(2)}$ to unite (116)–(117). Note that the optimization variable $\widetilde{P}_{X_1Y}$ can be absorbed into $\widetilde{P}'_{X_1X_2Y}$ due to the constraint $\widetilde{P}'_{X_1Y} = \widetilde{P}_{X_1Y}$.

V. Conclusion

We have obtained error exponents and achievable rates for both the standard and cognitive MAC using a mismatched multi-letter successive decoding rule. For the cognitive case, we have proved ensemble tightness, thus allowing us to conclusively establish that there are cases in which neither the mismatched successive decoding region nor the mismatched maximum-metric decoding region [3] dominate each other in the random coding setting.

An immediate direction for further work is to establish the ensemble tightness of the achievable rate region for the standard MAC in Theorem 1. A more challenging open question is to determine whether either of the *true* mismatched capacity regions (rather than just achievable random-coding regions) for the two decoding rules contain each other in general.

Appendix A
Behavior of Successive Decoder with $q = W$

Here we show that a rate pair (R_1, R_2) or error exponent $E(R_1, R_2)$ is achievable under maximum-likelihood (ML) decoding if and only if it is achievable under the successive rule in (2)–(3) with $q(x_1, x_2, y) = W(y|x_1, x_2)$. This is shown in the same way for the standard MAC and the cognitive MAC, so we focus on the former.

It suffices to show that, for any fixed codebooks $\mathcal{C}_1 = \{\boldsymbol{x}_1^{(i)}\}_{i=1}^{M_1}$ and $\mathcal{C}_2 = \{\boldsymbol{x}_2^{(j)}\}_{j=1}^{M_2}$, the error probability under ML decoding is lower bounded by a constant times the error probability under successive decoding. It also suffices to consider the variations where ties are broken as errors, since doing so reduces the error probability by at most a factor of two [30]. Formally, we consider the following:

1) The ML decoder maximizing $W^n(\boldsymbol{y}|\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)})$;
2) The successive decoder in (2)–(3) with $q = W$;
3) The genie-aided successive decoder using the true value of m_1 on the second step rather than $\hat{\mathsf{m}}_1$ [11]:

$$\hat{\mathsf{m}}_1 = \arg\max_i \sum_j W^n(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y}), \quad (118)$$

$$\hat{\mathsf{m}}_2 = \arg\max_j W^n(\boldsymbol{x}_1^{(\mathsf{m}_1)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y}). \quad (119)$$

We denote the probabilities under these decoders by $\mathbb{P}^{(\mathrm{ML})}[\cdot]$, $\mathbb{P}^{(\mathrm{S})}[\cdot]$ and $\mathbb{P}^{(\mathrm{Genie})}[\cdot]$ respectively. Denoting the random message pair by $(\mathsf{m}_1, \mathsf{m}_2)$, the resulting estimate by $(\hat{\mathsf{m}}_1, \hat{\mathsf{m}}_2)$, and the output sequence by $\boldsymbol{Y}$, we have

$$\begin{aligned}
&\mathbb{P}^{(\mathrm{ML})}[(\hat{\mathsf{m}}_1, \hat{\mathsf{m}}_2) \ne (\mathsf{m}_1, \mathsf{m}_2)] \\
&\ge \max\Bigg\{ \mathbb{P}^{(\mathrm{ML})}[\hat{\mathsf{m}}_1 \ne \mathsf{m}_1], \\
&\qquad \mathbb{P}^{(\mathrm{ML})}\Bigg[\bigcup_{j \ne \mathsf{m}_2} \Bigg\{ \frac{W^n(\boldsymbol{x}_1^{(\mathsf{m}_1)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{Y})}{W^n(\boldsymbol{x}_1^{(\mathsf{m}_1)}, \boldsymbol{x}_2^{(\mathsf{m}_2)}, \boldsymbol{Y})} \ge 1 \Bigg\} \Bigg] \Bigg\} && (120) \\
&\ge \max\Bigg\{ \mathbb{P}^{(\mathrm{Genie})}[\hat{\mathsf{m}}_1 \ne \mathsf{m}_1], \\
&\qquad \mathbb{P}^{(\mathrm{Genie})}\Bigg[\bigcup_{j \ne \mathsf{m}_2} \Bigg\{ \frac{W^n(\boldsymbol{x}_1^{(\mathsf{m}_1)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{Y})}{W^n(\boldsymbol{x}_1^{(\mathsf{m}_1)}, \boldsymbol{x}_2^{(\mathsf{m}_2)}, \boldsymbol{Y})} \ge 1 \Bigg\} \Bigg] \Bigg\} && (121) \\
&\ge \frac{1}{2} \mathbb{P}^{(\mathrm{Genie})}[(\hat{\mathsf{m}}_1, \hat{\mathsf{m}}_2) \ne (\mathsf{m}_1, \mathsf{m}_2)] && (122) \\
&= \frac{1}{2} \mathbb{P}^{(\mathrm{S})}[(\hat{\mathsf{m}}_1, \hat{\mathsf{m}}_2) \ne (\mathsf{m}_1, \mathsf{m}_2)], && (123)
\end{aligned}$$

where (121) follows since the two steps of the genie-aided decoder correspond to minimizing the two terms in the $\max\{\cdot,\cdot\}$, (122) follows by writing $\max\{\mathbb{P}[A], \mathbb{P}[B]\} \ge \frac{1}{2}(\mathbb{P}[A] + \mathbb{P}[B]) \ge \frac{1}{2}\mathbb{P}[A \cup B]$, and (123) follows since the overall error probability is unchanged by the genie [11].

Appendix B
Formulations of (16) and (28) in Terms of Convex Optimization Problems

In this section, we provide an alternative formulation of (16) that is written in terms of convex optimization problems. We start with the alternative formulation in (73)–(74). We first note that (74) holds if and only if

$$R_1 \le \min_{(\widetilde{P}_{X_1X_2Y}, \widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_1^{(2)}(Q_1 \times Q_2 \times W, R_2)} I_{\widetilde{P}}(X_1; X_2, Y) + \big[I_{\widetilde{P}'}(X_2; X_1, Y) - R_2 \big]^+, \quad (124)$$

since the argument to $[\cdot]^+$ is non-negative when $I_{\widetilde{P}'}(X_2; X_1, Y) \ge R_2$. Next, we claim that when combining (73) and (124), the rate region is unchanged if the constraint $I_{\widetilde{P}'}(X_2; X_1, Y) \ge R_2$ is omitted from (124). This is seen by noting that whenever $I_{\widetilde{P}'}(X_2; X_1, Y) < R_2$, the objective in

$$\mathcal{T}_1^{(1,1)}(P_{X_1X_2Y},R_2) \triangleq \Big\{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) : \widetilde{P}_{X_1X_2Y} \in \mathcal{S}_1(Q_1,P_{X_2Y}), \\ \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_1(Q_2,\widetilde{P}_{X_1Y}), I_{\widetilde{P}'}(X_2;X_1,Y) \le R_2, \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \ge \underline{F}(P_{X_1X_2Y},R_2)\Big\}, \quad (125)$$

$$\mathcal{T}_1^{(1,2)}(P_{X_1X_2Y},R_2) \triangleq \Big\{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) : \widetilde{P}_{X_1X_2Y} \in \mathcal{S}_1(Q_1,P_{X_2Y}), \\ \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_1(Q_2,\widetilde{P}_{X_1Y}), I_{\widetilde{P}'}(X_2;X_1,Y) \le R_2, \\ \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)] + R_2 - I_{\widetilde{P}'}(X_2;X_1,Y) \ge \underline{F}(P_{X_1X_2Y},R_2)\Big\}, \quad (126)$$

$$\mathcal{T}_1^{(2,1)}(P_{X_1X_2Y},R_2) \triangleq \Big\{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) : \widetilde{P}_{X_1X_2Y} \in \mathcal{S}_1(Q_1,P_{X_2Y}), \\ \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_1(Q_2,\widetilde{P}_{X_1Y}), \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \ge \underline{F}(P_{X_1X_2Y},R_2)\Big\}, \quad (127)$$

$$\mathcal{T}_1^{(2,2)}(P_{X_1X_2Y},R_2) \triangleq \Big\{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) : \widetilde{P}_{X_1X_2Y} \in \mathcal{S}_1(Q_1,P_{X_2Y}), \\ \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_1(Q_2,\widetilde{P}_{X_1Y}), \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)] \ge \underline{F}(P_{X_1X_2Y},R_2)\Big\}. \quad (128)$$

(124) coincides with that of (73), whereas the latter has a less restrictive constraint since $\overline{F}_1 > \overline{F}_2$ (see (66)–(67)).

We now deal with the non-concavity of the functions $\overline{F}_1$ and $\overline{F}_2$ appearing in the sets $\mathcal{T}_1^{(1)}$ and $\mathcal{T}_1^{(2)}$. Using the identity

$$\min_{x \le \max\{a,b\}} f(x) = \min\Big\{\min_{x\le a} f(x), \min_{x \le b} f(x)\Big\}, \quad (129)$$

we obtain the following rate conditions from (73) and (124):

$$R_1 \le \min_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_1^{(1,1)}(Q_1\times Q_2\times W,R_2)} I_{\widetilde{P}}(X_1;X_2,Y), \quad (130)$$

$$R_1 \le \min_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_1^{(1,2)}(Q_1\times Q_2\times W,R_2)} I_{\widetilde{P}}(X_1;X_2,Y), \quad (131)$$

$$R_1 \le \min_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_1^{(2,1)}(Q_1\times Q_2\times W,R_2)} I_{\widetilde{P}}(X_1;X_2,Y) + \big[I_{\widetilde{P}'}(X_2;X_1,Y) - R_2\big]^+, \quad (132)$$

$$R_1 \le \min_{(\widetilde{P}_{X_1X_2Y},\widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_1^{(2,2)}(Q_1\times Q_2\times W,R_2)} I_{\widetilde{P}}(X_1;X_2,Y) + \big[I_{\widetilde{P}'}(X_2;X_1,Y) - R_2\big]^+, \quad (133)$$

where the constraint sets are defined in (125)–(128) at the top of the page. These are obtained from $\mathcal{T}^{(k)}$ ($k=1,2$) by keeping only one term in the definition of $\overline{F}_k$ (see (66)–(67)), and by removing the constraint $I_{\widetilde{P}'}(X_2;X_1,Y) \ge R_2$ when $k=2$ in accordance with the discussion following (124).

The variable $\widetilde{P}'_{X_1X_2Y}$ can be removed from both (130) and (132), since in each case the choice $\widetilde{P}'_{X_1X_2Y}(x_1,x_2,y) = Q_2(x_2)\widetilde{P}_{X_1Y}(x_1,y)$ is feasible and yields $I_{\widetilde{P}'}(X_2;X_1,Y) = 0$. It follows that (130) and (132) yield the same value, and we conclude that (16) can equivalently be expressed in terms of three conditions: (131), (133), and

$$R_1 \le \min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_1^{(1,1')}(Q_1\times Q_2\times W,R_2)} I_{\widetilde{P}}(X_1;X_2,Y), \quad (134)$$

where the set

$$\mathcal{T}_1^{(1,1')}(P_{X_1X_2Y},R_2) \triangleq \Big\{\widetilde{P}_{X_1X_2Y} \in \mathcal{S}_1(Q_1,P_{X_2Y}) : \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \ge \underline{F}(P_{X_1X_2Y},R_2)\Big\} \quad (135)$$

is obtained by eliminating $\widetilde{P}'_{X_1X_2Y}$ from either (125) or (127). These three conditions are all written as convex optimization problems, as desired.

Starting with (116)–(117), one can follow a (a simplified version of) the above arguments for the cognitive MAC to show that (28) holds if and only if

$$R_1 \le \min_{(\widetilde{P}_{X_1Y},\widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_{1\mathrm{c}}^{(1)}(Q_{X_1X_2}\times W,R_2)} I_{\widetilde{P}}(X_1;Y), \quad (136)$$

$$R_1 \le \min_{(\widetilde{P}_{X_1Y},\widetilde{P}'_{X_1X_2Y}) \in \mathcal{T}_{1\mathrm{c}}^{(2')}(Q_{X_1X_2}\times W,R_2)} I_{\widetilde{P}}(X_1;Y) + \big[I_{\widetilde{P}'}(X_2;Y|X_1) - R_2\big]^+. \quad (137)$$

where

$$\mathcal{T}_{1\mathrm{c}}^{(2')}(P_{X_1X_2Y},R_2) \triangleq \Big\{(\widetilde{P}_{X_1Y},\widetilde{P}'_{X_1X_2Y}) : \\ \widetilde{P}_{X_1Y} \in \mathcal{S}_{1\mathrm{c}}(Q_{X_1},P_Y), \widetilde{P}'_{X_1X_2Y} \in \mathcal{S}'_{1\mathrm{c}}(Q_{X_1X_2},\widetilde{P}_{X_1Y}), \\ \mathbb{E}_{\widetilde{P}'}[\log q(X_1,X_2,Y)] \ge \underline{F}(P_{X_1X_2Y},R_2)\Big\}, \quad (138)$$

and where $\mathcal{T}_{1\mathrm{c}}^{(1)}$, $\mathcal{S}_{1\mathrm{c}}$ and $\mathcal{S}'_{1\mathrm{c}}$ are defined in (112)–(115).

References

[1] I. Csiszár and P. Narayan, "Channel capacity for a given decoding metric," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 35–43, Jan. 1995.

[2] A. Ganti, A. Lapidoth, and E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.

[3] A. Lapidoth, "Mismatched decoding and the multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1439–1452, Sept. 1996.

[4] A. Somekh-Baruch, "On achievable rates and error exponents for channels with mismatched decoding," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 727–740, Feb. 2015.

[5] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Multiuser random coding techniques for mismatched decoding," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3950–3970, July 2016.

[6] A. Feinstein, “A new basic theorem of information theory,” *IRE Prof. Group. on Inf. Theory*, vol. 4, no. 4, pp. 2–22, Sept. 1954.

[7] A. Somekh-Baruch, “A general formula for the mismatch capacity,” *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4554–4568, Sept 2015.

[8] M. H. Yassaee, M. R. Aref, and A. Gohari, “A technique for deriving one-shot achievability results in network information theory,” in *IEEE Int. Symp. Inf. Theory*, 2013.

[9] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, “The likelihood decoder: Error exponents and mismatch,” in *IEEE Int. Symp. Inf. Theory*, Hong Kong, 2015.

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

[11] A. Grant, B. Rimoldi, R. Urbanke, and P. Whiting, “Rate-splitting multiple access for discrete memoryless channels,” *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 873–890, 2001.

[12] G. Kaplan and S. Shamai, “Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment,” *Arch. Elek. Über.*, vol. 47, no. 4, pp. 228–239, 1993.

[13] J. Hui, “Fundamental issues of multiple accessing,” Ph.D. dissertation, MIT, 1983.

[14] I. Csiszár and J. Körner, “Graph decomposition: A new key to coding theorems,” *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 5–12, Jan. 1981.

[15] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, “On information rates for mismatched decoders,” *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.

[16] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, “Mismatched decoding: Error exponents, second-order rates and saddlepoint approximations,” *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2647–2666, May 2014.

[17] A. Nazari, A. Anastasopoulos, and S. S. Pradhan, “Error exponent for multiple-access channels: Lower bounds,” *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5095–5115, Sep. 2014.

[18] R. Etkin, N. Merhav, and E. Ordentlich, “Error exponents of optimum decoding for the interference channel,” *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 40–56, Jan. 2010.

[19] Y. Kaspi and N. Merhav, “Error exponents for broadcast channels with degraded message sets,” *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 101–123, Jan. 2011.

[20] W. Huleihel and N. Merhav, “Random coding error exponents for the two-user interference channel,” *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1019–1042, Feb. 2017.

[21] M. Alsan, “Performance of mismatched polar codes over BSCs,” in *Int. Symp. Inf. Theory Apps.*, 2012.

[22] ——, “A lower bound on achievable rates by polar codes with mismatch polar decoding,” in *Inf. Theory Workshop*, 2013.

[23] M. Alsan and E. Telatar, “Polarization as a novel architecture to boost the classical mismatched capacity of B-DMCs,” in *Inf. Theory Workshop*, 2014.

[24] M. Alsan, “Re-proving channel polarization theorems: An extremality and robustness analysis,” Ph.D. dissertation, EPFL, 2014.

[25] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.

[26] R. Etkin, N. Merhav, and E. Ordentlich, “Error exponents of optimum decoding for the interference channel,” *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 40–56, Jan. 2010.

[27] A. Somekh-Baruch and N. Merhav, “Exact random coding exponents for erasure decoding,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6444–6454, Oct. 2011.

[28] N. Merhav, “Error exponents of erasure/list decoding revisited via moments of distance enumerators,” *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4439–4447, Oct. 2008.

[29] N. Shulman, “Communication over an unknown channel via common broadcasting,” Ph.D. dissertation, Tel Aviv University, 2003.

[30] P. Elias, “Coding for two noisy channels,” in *Third London Symp. Inf. Theory*, 1955.

Jonathan Scarlett (S’14 – M’15) received the B.Eng. degree in electrical engineering and the B.Sci. degree in computer science from the University of Melbourne, Australia in 2010. From October 2011 to August 2014, he was a Ph.D. student in the Signal Processing and Communications Group at the University of Cambridge, United Kingdom. From September 2014 to September 2017, he was post-doctoral researcher with the Laboratory for Information and Inference Systems at the École Polytechnique Fédérale de Lausanne, Switzerland. Since January 2018, he has been an assistant professor in the Department of Computer Science and Department of Mathematics, National University of Singapore. His research interests are in the areas of information theory, machine learning, signal processing, and high-dimensional statistics. He received the Cambridge Australia Poynton International Scholarship, and the ‘EPFL Fellows’ postdoctoral fellowship co-funded by Marie Curie.

Alfonso Martinez (SM’11) was born in Zaragoza, Spain, in October 1973. He is currently a Ramón y Cajal Research Fellow at Universitat Pompeu Fabra, Barcelona, Spain. He obtained his Telecommunications Engineering degree from the University of Zaragoza in 1997. In 1998-2003 he was a Systems Engineer at the research centre of the European Space Agency (ESAESTEC) in Noordwijk, The Netherlands. His work on APSK modulation was instrumental in the definition of the physical layer of DVB-S2. From 2003 to 2007 he was a Research and Teaching Assistant at Technische Universiteit Eindhoven, The Netherlands, where he conducted research on digital signal processing for MIMO optical systems and on optical communication theory. Between 2008 and 2010 he was a post-doctoral fellow with the Information-Theoretic Learning Group at Centrum Wiskunde & Informatica (CWI), in Amsterdam, The Netherlands. In 2011 he was a Research Associate with the Signal Processing and Communications Lab at the Department of Engineering, University of Cambridge, Cambridge, U.K. His research interests lie in the fields of information theory and coding, with emphasis on digital modulation and the analysis of mismatched decoding; in this area he has coauthored a monograph on “Bit-Interleaved Coded Modulation.” More generally, he is intrigued by the connections between information theory, optical communications, and physics, particularly by the links between classical and quantum information theory.

Albert Guillén i Fàbregas (S’01 – M’05 – SM’09) received the Telecommunication Engineering Degree and the Electronics Engineering Degree from Universitat Politècnica de Catalunya and Politecnico di Torino, respectively in 1999, and the Ph.D. in Communication Systems from École Polytechnique Fédérale de Lausanne (EPFL) in 2004. Since 2011 he has been an ICREA Research Professor at Universitat Pompeu Fabra. He is also an Adjunct Researcher at the University of Cambridge. He has held appointments at the New Jersey Institute of Technology, Telecom Italia, European Space Agency (ESA), Institut Eurècom, University of South Australia, University of Cambridge, as well as visiting appointments at EPFL, École Nationale des Télécommunications (Paris), Universitat Pompeu Fabra, University of South Australia, Centrum Wiskunde & Informatica and Texas A&M University in Qatar. His research interests are in the areas of information theory, coding theory and communication theory. Dr. Guillén i Fàbregas is a Member of the Young Academy of Europe, and received the Starting Grant from the European Research Council, the Young Authors Award of the 2004 European Signal Processing Conference, the 2004 Best Doctoral Thesis Award from the Spanish Institution of Telecommunications Engineers, and a Research Fellowship of the Spanish Government to join ESA. He is also an Associate Editor of the IEEE Transactions on Information Theory, an Editor of the Foundations and Trends in Communications and Information Theory, Now Publishers and was an Editor of the IEEE Transactions on Wireless Communications.