Expurgated Bounds for the Asymmetric Broadcast Channel

Ran Averbuch, Nir Weinberger*, and Neri Merhav

December 10, 2017

The Andrew & Erna Viterbi Faculty of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 3200003, ISRAEL
nir.wein@gmail.com, {rans@campus, merhav@ee}.technion.ac.il

Abstract

This work contains two main contributions concerning the expurgation of hierarchical ensembles for the asymmetric broadcast channel. The first is an analysis of the optimal maximum likelihood (ML) decoders for the weak and strong user. Two different methods of code expurgation will be used, that will provide two competing error exponents. The second is the derivation of expurgated exponents under the generalized stochastic likelihood decoder (GLD). We prove that the GLD exponents are at least as tight as the maximum between the random coding error exponents derived in an earlier work by Averbuch and Merhav (2017) and one of our ML–based expurgated exponents. By that, we actually prove the existence of hierarchical codebooks that achieve the best of the random coding exponent and the expurgated exponent simultaneously for both users.

Index Terms: Asymmetric broadcast channel, error exponent, expurgated exponent, likelihood decoder.

* This work was done while N. Weinberger was at the Technion. Currently, he is with the School of Electrical Engineering at Tel–Aviv University.

1 Introduction

One of the most elementary system configuation models in multi-user information theory is the broadcast channel (BC). It has been introduced more than four decades ago by Cover [1], and since then, a vast amount of papers and books, analyzing different aspects of the broadcast model, have been published. Although the characterization of the capacity region of the general BC is still an open problem, some special cases have been solved. Most notably, the broadcast channel with degraded message sets, also known as the asymmetric broadcast channel (ABC), was introduced and solved by Körner and Marton [3].

The direct part of their coding theorem relies on Bergmans' scheme [2], which suggested the use of an hierarchical random code: first generate "cloud centers", which designate messages intended to both the receiver with the relatively high channel quality, henceforth referred to as the *strong user*, and the receiver with the relatively low channel quality, henceforth referred to as the *weak user*. Then, in the second step, "around" each cloud center, generate a "sattelite" codeword for each message that is intended to the strong user only. The transmitter sends a codeword pertaining to one of the clouds. The strong decoder fully decodes both the common message (cloud center) and his private message (sattelite), whereas the weak decoder decodes the common message only.

While the capacity region of the ABC has been known for many years, only little is known about its reliability functions. The earliest work on error exponents for the general ABC is of Körner and Sgarro [4]. Later, Kaspi and Merhav [5] have derived tighter lower bounds to the reliability functions of both users by analyzing random coding error exponents of their optimal decoders. Most recently [6], the exact random coding error exponents have been determined for both the strong user and the weak user, for the ensemble of fixed composition codes. Upper bounds to the reliability functions are discussed in [4], and are beyond the scope of this paper.

Even in the single–user case, it is known for many years that the random coding error exponent is not tight (with respect to the reliability function) for relatively low coding rates, and may be improved by expurgation [7], [15]. Specifically, improved bounds are obtained by eliminating codewords that contribute relatively highly to the error probability, and asserting that some upper bound holds for all remaining codewords. More recent papers, where the method of expurgation is studied are [8], [9] (list decoding), [10] (Gaussian BC) and [11] (discrete

memoryless multiple–access channels), among many others.

The main objective of this paper is to study expurgation techniques for the hierarchical ensemble used over the ABC. Expurgating a code for the ABC is not a trivial extension of expurgation in the single-user case, because there might be conflicting goals from the viewpoints of the two users. Nonetheless, we were able to define expurgation procedures that guarantee no harm to the performance of either user. This has paved the way to derive tighter lower bounds on the reliability functions of the ABC.

We start by analyzing the optimal maximum likelihood (ML) decoder, and derive some expurgated bounds, that are natural generalizations of the single–user expurgated bound due to Csiszár, Körner and Marton (CKM) [7]. Although our first process of code expurgation is fairly intuitive, there is at least one specific step in our first derivation where exponential tightness might be compromised. This point gives rise to a possible room for improvement upon the results of our first theorem, and indeed, such an improvement is achieved by a second method of expurgation. Here, one starts by expurgating cloud centers, and only afterwards, single codewords. The intuition behind this technique is the following. When the exponential rate of the codewords within a cloud is too high, the weak user can still make a good estimation, merely by relying on the set of cloud centers. The expurgated bounds of our second method, however, are not always tighter than those of the first method, because of other differences in their derivations.

We then expand the scope and consider the generalized likelihood decoder (GLD), which is a more general family of stochastic likelihood decoders. For such decoders, the probability of deciding on a given message is proportional to a general exponential function of the joint empirical distribution of the cloud–center, the codeword and the received channel output vector. The random coding error exponent of the ordinary and the mismatched likelihood decoders for the single–user channel have been derived by Scarlett *et al.* [12]. In a more recent paper by Merhav [13], the expurgated exponent of the GLD has been derived and compared to the classical expurgated bound of [7], showing an explicit improvement at relatively high coding rates. In this paper, we consider GLD's for both the strong and the weak users of an ABC, and derive expurgated exponents under these decoders. These bounds generalize the bound of [13], and prove that they are at least as tight as the maximum between the random coding error exponents of [6] and the expurgated bounds of our first theorem, which are based on the

ML decoder. By that, we actually prove the existence of hierarchical codebooks that attain the best of the random coding exponent and the expurgated exponent simultaneously for both users. The main drawback of those error exponents is that they are not easy to calculate since they involve minimizations over relatively cumbersome auxiliary channels, and hence, efficient computation algorithms for the GLD bound are left for further research. From this viewpoint, the exponents of our first theorems are much more attractive.

The remaining part of the paper is organized as follows. In Section 2, we establish notation conventions, and review some preliminaries. In Section 3, we formalize the model, the decoders and reliability functions. In Section 4, we summarize the main theoretical results of this paper. Sections 5, 6 and 7 provide the proofs of our main theorems.

2 Preliminaries

2.1 Notation Conventions

Throughout the paper, random variables will be denoted by capital letters, specific values they may take will be denoted by the corresponding lower case letters, and their alphabets will be denoted by calligraphic letters. Random vectors and their realizations will be denoted, respectively, by capital letters and the corresponding lower case letters, both in the bold face font. Their alphabets will be superscripted by their dimensions. For example, the random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$, ($n$ - positive integer) may take a specific vector value $\boldsymbol{x} = (x_1, \ldots, x_n)$ in $\mathcal{X}^n$, the n-th order Cartesian power of $\mathcal{X}$, which is the alphabet of each component of this vector. Sources and channels will be subscripted by the names of the relevant random variables/vectors and their conditionings, whenever applicable, following the standard notation conventions, e.g., Q_X, $Q_{Y|X}$, and so on. When there is no room for ambiguity, these subscripts will be omitted. For a generic joint distribution $Q_{XY} = \{Q_{XY}(x,y), x \in \mathcal{X}, y \in \mathcal{Y}\}$, which will sometimes be abbreviated by Q, information measures will be denoted in the conventional manner, but with a subscript Q, that is, $H_Q(X)$ is the marginal entropy of X, $H_Q(X|Y)$ is the conditional entropy of X given Y, $I_Q(X;Y) = H_Q(X) - H_Q(X|Y)$ is the mutual information between X and Y, and so on. The weighted divergence between two conditional distributions (channels), say, $Q_{Z|X}$

and $W = \{W(z|x), x \in \mathcal{X}, z \in \mathcal{Z}\}$, with weighting Q_X is defined as

$$D(Q_{Z|X}||W|Q_X) = \sum_{x\in\mathcal{X}} Q_X(x) \sum_{z\in\mathcal{Z}} Q_{Z|X}(z|x) \log \frac{Q_{Z|X}(z|x)}{W(z|x)}, \tag{1}$$

where logarithms, here and throughout the sequel, are taken to the natural base, and hence, all of the coding rates are given in *nats per channel use*. The probability of an event $\mathcal{E}$ will be denoted by $\Pr\{\mathcal{E}\}$, and the expectation operator with respect to (w.r.t.) a probability distribution P will be denoted by $\mathbb{E}_P\{\cdot\}$, where the subscript will often be omitted. For two positive sequences a_n and b_n, the notation $a_n \doteq b_n$ will stand for equality in the exponential scale, that is, $\lim_{n\to\infty} \frac{1}{n} \log \frac{a_n}{b_n} = 0$. The indicator function of an event $\mathcal{E}$ will be denoted by $\mathcal{I}\{\mathcal{E}\}$. The notation $[x]_+$ will stand for $\max\{0, x\}$.

The empirical distribution of a sequence $\boldsymbol{x} \in \mathcal{X}^n$, which will be denoted by $\hat{P}_{\boldsymbol{x}}$, is the vector of relative frequencies, $\hat{P}_{\boldsymbol{x}}(x)$, of each symbol $x \in \mathcal{X}$ in $\boldsymbol{x}$. The type class of $\boldsymbol{x} \in \mathcal{X}^n$, denoted $\mathcal{T}(\boldsymbol{x})$, is the set of all vectors $\boldsymbol{x}'$ with $\hat{P}_{\boldsymbol{x}'} = \hat{P}_{\boldsymbol{x}}$. When we wish to emphasize the dependence of the type class on the empirical distribution $\hat{P}$, we will denote it by $\mathcal{T}(\hat{P})$. Information measures associated with empirical distributions will be denoted with 'hats' and will be subscripted by the sequences from which they are induced. For example, the entropy associated with $\hat{P}_{\boldsymbol{x}}$, which is the empirical entropy of $\boldsymbol{x}$, will be denoted by $\hat{H}_{\boldsymbol{x}}(X)$. Similar conventions will apply to the joint empirical distribution, the joint type class, the conditional empirical distributions and the conditional type classes associated with pairs (and multiples) of sequences of length n. Accordingly, $\hat{P}_{\boldsymbol{xy}}$ would be the joint empirical distribution of $(\boldsymbol{x}, \boldsymbol{y}) = \{(x_i, y_i)\}_{i=1}^n$, $\mathcal{T}(\hat{P}_{\boldsymbol{xy}})$ will denote the joint type class of $(\boldsymbol{x}, \boldsymbol{y})$, $\mathcal{T}(\boldsymbol{x}|\boldsymbol{y})$ will stand for the conditional type class of $\boldsymbol{x}$ given $\boldsymbol{y}$, $\hat{I}_{\boldsymbol{xy}}(X;Y)$ will denote the empirical mutual information, and so on. When we wish to emphasize the dependence of $\mathcal{T}(\boldsymbol{x}|\boldsymbol{y})$ upon $\boldsymbol{y}$ and the relevant empirical conditional distribution, $Q_{X|Y} = \hat{P}_{\boldsymbol{x}|\boldsymbol{y}}$, we denote it by $\mathcal{T}(Q_{X|Y}|\boldsymbol{y})$. Similar conventions will apply to triples of sequences, say, $\{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\}$, etc. Likewise, when we wish to emphasize the dependence of empirical information measures upon a given empirical distribution, Q, we denote them using the subscript Q, as described above.

2.2 Mathematical Tools

Along the proofs in the current paper, some mathematical results are used extensively. Instead of explaining them each time repeatedly, let us summerize them:

• We abbreviate the union bounds and Markov's inequality by UB and MI, respectively.

• We refer to the following inequality as the *power distribution inequality* (PD),

$$\left(\sum_{j\in\mathcal{J}} a_j\right)^s \leq \sum_{j\in\mathcal{J}} a_j^s \tag{2}$$

which holds whenever $s \in [0,1]$ and $a_j \geq 0$ for all $j \in \mathcal{J}$ [15, Exercise 4.15(f)]. A special case occurs when the cardinality of $\mathcal{J}$ is subexponential in the blocklength n (e.g., when we sum over type–classes) and then

$$\left(\sum_{j\in\mathcal{J}} e^{na_j}\right)^s \doteq \left(\max_{j\in\mathcal{J}} e^{na_j}\right)^s = \max_{j\in\mathcal{J}} e^{nsa_j} \doteq \sum_{j\in\mathcal{J}} e^{nsa_j}. \tag{3}$$

• Let N be a binomial random variable with e^{nR} $(R \geq 0)$ trials and success rate of the exponential order of e^{-nI} $(I \geq 0)$. It is shown in [16, Chap. 6.3] that for $s > 0$

$$\mathbb{E}\left\{N^s\right\} \doteq \begin{cases} \exp\left\{n[R-I]s\right\} & R \geq I \\ \exp\left\{n[R-I]\right\} & R < I. \end{cases} \tag{4}$$

3 Definitions and Problem Formulation

We consider a memoryless ABC with a finite input alphabet $\mathcal{X}$ and finite output alphabets $\mathcal{Y}$ and $\mathcal{Z}$. Let $W_{Y|X} \equiv W_1 = \{W_1(y|x),\ x \in \mathcal{X},\ y \in \mathcal{Y}\}$ and $W_{Z|X} \equiv W_2 = \{W_2(z|x),\ x \in \mathcal{X},\ z \in \mathcal{Z}\}$ denote the single–letter input–output transition probability matrices, associated with the strong user and the weak user, respectively. When these channels are fed by an input vector $\boldsymbol{x} \in \mathcal{X}^n$, they produce the corresponding output vectors $\boldsymbol{y} \in \mathcal{Y}^n$ and $\boldsymbol{z} \in \mathcal{Z}^n$, according to

$$W_1(\boldsymbol{y}|\boldsymbol{x}) = \prod_{t=1}^{n} W_1(y_t|x_t), \tag{5}$$

$$W_2(\boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^{n} W_2(z_t|x_t). \tag{6}$$

We are interested in sending one out of $M_y M_z$ messages to the strong user, that observes $\boldsymbol{y}$, and one out of M_z messages to the weak user, that observes $\boldsymbol{z}$. The two messages are chosen under the uniform distribution. Although our results prove the existence of a single sequence of *deterministic* hierarchical constant composition (HCC) codebooks, whose error probabilities are provably bounded, our proof techniques use extensively the following mechanism of random selection of an HCC code for the ABC. Let $\mathcal{U}$ be a finite alphabet, let P_U be a given

probability distribution on $\mathcal{U}$, and let $P_{X|U}$ be a given matrix of conditional probabilities of X given U, such that the type–class $\mathcal{T}(P_U)$ and the conditional type–class $\mathcal{T}(P_{X|U}|\boldsymbol{u})$ are non–empty. We first select, independently at random, $M_z = \lceil e^{nR_z} \rceil$ n-vectors ("cloud centers"), $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{M_z-1}$, all under the uniform distribution over the type–class $\mathcal{T}(P_U)$. Next, for each $m = 0, 1, \ldots, M_z-1$, we select conditionally independently (given $\boldsymbol{u}_m$), $M_y = \lceil e^{nR_y} \rceil$ codewords, $\boldsymbol{x}_{m,0}, \boldsymbol{x}_{m,1}, \ldots, \boldsymbol{x}_{m,(M_y-1)}$, under the uniform distribution across the conditional type–class $\mathcal{T}(P_{X|U}|\boldsymbol{u}_m)$. We denote the sub–code for each cloud by $\mathcal{C}_m(n) = \{\boldsymbol{x}_{m,0}, \boldsymbol{x}_{m,1}, \ldots, \boldsymbol{x}_{m,(M_y-1)}\}$. Thus, the communication rate to the weak user is R_z, while the total communication rate to the strong user is $R_z + R_y$. Once selected, the entire codebook $\mathcal{C}(n) = \cup_{m=0}^{M_z-1} \mathcal{C}_m(n)$, and the collection of cloud centers, $\{\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{M_z-1}\}$, are revealed to the encoder and to both decoders. We usually omit the dependence on n from the notation of the code, and use $\mathcal{C}$ and $\mathcal{C}_m$, for short. We denote by $\mathscr{C}$ a sequence of HCC codes, $\{\mathcal{C}(n),\ n = 1, 2, \ldots\}$.

For any of the following described decoding rules, denote by $[\hat{m}(\boldsymbol{y}), \hat{i}(\boldsymbol{y})]$ the decoded pair of the strong user, and by $\tilde{m}(\boldsymbol{z})$ the decoded cloud of the weak user. The ML decoder for the strong user is given by

$$[\hat{m}(\boldsymbol{y}), \hat{i}(\boldsymbol{y})] = \mathop{\arg\max}_{0 \le m \le M_z-1, 0 \le i \le M_y-1} W_1(\boldsymbol{y}|\boldsymbol{x}_{mi}), \tag{7}$$

and the optimal ML decoder for the weak user (the bin index decoder) is given by

$$\tilde{m}(\boldsymbol{z}) = \mathop{\arg\max}_{0 \le m \le M_z-1} W_2(\boldsymbol{z}|\mathcal{C}_m), \tag{8}$$

where

$$W_2(\boldsymbol{z}|\mathcal{C}_m) \stackrel{\Delta}{=} \frac{1}{M_y} \sum_{\boldsymbol{x} \in \mathcal{C}_m} W_2(\boldsymbol{z}|\boldsymbol{x}) = \frac{1}{M_y} \sum_{i=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{mi}). \tag{9}$$

The likelihood decoder is a stochastic decoder, that chooses the decoded message according to the posterior probability mass function, induced by the channel output (either $\boldsymbol{y}$ or $\boldsymbol{z}$). For the strong user, the ordinary likelihood decoder randomly selects the estimated message $(\hat{m}, \hat{i})$ according to the following posterior distribution

$$\Pr\left\{\hat{m} = m, \hat{i} = i \Big| \boldsymbol{y}\right\} = \frac{W_1(\boldsymbol{y}|\boldsymbol{x}_{mi})}{\sum_{m'=0}^{M_z-1} \sum_{i'=0}^{M_y-1} W_1(\boldsymbol{y}|\boldsymbol{x}_{m'i'})}. \tag{10}$$

The generalized likelihood decoder (GLD) for the strong user is defined by

$$\Pr\left\{\hat{m} = m, \hat{i} = i \Big| \boldsymbol{y}\right\} = \frac{\exp\{n g_{\mathrm{s}}(\hat{P}_{\boldsymbol{u}_m \boldsymbol{x}_{mi} \boldsymbol{y}})\}}{\sum_{m'=0}^{M_z-1} \sum_{i'=0}^{M_y-1} \exp\{n g_{\mathrm{s}}(\hat{P}_{\boldsymbol{u}_{m'} \boldsymbol{x}_{m'i'} \boldsymbol{y}})\}}, \tag{11}$$

where $\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{y}}$ is the empirical distribution of $(\boldsymbol{u}_m, \boldsymbol{x}_{mi}, \boldsymbol{y})$, and $g_{\rm s}(\cdot)$ is a given continuous, real valued functional of this empirical distribution. In the same manner, the ordinary likelihood decoder for the weak user randomly selects the estimated cloud $\tilde{m}$ according to

$$\Pr\{\tilde{m}=m|\boldsymbol{z}\}=\frac{\sum_{i=0}^{M_y-1}W_2(\boldsymbol{z}|\boldsymbol{x}_{mi})}{\sum_{m'=0}^{M_z-1}\sum_{i'=0}^{M_y-1}W_2(\boldsymbol{z}|\boldsymbol{x}_{m'i'})}, \tag{12}$$

while the GLD for the weak user is defined by

$$\Pr\{\tilde{m}=m|\boldsymbol{z}\}=\frac{\sum_{i=0}^{M_y-1}\exp\{ng_{\rm w}(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{z}})\}}{\sum_{m'=0}^{M_z-1}\sum_{i'=0}^{M_y-1}\exp\{ng_{\rm w}(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'i'}\boldsymbol{z}})\}}. \tag{13}$$

Exactly as the universal decoders derived in [6], generalized decoders may also depend on the cloud–centers, which may be helpful, since all of the codewords in each sub–code are highly correlated via their cloud–center. One of the most important properties of the GLD is the following. The union bound, which is used in the first steps of the derivations for both users, actually provides an exact expression for the probability of error, unlike in the analyses of the ML decoders, where the union bound harms the exponential tightness, at least for relatively high rates.

The generalized likelihood decoders cover several important special cases for the strong user. The choice

$$g_{\rm s}(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{y}})=\sum_{u,x,y}\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{y}}(u,x,y)\log W_1(y|x)=\sum_{x,y}\hat{P}_{\boldsymbol{x}_{mi}\boldsymbol{y}}(x,y)\log W_1(y|x) \tag{14}$$

corresponds to the ordinary likelihood decoder. More generally, one may introduce an "inverse temperature" parameter $\beta\geq 0$ and define

$$g_{\rm s}(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{y}})=\beta\sum_{x,y}\hat{P}_{\boldsymbol{x}_{mi}\boldsymbol{y}}(x,y)\log W_1(y|x). \tag{15}$$

Here, β controls the degree of skewedness of the distribution (11): while $\beta=1$ corresponds to the ordinary likelihood decoder, $\beta\to\infty$ leads to the deterministic ML decoder. In the same manner,

$$g_{\rm s}(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{y}})=\beta\sum_{x,y}\hat{P}_{\boldsymbol{x}_{mi}\boldsymbol{y}}(x,y)\log W'(y|x), \tag{16}$$

for W' being different from W_1, defines a family of mismatched likelihood decoders. Yet another interesting choice is

$$g_{\rm s}(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{y}})=\beta\hat{I}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{y}}(UX;Y), \tag{17}$$

which is a parametric family of stochastic mutual information decoders, where the limit of $\beta \to \infty$ yields the ordinary maximum mutual information (MMI) universal decoder [14]. Similarly for the weak user, the choice

$$g_{\mathrm{w}}(\hat{P}_{\boldsymbol{u}_m \boldsymbol{x}_{mi} \boldsymbol{z}}) = \sum_{x,z} \hat{P}_{\boldsymbol{x}_{mi} \boldsymbol{z}}(x,z) \log W_2(z|x) \tag{18}$$

corresponds to the ordinary likelihood decoder for the weak user. From now on, we will assume that $g_{\mathrm{s}}(\cdot)$ and $g_{\mathrm{w}}(\cdot)$ are the same functional, and denote both of them by $g(\cdot)$.

Let $\boldsymbol{Y} \in \mathcal{Y}^n$ and $\boldsymbol{Z} \in \mathcal{Z}^n$ be the random channel outputs resulting from the transmission of $\boldsymbol{x}_{mi}$. For a given code $\mathcal{C}(n)$, define the error probabilities as

$$P_{\mathrm{e}|mi}(\mathcal{C}(n)) = \Pr\left\{ [\hat{m}(\boldsymbol{Y}), \hat{i}(\boldsymbol{Y})] \neq (m,i) \Big| \boldsymbol{x}_{mi} \text{ sent} \right\}, \tag{19}$$

and

$$P_{\mathrm{e}|m}(\mathcal{C}(n)) = \frac{1}{M_y} \sum_{i=0}^{M_y - 1} \Pr\left\{ \tilde{m}(\boldsymbol{Z}) \neq m | \boldsymbol{x}_{mi} \text{ sent} \right\}, \tag{20}$$

where in both definitions, $\Pr\{\cdot\}$ designates the probability measure associated with the randomness of the channel outputs given its input, and the randomness of the stochastic decoders. Moreover, the error probabilities are defined to be zero whenever the blocklength is such that no code can be generated. Our main objective is to prove the existence of sequences of HCC codes $\mathscr{C} = \{\mathcal{C}(n)\}_{n=1}^{\infty}$ and obtain the tightest possible single–letter expressions that lower bound the following limits

$$E_{\mathrm{su}}(\mathscr{C}) = \liminf_{n \to \infty} \left[-\frac{\log \max_{m,i} P_{\mathrm{e}|mi}(\mathcal{C}(n))}{n} \right], \tag{21}$$

and

$$E_{\mathrm{wu}}(\mathscr{C}) = \liminf_{n \to \infty} \left[-\frac{\log \max_{m} P_{\mathrm{e}|m}(\mathcal{C}(n))}{n} \right], \tag{22}$$

both for the ML decoder and the GLD.

As already mentioned in the Introduction, the exact random coding error exponents for both users of the ABC have been derived and analyzed in [6]. For the weak user, it is given by

$$E_{\mathrm{wu}}^{\mathrm{RC}}(R_y, R_z) = \min_{Q_{Z|UX}} \left\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + \Big[I_Q(U;Z) + [I_Q(X;Z|U) - R_y]_+ - R_z \Big]_+ \right\}, \tag{23}$$

while the exact error exponent of the strong user is given by

$$E_{\rm su}^{\rm RC}(R_y,R_z)=\min_{Q_{Y|UX}}\Bigg\{D(Q_{Y|UX}\|W_{Y|X}|Q_{UX}) \\ +\min\Big\{\big[I_Q(UX;Y)-R_y-R_z\big]_+,\big[I_Q(X;Y|U)-R_y\big]_+\Big\}\Bigg\}. \quad (24)$$

We may expect to improve these error exponents, at least when one of the coding rates is low, by code expurgation. In this paper, we derive expurgated exponents for the ABC under ML decoding in two different methods. In addition, we discuss the GLD, that enables us to achieve the best between the random coding bound and one of the ML–based expurgated bounds.

4 Main Results

For maximum likelihood decoding, we distinguish between two different methods of expurgation for the HCC ensemble. The first methodology is based on the following technique of expurgation: we randomly draw a HCC codebook, and then simultaneously expurgate both bad clouds and bad codewords within the remaining clouds. The resulting expurgated bounds are given in Theorem 1, which is proved in Section 5. A full discussion of the first two theorems follows the second theorem. In order to state our first theorem, we start with the following definitions. We define the following sets of distributions

$$\mathcal{S}\stackrel{\Delta}{=}\big\{Q_{UXX'}:\ Q_{UX'}=Q_{UX}=P_{UX}\big\}, \quad (25)$$

$$\mathcal{P}\stackrel{\Delta}{=}\big\{Q_{UU'XX'}:\ Q_{U'X'}=Q_{UX}=P_{UX}\big\}. \quad (26)$$

For a general channel W, we define the averaged Chernoff distance function by

$$D_s(Q_{XX'})\stackrel{\Delta}{=}-\sum_{(x,x')\in\mathcal{X}^2}Q_{XX'}(x,x')\log\left[\sum_{y\in\mathcal{Y}}W^{1-s}(y|x)\cdot W^s(y|x')\right], \quad (27)$$

where in the following, we choose $W=W_1$ or $W=W_2$, depending on the user we are relating to. For the weak user, define an error exponent function as

$$E_{\rm wu}^{\rm ML1}(R_y,R_z)\stackrel{\Delta}{=}\max_{0\le t\le 1}\min_{\substack{Q_{UU'XX'}\in\mathcal{P}\\ I_Q(UX;U'X')\le 2R_y+R_z}}\big[I_Q(UX;U'X')+D_t(Q_{XX'})\big]-R_y-R_z. \quad (28)$$

Next, for the strong user we define the following error exponent functions

$$E_{\text{su-1}}^{\text{ML1}}(R_y,s) \stackrel{\Delta}{=} \min_{\substack{Q_{UXX'}\in\mathcal{S}\\ I_Q(X;X'|U)\le R_y}} \big[I_Q(X;X'|U)+D_s(Q_{XX'})\big]-R_y, \tag{29}$$

$$E_{\text{su-2}}^{\text{ML1}}(R_y,R_z,s) \stackrel{\Delta}{=} \min_{\substack{Q_{UU'XX'}\in\mathcal{P}\\ I_Q(UX;U'X')\le R_y+R_z}} \big[I_Q(UX;U'X')+D_s(Q_{XX'})\big]-R_y-R_z, \tag{30}$$

$$E_{\text{su}}^{\text{ML1}}(R_y,R_z) \stackrel{\Delta}{=} \max_{0\le s\le 1} \min\left\{E_{\text{su-1}}^{\text{ML1}}(R_y,s), E_{\text{su-2}}^{\text{ML1}}(R_y,R_z,s)\right\}. \tag{31}$$

Theorem 1. There exists a sequence of HCC codes, $\mathscr{C}=\{\mathcal{C}(n),\ n=1,2,\dots\}$, with a rate pair (R_y,R_z) for which both

$$E_{\text{su}}(\mathscr{C}) \ge E_{\text{su}}^{\text{ML1}}(R_y,R_z), \tag{32}$$

$$E_{\text{wu}}(\mathscr{C}) \ge E_{\text{wu}}^{\text{ML1}}(R_y,R_z). \tag{33}$$

The second method is somewhat different, and the idea behind it is the following. At the first step, we expurgate sub–codes, merely according to their cloud–centers. Then, at the second step, we fix the set of cloud–centers of the remaining clouds from the first step, and then expurgate specific codewords, as well as clouds, according to some collective behavior of their codewords. The resulting expurgated bounds are given in Theorem 2, and as can be seen below, the expressions are more complicated than those of Theorem 1, at least for the weak user. The proof can be found in Section 6. In order to state our second theorem, we need a few definitions.

For a given marginal Q_{UZ}, let $\mathcal{S}(Q_{UZ})$ denote the set of conditional distributions $\{Q_{X|UZ}\}$ such that $\sum_z Q_{UZ}(u,z)Q_{X|UZ}(x|u,z)=P_{UX}(u,x)$ for every $(u,x)\in\mathcal{U}\times\mathcal{X}$, where $P_{UX}=P_U\times P_{X|U}$. We denote $\bar{t}=1-t$. For the weak user, define

$$\begin{aligned}&\hat{D}_t(R_y,Q_{UU'}) \stackrel{\Delta}{=}\\ &\min_{Q_{Z|UU'}}\min_{Q_{X|UZ}\in\mathcal{S}(Q_{UZ})}\min_{Q_{X'|U'Z}\in\mathcal{S}(Q_{U'Z})}\Big\{\bar{t}\cdot D(Q_{Z|UX}\|W_{Z|X}|Q_{UX})+t\cdot D(Q_{Z|U'X'}\|W_{Z|X'}|Q_{U'X'})\\ &+\bar{t}\cdot I_Q(Z;U'|U)+t\cdot I_Q(Z;U|U')+t\cdot[I_Q(X;Z|U)-R_y]_+ +\bar{t}\cdot[I_Q(X';Z|U')-R_y]_+\Big\}.\end{aligned} \tag{34}$$

We define the following set of distributions

$$\mathcal{Q} \stackrel{\Delta}{=} \big\{Q_{UU'}:\ Q_U=Q_{U'}=P_U\big\}, \tag{35}$$

and an error exponent function

$$E_{\text{wu}}^{\text{ML2}}(R_y, R_z) \stackrel{\Delta}{=} \max_{0\le t\le 1} \min_{\substack{Q_{UU'}\in\mathcal{Q} \\ I_Q(U;U')\le R_z}} \left[I_Q(U;U') + \hat{D}_t(R_y, Q_{UU'})\right] - R_z. \tag{36}$$

Next, for the strong user we define the following error exponent functions

$$E_{\text{su-1}}^{\text{ML2}}(R_y, s) \stackrel{\Delta}{=} \min_{\substack{Q_{UXX'}\in\mathcal{S} \\ I_Q(X;X'|U)\le R_y}} \left[I_Q(X;X'|U) + D_s(Q_{XX'})\right] - R_y, \tag{37}$$

$$E_{\text{su-2}}^{\text{ML2}}(R_y, R_z, s) \stackrel{\Delta}{=} \min_{\substack{Q_{UU'XX'}\in\mathcal{P} \\ I_Q(U;U')\le R_z}} \left[I_Q(UX;U'X') + D_s(Q_{XX'})\right] - R_y - R_z, \tag{38}$$

$$E_{\text{su}}^{\text{ML2}}(R_y, R_z) \stackrel{\Delta}{=} \max_{0\le s\le 1} \min\left\{E_{\text{su-1}}^{\text{ML2}}(R_y, s), E_{\text{su-2}}^{\text{ML2}}(R_y, R_z, s)\right\}. \tag{39}$$

Theorem 2. There exists a sequence of HCC codes, $\mathscr{C} = \{\mathcal{C}(n),\ n = 1, 2, \dots\}$, with a rate pair (R_y, R_z) for which both

$$E_{\text{su}}(\mathscr{C}) \ge E_{\text{su}}^{\text{ML2}}(R_y, R_z), \tag{40}$$

$$E_{\text{wu}}(\mathscr{C}) \ge E_{\text{wu}}^{\text{ML2}}(R_y, R_z). \tag{41}$$

Discussion

First, all of the expressions in Theorems 1 and 2 generalize the well–known CKM expurgated bound [7]. For example, it can be easily recovered from the expression $E_{\text{wu}}^{\text{ML1}}(R_y, R_z)$ of Theorem 1, when degenerating the hierarchical codebook by choosing $R_y = 0$, as well as $P_{X|U}(x|u) = \delta(x - u)$ (assuming that $\mathcal{X} = \mathcal{U}$), in order to get back to the CKM expurgated bound.

Concerning the strong user, each bound is given by the minimum between two different expressions. The first expression is related to error events within the cloud of the true codeword. In fact, we have that $E_{\text{su-1}}^{\text{ML1}}(R_y, s) = E_{\text{su-1}}^{\text{ML2}}(R_y, s)$, where the difference is given by the second components, $E_{\text{su-2}}^{\text{ML1}}(R_y, R_z, s)$ and $E_{\text{su-2}}^{\text{ML2}}(R_y, R_z, s)$, for which the method of expurgation is relevant and cause a change in the final expressions. Although the objectives in (30) and (38) are exactly the same, the constraints are different, and are not subsets of each other.

Concerning the weak user, the situation is much more complicated, because of the structure of the optimal decoder. The derivation in the proof of Theorem 1 contains a passage [(68) to (69)] that may harm the exponential tightness of the result. Specifically, we use the PD

inequality over the sums that stems from the definition of the optimal decoder, i.e.,

$$\left\{\frac{1}{M_y}\sum_{i=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{mi})\right\}^{1-t}\left\{\frac{1}{M_y}\sum_{j=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{m'j})\right\}^{t} \tag{42}$$

$$\leq \frac{1}{M_y}\sum_{i=0}^{M_y-1}\sum_{j=0}^{M_y-1} W_2^{1-t}(\boldsymbol{z}|\boldsymbol{x}_{mi})\cdot W_2^{t}(\boldsymbol{z}|\boldsymbol{x}_{m'j}), \tag{43}$$

and therefore, the resulting bound of Theorem 1 is, in fact, a natural generalization of the classical single–user expurgated bound, which only depends on the Chernoff distance between pairs of codewords. Because of this passage, the bound of Theorem 1 is inferior to the bound of Theorem 2, at relatively high values of R_y. However, the resulting exponent of Theorem 1 still outperforms the result of Theorem 2, at least for relatively low R_y values (see Fig. 1). The reason for the bound of Theorem 2 to be inferior at relatively low rates is because of the remaining parts of the two proofs. While both proofs use Markov's inequality in order to show the existence of good codebooks, some logical arguments that can only be claimed in the proof of Theorem 1 provide tighter upper bounds. More specifically, the proof of Theorem 1 relies merely on type–class enumerators, which takes only integer values. It is shown that there exists codebooks, for which those enumerators must equal to zero in some range of relatively low rates. On the other hand, the proof of Theorem 2 relies on some more complicated quantities, that are not necessarily integer–valued, and hence, they cannot be assured to be equal to zero at any range of rates. Since the derivation in the proof of Theorem 2 is exponentially tight after the first two steps, and does not compromise on the optimal decoders (as Theorem 1 does in the passage we mentioned above), it provides a better result at relatively high R_y. Specifically, the expression given in Theorem 2 reaches a plateau at high R_y, while the expression of Theorem 1 reaches zero. In this regime, there is no loss in the exponent of the weak user if its decoder treats the satellites codewords as noise. In this event, the satellite-rate is immaterial, and the exponent of the weak user only depends on R_z. One should note that the improvement at high rates is obtained by expressions which are more complicated to compute.

We next provide some numerical results, comparing our expurgated bounds for the weak user (Fig. 1) and for the strong user (Fig. 2), as given by Theorems 1 and 2. Let W_1 and W_2 be two binary symmetric channels (BSCs) with crossover parameters $p_y = 0.0005$ and $p_z = 0.001$, respectively. Let $\mathcal{U}$ be binary as well and let P_U be uniformly distributed over $\{0, 1\}$. Also, let $P_{X|U}$ be a BSC with crossover parameter $p_{x|u} = 0.15$.

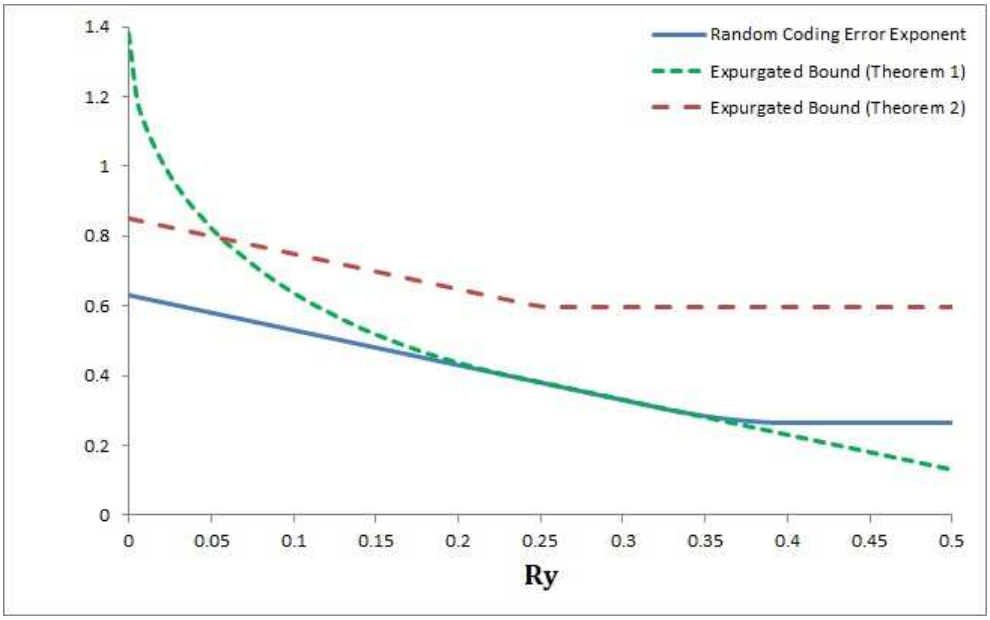

Figure 1: Expurgated bounds for the weak user ($R_z = 0$).

In Fig. 2, the orange (dot-dashed) curve describes the expression of $E_{\text{su-1}}^{\text{ML1}}(R_y, s)$, which is common in both exponents. As can be seen, each of the two exponents is dominated by this expression at relatively high sattelite rates. Note that at least for this specific example, $E_{\text{su}}^{\text{ML1}}(R_y, R_z)$ is higher than $E_{\text{su}}^{\text{ML2}}(R_y, R_z)$ at any pair of coding rates.

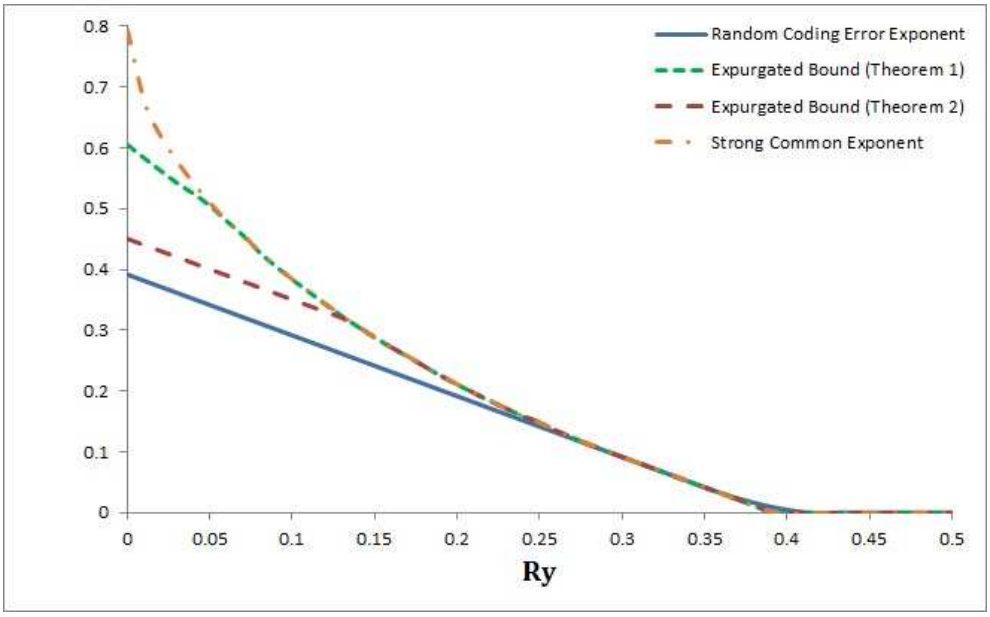

Figure 2: Expurgated bounds for the strong user ($R_z = 0.2$).

We now move on to discuss the GLD. As was already mentioned earlier, the GLD enables us to make a tighter derivation for the probability of error, and therefore, the resulting expurgated bounds are strictly tighter, at least at relatively high rates. The drawback of the expressions of Theorem 3 is that they are quite cumbersome, at least when compared to those of Theorems 1 or 2. In order to characterize the expurgated bounds of the GLD, we define first a few quantities. Let

$$\phi(R_y, Q_{UY}) = \max_{\{Q_{X|UY}:\ I_Q(X;Y|U)\leq R_y\}} [g(Q) - I_Q(X;Y|U)] + R_y, \tag{44}$$

and

$$\psi(R_y, R_z, Q_Y) = \max_{\{Q_{UX|Y}:\ I_Q(U;Y)\leq R_z,\ I_Q(UX;Y)\leq R_z+R_y\}} [g(Q) - I_Q(UX;Y)] + R_z + R_y. \tag{45}$$

Also, define

$$\Upsilon(Q_{UXX'}, R_y, R_z) = \min_{Q_{Y|UXX'}} \Bigg(D(Q_{Y|UX}\|W_{Y|X}|Q_{UX}) + I_Q(X';Y|UX) + \Big[\max\big\{g(Q_{UXY}), \phi(R_y, Q_{UY}), \psi(R_y, R_z, Q_Y)\big\} - g(Q_{UX'Y})\Big]_+\Bigg), \tag{46}$$

and

$$\Omega(Q_{UU'XX'}, R_y, R_z) = \min_{Q_{Y|UU'XX'}} \Bigg(D(Q_{Y|UX}\|W_{Y|X}|Q_{UX}) + I_Q(U'X';Y|UX) + \Big[\max\big\{g(Q_{UXY}), \phi(R_y, Q_{UY}), \psi(R_y, R_z, Q_Y)\big\} - g(Q_{U'X'Y})\Big]_+\Bigg). \tag{47}$$

We define the following error exponent functions. For the weak user,

$$E_{\text{wu}}^{\text{GLD}}(R_y, R_z) \stackrel{\Delta}{=} \min_{\substack{Q_{UU'XX'}\in\mathcal{P} \\ I_Q(UX;U'X')<2R_y+R_z, \\ I_Q(U;U')<R_z}} [I_Q(UX;U'X') + \Omega(Q, R_y, R_z)] - R_y - R_z, \tag{48}$$

and for the strong user

$$E_{\text{su-1}}^{\text{GLD}}(R_y, R_z) \stackrel{\Delta}{=} \min_{\substack{Q_{UXX'}\in\mathcal{S} \\ I_Q(X;X'|U)<R_y}} [I_Q(X;X'|U) + \Upsilon(Q, R_y, R_z)] - R_y, \tag{49}$$

$$E_{\text{su-2}}^{\text{GLD}}(R_y, R_z) \stackrel{\Delta}{=} \min_{\substack{Q_{UU'XX'}\in\mathcal{P} \\ I_Q(UX;U'X')<R_y+R_z, \\ I_Q(UX;U')<R_z}} [I_Q(UX;U'X') + \Omega(Q, R_y, R_z)] - R_y - R_z, \tag{50}$$

$$E_{\text{su}}^{\text{GLD}}(R_y, R_z) \stackrel{\Delta}{=} \min\left\{E_{\text{su-1}}^{\text{GLD}}(R_y, R_z), E_{\text{su-2}}^{\text{GLD}}(R_y, R_z)\right\}. \tag{51}$$

Theorem 3. There exists a sequence of HCC codes, $\mathscr{C} = \{\mathcal{C}(n),\ n = 1, 2, \dots\}$, with a rate pair (R_y, R_z) for which both

$$E_{\rm su}(\mathscr{C}) \geq E_{\rm su}^{\rm GLD}(R_y, R_z), \tag{52}$$

$$E_{\rm wu}(\mathscr{C}) \geq E_{\rm wu}^{\rm GLD}(R_y, R_z). \tag{53}$$

Discussion

• An expurgated bound for the GLD in the single–user regime has been derived by Merhav [13]. Theorem 3 generalizes the analysis of the GLD from the single-user case considered in [13] to the asymmetric broadcast channel studied here. While this generalization is not trivial, there are some parts in the proof which can be easily deduced from the single-user case. Those parts were omitted for the sake of brevity, but can be found in [18], which is an extended version of the current manuscript, that contains all the full proofs.

• The expression of eq. (48) has the same structure as the bound given in Theorem 1, except that here the functional $\Omega(Q, R_y, R_z)$ replaces the expected Chernoff distance, and an additional constraint ($I_Q(U;U') < R_z$) has been added. We prove in Appendix A that at least for the choice $g(Q) = \mathbb{E}_Q \log W_2(Z|X)$, the expurgated bound of Theorem 3, $E_{\rm wu}^{\rm GLD}(R_y, R_z)$, is at least as tight as the bound of Theorem 1, $E_{\rm wu}^{\rm ML1}(R_y, R_z)$.

• One of the main advantages of the GLD, is the fact that the derivation of its probability of error may be exponentially tighter than the derivations in the proofs of Theorems 1 or 2, because its circumvents the use of the union bound in the proofs of Theorems 1 and 2 [eqs. (55)–(56)]. As a consequence, we show in Appendix B that $E_{\rm wu}^{\rm GLD}(R_y, R_z)$ cannot be smaller than the random coding error exponent of the weak user at any pair of rates, by examining the former for the suboptimal GLD based on the universal metric $g(Q) = I_Q(UX; Z)$. We conclude that $E_{\rm wu}^{\rm GLD}(R_y, R_z)$ is at least as tight as the maximum between $E_{\rm wu}^{\rm ML1}(R_y, R_z)$ and the random coding error exponent, $E_{\rm wu}^{\rm RC}(R_y, R_z)$. This generalizes the result of [13] from the single-user channel (as was obtained by [19] using different arguments) to the ABC.

• The same can be proved for the strong user, i.e., that $E_{\rm su}^{\rm GLD}(R_y, R_z)$ is at least as tight as the maximum between $E_{\rm su}^{\rm ML1}(R_y, R_z)$ and the random coding error exponent, $E_{\rm su}^{\rm RC}(R_y, R_z)$. We conclude, that there exist a HCC codebook, for which one user works in the "expurgated region" (slope greater than 1), while the other user works in the "random coding region" (slope smaller than 1). For example, it may be the case when the channel to the strong user is quite

clean, while the channel to the weak user is very noisy, compared to the required rates.

• Keeping in mind the discussion after Theorem 2, it is now clear that the expurgated bound for the weak user of Theorem 3, $E_{\rm wu}^{\rm GLD}(R_y,R_z)$, is strictly tighter than the bound $E_{\rm wu}^{\rm ML2}(R_y,R_z)$ of Theorem 2, at relatively low rates, and is strictly tighter than the bound $E_{\rm wu}^{\rm ML1}(R_y,R_z)$ of Theorem 1, at relatively high rates. However, we were not able to determine whether $E_{\rm wu}^{\rm GLD}(R_y,R_z)$ is at least as tight as $E_{\rm wu}^{\rm ML2}(R_y,R_z)$ of Theorem 2, at relatively high rates. In other words, it is not clear whether the bound of Theorem 3 is at least as tight as the maximum between the bounds of the first two theorems, although we conjecture that it is indeed the case when choosing one of the decoding metrics $g(Q)=\beta\mathbb{E}_Q\log W_2(Z|X)$ or $g(Q)=\beta I_Q(UX;Z)$, and letting $\beta\to\infty$.

5 Proof of Theorem 1

The proof has three main parts: In Section 5.1, we upper bound the error probabilities and express each one of them using suitable type-class enumerators. In Section 5.2, which is the main step of the proof, we show the existence of an hierarchical codebook for which these enumerators are upper bounded by specific deterministic functions of information measures. In Section 5.3, we put back these deterministic bounds in order to get the desired results.

5.1 Upper Bounding Error Probabilities Using Type–Enumerators

We define the Chernoff distance between $\boldsymbol{x}$ and $\boldsymbol{x}'$ by ($W=W_1$ or $W=W_2$)

$$d_s(\boldsymbol{x},\boldsymbol{x}')\stackrel{\Delta}{=}-\log\left[\sum_{\boldsymbol{y}}W^{1-s}(\boldsymbol{y}|\boldsymbol{x})\cdot W^s(\boldsymbol{y}|\boldsymbol{x}')\right]. \tag{54}$$

For the strong user, we have the following upper bound for any $s\in[0,1]$

$$P_{{\rm e}|mi}(\mathcal{C})=\sum_{\boldsymbol{y}}W_1(\boldsymbol{y}|\boldsymbol{x}_{mi})\cdot\mathcal{I}\left\{\bigcup_{(m',j)\neq(m,i)}\Big\{W_1(\boldsymbol{y}|\boldsymbol{x}_{m'j})\geq W_1(\boldsymbol{y}|\boldsymbol{x}_{mi})\Big\}\right\} \tag{55}$$

$$\stackrel{\rm UB}{\leq}\sum_{\boldsymbol{y}}W_1(\boldsymbol{y}|\boldsymbol{x}_{mi})\cdot\left[\sum_{(m',j)\neq(m,i)}\frac{W_1(\boldsymbol{y}|\boldsymbol{x}_{m'j})}{W_1(\boldsymbol{y}|\boldsymbol{x}_{mi})}\right]^s \tag{56}$$

$$\stackrel{\rm PD}{\leq}\sum_{\boldsymbol{y}}W_1(\boldsymbol{y}|\boldsymbol{x}_{mi})\cdot\frac{\sum_{(m',j)\neq(m,i)}W_1^s(\boldsymbol{y}|\boldsymbol{x}_{m'j})}{W_1^s(\boldsymbol{y}|\boldsymbol{x}_{mi})} \tag{57}$$

$$= \sum_{(m',j)\neq(m,i)} \sum_{\boldsymbol{y}} W_1^{1-s}(\boldsymbol{y}|\boldsymbol{x}_{mi}) \cdot W_1^{s}(\boldsymbol{y}|\boldsymbol{x}_{m'j}) \tag{58}$$

$$= \sum_{(m',j)\neq(m,i)} \exp\left\{-d_s(\boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j})\right\} \tag{59}$$

$$= \sum_{j\neq i} \exp\left\{-d_s(\boldsymbol{x}_{mi}, \boldsymbol{x}_{mj})\right\} + \sum_{m'\neq m} \sum_{j=0}^{M_y-1} \exp\left\{-d_s(\boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j})\right\} \tag{60}$$

$$= \sum_{Q_{UXX'}\in\mathcal{S}} N_{mi}^{\text{IN}}(Q_{UXX'}, \mathcal{C}) \cdot e^{-nD_s(Q_{XX'})} + \sum_{Q_{UU'XX'}\in\mathcal{P}} N_{mi}^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) \cdot e^{-nD_s(Q_{XX'})}, \tag{61}$$

where $N_{mi}^{\text{IN}}(Q_{UXX'}, \mathcal{C})$ denotes the number of codewords $\boldsymbol{x}_{mj} \in \mathcal{C}_m$, for any $j \neq i$, such that the joint empirical distribution of $\boldsymbol{x}_{mj}$ with $(\boldsymbol{u}_m, \boldsymbol{x}_{mi})$ is $Q_{UXX'}$, that is

$$N_{mi}^{\text{IN}}(Q_{UXX'}, \mathcal{C}) = \sum_{j\neq i} \mathcal{I}\Big\{(\boldsymbol{u}_m, \boldsymbol{x}_{mi}, \boldsymbol{x}_{mj}) \in \mathcal{T}(Q_{UXX'})\Big\}, \tag{62}$$

and $N_{mi}^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C})$ denotes the number of pairs $(\boldsymbol{u}_{m'}, \boldsymbol{x}_{m'j})$, $\boldsymbol{x}_{m'j} \in \mathcal{C}_{m'}$, for any $m' \neq m$, such that the joint empirical distribution of them with $(\boldsymbol{u}_m, \boldsymbol{x}_{mi})$ is $Q_{UU'XX'}$, that is

$$N_{mi}^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) = \sum_{m'\neq m} \sum_{j=0}^{M_y-1} \mathcal{I}\Big\{(\boldsymbol{u}_m, \boldsymbol{x}_{mi}, \boldsymbol{u}_{m'}, \boldsymbol{x}_{m'j}) \in \mathcal{T}(Q_{UU'XX'})\Big\}. \tag{63}$$

For the weak user, we have for any $t \in [0, 1]$

$$P_{\text{e}|m}(\mathcal{C}) = \frac{1}{M_y} \sum_{i=0}^{M_y-1} \sum_{\boldsymbol{z}} W_2(\boldsymbol{z}|\boldsymbol{x}_{mi}) \cdot \mathcal{I}\left\{ \bigcup_{m'\neq m} \Big\{W_2(\boldsymbol{z}|\mathcal{C}_{m'}) \geq W_2(\boldsymbol{z}|\mathcal{C}_m)\Big\} \right\} \tag{64}$$

$$\overset{\text{UB}}{\leq} \frac{1}{M_y} \sum_{i=0}^{M_y-1} \sum_{\boldsymbol{z}} W_2(\boldsymbol{z}|\boldsymbol{x}_{mi}) \cdot \left[\sum_{m'\neq m} \frac{W_2(\boldsymbol{z}|\mathcal{C}_{m'})}{W_2(\boldsymbol{z}|\mathcal{C}_m)} \right]^t \tag{65}$$

$$\overset{\text{PD}}{\leq} \frac{1}{M_y} \sum_{i=0}^{M_y-1} \sum_{\boldsymbol{z}} W_2(\boldsymbol{z}|\boldsymbol{x}_{mi}) \cdot \frac{\sum_{m'\neq m} W_2^t(\boldsymbol{z}|\mathcal{C}_{m'})}{W_2^t(\boldsymbol{z}|\mathcal{C}_m)} \tag{66}$$

$$= \sum_{m'\neq m} \sum_{\boldsymbol{z}} W_2^{1-t}(\boldsymbol{z}|\mathcal{C}_m) \cdot W_2^t(\boldsymbol{z}|\mathcal{C}_{m'}) \tag{67}$$

$$= \sum_{m'\neq m} \sum_{\boldsymbol{z}} \left[\left\{ \frac{1}{M_y} \sum_{i=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{mi}) \right\}^{1-t} \left\{ \frac{1}{M_y} \sum_{j=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{m'j}) \right\}^t \right] \tag{68}$$

$$\overset{\text{PD}}{\leq} \sum_{m'\neq m} \sum_{\boldsymbol{z}} \frac{1}{M_y} \sum_{i=0}^{M_y-1} \sum_{j=0}^{M_y-1} W_2^{1-t}(\boldsymbol{z}|\boldsymbol{x}_{mi}) \cdot W_2^t(\boldsymbol{z}|\boldsymbol{x}_{m'j}) \tag{69}$$

$$= \frac{1}{M_y} \sum_{m'\neq m} \sum_{i=0}^{M_y-1} \sum_{j=0}^{M_y-1} \sum_{\boldsymbol{z}} W_2^{1-t}(\boldsymbol{z}|\boldsymbol{x}_{mi}) \cdot W_2^t(\boldsymbol{z}|\boldsymbol{x}_{m'j}) \tag{70}$$

$$= \frac{1}{M_y} \sum_{m' \neq m} \sum_{i=0}^{M_y-1} \sum_{j=0}^{M_y-1} \exp\left\{-d_t(\boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j})\right\} \tag{71}$$

$$= \frac{1}{M_y} \sum_{Q_{UU'XX'} \in \mathcal{P}} \hat{N}_m(Q_{UU'XX'}, \mathcal{C}) e^{-nD_t(Q_{XX'})}, \tag{72}$$

where $\hat{N}_m(Q_{UU'XX'}, \mathcal{C})$ denotes the number of triplets $(\boldsymbol{u}_{m'}, \boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j})$, $\boldsymbol{x}_{mi} \in \mathcal{C}_m$, $\boldsymbol{x}_{m'j} \in \mathcal{C}_{m'}$, for any $m' \neq m$, such that the joint empirical distribution of them with $\boldsymbol{u}_m$ is $Q_{UU'XX'}$, that is

$$\hat{N}_m(Q_{UU'XX'}, \mathcal{C}) = \sum_{m' \neq m} \sum_{i=0}^{M_y-1} \sum_{j=0}^{M_y-1} \mathcal{I}\Big\{(\boldsymbol{u}_m, \boldsymbol{u}_{m'}, \boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j}) \in \mathcal{T}(Q_{UU'XX'})\Big\}. \tag{73}$$

Remark: It is important to notice that the main weakness of this derivation is the passage between (68) and (69), where we are using the PD inequality over the two exponential sums of sizes M_y, but in spite of that, the resulting exponent is still higher at relatively low coding rates, as can be seen in Fig. 1.

5.2 Proving the Existence of Codebooks with Specified Type–Enumerators

At this part we prove the existence of codebooks with prescribed upper bounds on the enumerators. The main idea here is that for every $\epsilon > 0$ and all sufficiently large n, there exists a code $\mathcal{C}$ with a rate pair (R_y, R_z), that satisfies, for every codeword (m, i) and every $Q_{UU'XX'}$,

$$\begin{aligned} &N_{mi}^{\text{IN}}(Q_{UXX'}, \mathcal{C}) \leq N_{\text{IN}}^*(Q_{UXX'}) \\ &\stackrel{\Delta}{=} \begin{cases} \exp\Big\{n[R_y - I_Q(X; X'|U) + \epsilon]\Big\} & R_y \geq I_Q(X; X'|U) - \epsilon \\ 0 & R_y < I_Q(X; X'|U) - \epsilon, \end{cases} \end{aligned} \tag{74}$$

and

$$\begin{aligned} &N_{mi}^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) \leq N_{\text{OUT}}^*(Q_{UU'XX'}) \\ &\stackrel{\Delta}{=} \begin{cases} \exp\Big\{n[R_y + R_z - I_Q(UX; U'X') + \epsilon]\Big\} & R_y + R_z \geq I_Q(UX; U'X') - \epsilon \\ 0 & R_y + R_z < I_Q(UX; U'X') - \epsilon, \end{cases} \end{aligned} \tag{75}$$

and furthermore, for every cloud m and every $Q_{UU'XX'}$,

$$\begin{aligned} &\hat{N}_m(Q_{UU'XX'}, \mathcal{C}) \leq \hat{N}^*(Q_{UU'XX'}) \\ &\stackrel{\Delta}{=} \begin{cases} \exp\Big\{n[2R_y + R_z - I_Q(UX; U'X') + \epsilon]\Big\} & 2R_y + R_z \geq I_Q(UX; U'X') - \epsilon \\ 0 & 2R_y + R_z < I_Q(UX; U'X') - \epsilon. \end{cases} \end{aligned} \tag{76}$$

To show this, we first compute ensemble averages. Define

$$\hat{N}(Q_{UU'XX'},\mathcal{C}) \triangleq \frac{1}{M_z}\sum_{m=0}^{M_z-1}\hat{N}_m(Q_{UU'XX'},\mathcal{C}), \tag{77}$$

$$N^{\text{IN}}(Q_{UXX'},\mathcal{C}) \triangleq \frac{1}{M_yM_z}\sum_{m=0}^{M_z-1}\sum_{i=0}^{M_y-1}N^{\text{IN}}_{mi}(Q_{UXX'},\mathcal{C}), \tag{78}$$

$$N^{\text{OUT}}(Q_{UU'XX'},\mathcal{C}) \triangleq \frac{1}{M_yM_z}\sum_{m=0}^{M_z-1}\sum_{i=0}^{M_y-1}N^{\text{OUT}}_{mi}(Q_{UU'XX'},\mathcal{C}). \tag{79}$$

For the weak user, we have that

$$\mathbb{E}\{\hat{N}(Q_{UU'XX'},\mathcal{C})\} = \frac{1}{M_z}\sum_{m=0}^{M_z-1}\mathbb{E}\{\hat{N}_m(Q_{UU'XX'},\mathcal{C})\} \tag{80}$$

$$= \mathbb{E}\{\hat{N}_0(Q_{UU'XX'},\mathcal{C})\} \tag{81}$$

$$= (M_z-1)\cdot M_y^2\cdot\Pr\big\{(\boldsymbol{U},\boldsymbol{U}',\boldsymbol{X},\boldsymbol{X}')\in\mathcal{T}(Q_{UU'XX'})\big\} \tag{82}$$

$$= (M_z-1)\cdot M_y^2\cdot\frac{|\mathcal{T}(Q_{UU'XX'})|}{|\mathcal{T}(P_U)|^2\cdot|\mathcal{T}(P_{X|U})|^2} \tag{83}$$

$$\doteq (M_z-1)\cdot M_y^2\cdot\frac{\exp\{nH_Q(U,U',X,X')\}}{e^{2nH(U)}\cdot e^{2nH(X|U)}} \tag{84}$$

$$= \exp\Big\{n[2R_y+R_z-I_Q(UX;U'X')]\Big\}. \tag{85}$$

Similarly, for the strong user, we have that

$$\mathbb{E}\{N^{\text{IN}}(Q_{UXX'},\mathcal{C})\} = \frac{1}{M_yM_z}\sum_{m=0}^{M_z-1}\sum_{i=0}^{M_y-1}\mathbb{E}\big\{N^{\text{IN}}_{mi}(Q_{UXX'},\mathcal{C})\big\} \tag{86}$$

$$\doteq \exp\Big\{n[R_y-I_Q(X;X'|U)]\Big\}, \tag{87}$$

and

$$\mathbb{E}\{N^{\text{OUT}}(Q_{UU'XX'},\mathcal{C})\} = \frac{1}{M_yM_z}\sum_{m=0}^{M_z-1}\sum_{i=0}^{M_y-1}\mathbb{E}\big\{N^{\text{OUT}}_{mi}(Q_{UU'XX'},\mathcal{C})\big\} \tag{88}$$

$$\doteq \exp\Big\{n[R_y+R_z-I_Q(UX;U'X')]\Big\}. \tag{89}$$

It then follows that

$$\Pr\left\{\mathcal{C}:\bigcup_{Q_{UU'XX'}}\Big\{\hat{N}(Q_{UU'XX'},\mathcal{C})>\mathbb{E}\{\hat{N}(Q_{UU'XX'},\mathcal{C})\}\cdot e^{n\epsilon/2}\Big\}\right\}$$

$$\bigcup\left\{\bigcup_{Q_{UXX'}}\Big\{N^{\text{IN}}(Q_{UXX'},\mathcal{C})>\mathbb{E}\left\{N^{\text{IN}}(Q_{UXX'},\mathcal{C})\right\}\cdot e^{n\epsilon/2}\Big\}\right\}$$

$$\bigcup \left\{ \bigcup_{Q_{UU'XX'}} \left\{ N^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) > \mathbb{E}\left\{ N^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) \right\} \cdot e^{n\epsilon/2} \right\} \right\} \tag{90}$$

$$\begin{aligned} &\overset{\text{UB}}{\leq} \sum_{Q_{UU'XX'}} \Pr\left\{ \mathcal{C} :\ \hat{N}(Q_{UU'XX'}, \mathcal{C}) > \mathbb{E}\{\hat{N}(Q_{UU'XX'}, \mathcal{C})\} \cdot e^{n\epsilon/2} \right\} \\ &\quad + \sum_{Q_{UXX'}} \Pr\left\{ \mathcal{C} :\ N^{\text{IN}}(Q_{UXX'}, \mathcal{C}) > \mathbb{E}\left\{ N^{\text{IN}}(Q_{UXX'}, \mathcal{C}) \right\} \cdot e^{n\epsilon/2} \right\} \\ &\quad + \sum_{Q_{UU'XX'}} \Pr\left\{ \mathcal{C} :\ N^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) > \mathbb{E}\left\{ N^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) \right\} \cdot e^{n\epsilon/2} \right\} \end{aligned} \tag{91}$$

$$\overset{\text{MI}}{\leq} \sum_{Q_{UU'XX'}} e^{-n\epsilon/2} + \sum_{Q_{UXX'}} e^{-n\epsilon/2} + \sum_{Q_{UU'XX'}} e^{-n\epsilon/2} \tag{92}$$

$$\leq 3 \cdot (n+1)^{|\mathcal{U}|^2 \cdot |\mathcal{X}|^2} \cdot e^{-n\epsilon/2} \to 0, \tag{93}$$

which means that with high probability

$$\frac{1}{M_z} \sum_{m=0}^{M_z-1} \hat{N}_m(Q_{UU'XX'}, \mathcal{C}) \leq e^{n[2R_y + R_z - I_Q(UX;U'X') + \epsilon/2]} \quad \forall Q_{UU'XX'}, \tag{94}$$

$$\frac{1}{M_y M_z} \sum_{m=0}^{M_z-1} \sum_{i=0}^{M_y-1} N_{mi}^{\text{IN}}(Q_{UXX'}, \mathcal{C}) \leq e^{n[R_y - I_Q(X;X'|U) + \epsilon/2]} \quad \forall Q_{UXX'}, \tag{95}$$

$$\frac{1}{M_y M_z} \sum_{m=0}^{M_z-1} \sum_{i=0}^{M_y-1} N_{mi}^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) \leq e^{n[R_y + R_z - I_Q(UX;U'X') + \epsilon/2]} \quad \forall Q_{UU'XX'}. \tag{96}$$

For a given such code and every given $Q_{UU'XX'}$, there must exist at least $(1 - 3e^{-n\epsilon/2}) \cdot M_z$ values of m and at least $(1 - 3e^{-n\epsilon/2}) \cdot M_y M_z$ pairs (m, i) such that

$$\hat{N}_m(Q_{UU'XX'}, \mathcal{C}) \leq e^{n[2R_y + R_z - I_Q(UX;U'X') + \epsilon]}, \tag{97}$$

$$N_{mi}^{\text{IN}}(Q_{UXX'}, \mathcal{C}) \leq e^{n[R_y - I_Q(X;X'|U) + \epsilon]}, \tag{98}$$

$$N_{mi}^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) \leq e^{n[R_y + R_z - I_Q(UX;U'X') + \epsilon]}. \tag{99}$$

Upon eliminating the exceptional clouds and exceptional codewords from the code, for all $Q_{UU'XX'}$, we end up with at least $[1 - 3(n+1)^{|\mathcal{U}|^2 \cdot |\mathcal{X}|^2} \cdot e^{-n\epsilon/2}] \cdot M_z$ clouds and at least $[1 - 3(n+1)^{|\mathcal{U}|^2 \cdot |\mathcal{X}|^2} \cdot e^{-n\epsilon/2}] \cdot M_y M_z$ codewords for which

$$\hat{N}_m(Q_{UU'XX'}, \mathcal{C}) \leq e^{n[2R_y + R_z - I_Q(UX;U'X') + \epsilon]} \quad \forall Q_{UU'XX'}, \tag{100}$$

$$N_{mi}^{\text{IN}}(Q_{UXX'}, \mathcal{C}) \leq e^{n[R_y - I_Q(X;X'|U) + \epsilon]} \quad \forall Q_{UXX'}, \tag{101}$$

$$N_{mi}^{\text{OUT}}(Q_{UU'XX'}, \mathcal{C}) \leq e^{n[R_y + R_z - I_Q(UX;U'X') + \epsilon]} \quad \forall Q_{UU'XX'}. \tag{102}$$

Let $\mathcal{C}'$ denote the sub-code formed by these $[1-3(n+1)^{|\mathcal{U}|^2\cdot|\mathcal{X}|^2}\cdot e^{-n\epsilon/2}]\cdot M_z$ remaining clouds and $[1-3(n+1)^{|\mathcal{U}|^2\cdot|\mathcal{X}|^2}\cdot e^{-n\epsilon/2}]\cdot M_yM_z$ remaining codewords. Since $\hat{N}_m(Q_{UU'XX'},\mathcal{C}')\leq \hat{N}_m(Q_{UU'XX'},\mathcal{C})$, $N_{mi}^{\text{IN}}(Q_{UXX'},\mathcal{C}')\leq N_{mi}^{\text{IN}}(Q_{UXX'},\mathcal{C})$ and $N_{mi}^{\text{OUT}}(Q_{UU'XX'},\mathcal{C}')\leq N_{mi}^{\text{OUT}}(Q_{UU'XX'},\mathcal{C})$, then the sub-code certainly satisfies

$$\hat{N}_m(Q_{UU'XX'},\mathcal{C}')\leq e^{n[2R_y+R_z-I_Q(UX;U'X')+\epsilon]}\quad\forall m,Q_{UU'XX'}, \tag{103}$$

$$N_{mi}^{\text{IN}}(Q_{UXX'},\mathcal{C}')\leq e^{n[R_y-I_Q(X;X'|U)+\epsilon]}\quad\forall (m,i),Q_{UXX'}, \tag{104}$$

$$N_{mi}^{\text{OUT}}(Q_{UU'XX'},\mathcal{C}')\leq e^{n[R_y+R_z-I_Q(UX;U'X')+\epsilon]}\quad\forall (m,i),Q_{UU'XX'}. \tag{105}$$

Finally, observe that since $\hat{N}_m(Q_{UU'XX'},\mathcal{C}')$ is a non-negative integer, then for $Q_{UU'XX'}$ with $2R_y+R_z-I_Q(UX;U'X')+\epsilon<0$, the inequality of (103) means $\hat{N}_m(Q_{UU'XX'},\mathcal{C}')=0$, in which case the right hand side of (103) becomes $\hat{N}^*(Q_{UU'XX'})$, and similarly for (104) with $N_{\text{IN}}^*(Q_{UXX'})$ and for (105) with $N_{\text{OUT}}^*(Q_{UU'XX'})$. Thus, we have shown that there exists a code $\mathcal{C}'$ consisting of $[1-3(n+1)^{|\mathcal{U}|^2\cdot|\mathcal{X}|^2}\cdot e^{-n\epsilon/2}]\cdot e^{nR_z}$ clouds and $[1-3(n+1)^{|\mathcal{U}|^2\cdot|\mathcal{X}|^2}\cdot e^{-n\epsilon/2}]\cdot e^{n(R_y+R_z)}$ codewords, for which all clouds satisfy $\hat{N}_m(Q_{UU'XX'},\mathcal{C}')\leq\hat{N}^*(Q_{UU'XX'})$, and all codewords satisfy $N_{mi}^{\text{IN}}(Q_{UXX'},\mathcal{C}')\leq N_{\text{IN}}^*(Q_{UXX'})$ and $N_{mi}^{\text{OUT}}(Q_{UU'XX'},\mathcal{C}')\leq N_{\text{OUT}}^*(Q_{UU'XX'})$ for all joint types $Q_{UU'XX'}$.

Since there are at least $\frac{1}{2}\cdot e^{nR_z}$ clouds in $\mathcal{C}'$ (we take the factor of one half just for simplicity), and at least $\frac{1}{2}\cdot e^{n(R_z+R_y)}$ codewords, then we conclude that at least $\frac{1}{4}\cdot e^{nR_z}$ clouds contain at least $\frac{1}{4}\cdot e^{nR_y}$ codewords each. Thus, we can eliminate some more clouds and codewords in order to obtain a codebook, for which each cloud contains exactly the same number of codewords.

5.3 Plugging Back Type–Enumerators to the Bounds

As a consequence of the above observation, we have seen the existence of a code $\mathcal{C}'$ for which

$$\begin{aligned}
&\max_m P_{\text{e}|m}(\mathcal{C}')\\
&\leq e^{-nR_y}\sum_{Q_{UU'XX'}\in\mathcal{P}}\hat{N}^*(Q_{UU'XX'})\exp\big\{-nD_t(Q_{XX'})\big\} &(106)\\
&=e^{-nR_y}\sum_{\substack{Q_{UU'XX'}\in\mathcal{P}\\ I_Q(UX;U'X')\leq 2R_y+R_z+\epsilon}}\exp\big\{n[2R_y+R_z-I_Q(UX;U'X')-D_t(Q_{XX'})+\epsilon]\big\} &(107)\\
&=\sum_{\substack{Q_{UU'XX'}\in\mathcal{P}\\ I_Q(UX;U'X')\leq 2R_y+R_z+\epsilon}}\exp\big\{n[R_y+R_z-I_Q(UX;U'X')-D_t(Q_{XX'})+\epsilon]\big\}, &(108)
\end{aligned}$$

as well as

$$\begin{aligned}
&\max_{m,i} P_{\mathrm{e}|mi}(\mathcal{C}') \\
&\leq \sum_{Q_{UXX'}\in\mathcal{S}} N^*_{\mathrm{IN}}(Q_{UXX'})\cdot e^{-nD_s(Q_{XX'})} + \sum_{Q_{UU'XX'}\in\mathcal{P}} N^*_{\mathrm{OUT}}(Q_{UU'XX'})\cdot e^{-nD_s(Q_{XX'})} \qquad (109)\\
&= \sum_{\substack{Q_{UXX'}\in\mathcal{S}\\ I_Q(X;X'|U)\leq R_y+\epsilon}} \exp\Big\{n[R_y - I_Q(X;X'|U) - D_s(Q_{XX'}) + \epsilon]\Big\} \\
&\quad + \sum_{\substack{Q_{UU'XX'}\in\mathcal{P}\\ I_Q(UX;U'X')\leq R_y+R_z+\epsilon}} \exp\Big\{n[R_y + R_z - I_Q(UX;U'X') - D_s(Q_{XX'}) + \epsilon]\Big\}, \qquad (110)
\end{aligned}$$

and due to the arbitrariness of $\epsilon > 0$, this means that there exists a sequence of codes with a rate pair (R_y, R_z) for which (32) and (33) hold. This completes the proof of Theorem 1.

6 Proof of Theorem 2

As we explained in Section 4, the expurgation of an hierarchical codebook can be done in two stages. In Section 6.1, we concentrate only on the cloud-centers, and discard a negligible amount of them in order to have a good set, where by "good" we mean that some specific structural properties are ensured to be valid. Then, we fix the set of cloud-centers, and move to the second stage, where we continue to expurgate clouds, as well as codewords, in order to obtain a complete codebook, for which some additional properties are assured. More specifically, in Section 6.2, we upper bound the error probabilities and express each one of them using suitable quantities. Section 6.3 is rather technical and contains the derivations of the ensemble averages. In Section 6.4, which is the main step of the proof, we show the existence of an hierarchical codebook, for which these quantities are upper bounded by specific deterministic functions of information measures of the fixed cloud centers. In Section 6.5, we put back these deterministic bounds in order to get the desired results.

6.1 Proving the Existence of Codebooks with Specified Clouds Type-Enumerators

Let $N_m(Q_{UU'}, \mathcal{C})$ denote the number of cloud centers $\boldsymbol{u}_{m'}$, for any $m' \neq m$, such that the joint empirical distribution of $\boldsymbol{u}_{m'}$ with $\boldsymbol{u}_m$ is $Q_{UU'}$, that is

$$N_m(Q_{UU'}, \mathcal{C}) = \sum_{m'\neq m} \mathcal{I}\Big\{(\boldsymbol{u}_m, \boldsymbol{u}_{m'}) \in \mathcal{T}(Q_{UU'})\Big\}. \qquad (111)$$

As in proof of Theorem 1, one can easily prove that for every $\epsilon > 0$ and all sufficiently large n, there exists a code $\mathcal{C}$ (cloud-centers) with a rate R_z (essentially), that satisfies, for every cloud m and every $Q_{UU'}$,

$$\begin{aligned} & N_m(Q_{UU'}, \mathcal{C}) \leq N^*(Q_{UU'}) \\ & \stackrel{\Delta}{=} \begin{cases} \exp\Big\{n[R_z - I_Q(U;U') + \epsilon]\Big\} & R_z \geq I_Q(U;U') - \epsilon \\ 0 & R_z < I_Q(U;U') - \epsilon. \end{cases} \end{aligned} \tag{112}$$

Although the proof is omitted here, one important point should be noticed. It is shown that there exists a code $\mathcal{C}'$ consisting of $M_z' \stackrel{\Delta}{=} \left[1 - (n+1)^{|\mathcal{U}|^2} \cdot e^{-n\epsilon/2}\right] \cdot M_z$ clouds, for which all of them satisfy $N_m(Q_{UU'}, \mathcal{C}') \leq N^*(Q_{UU'})$ for all joint types $Q_{UU'}$, and hence, we stick with M_z', instead of the original size of M_z. From now on, we assume without loss of generality that $\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}$ is the set of cloud centers for which this property holds. Now, let us continue to the second stage of expurgation.

6.2 Upper Bounding Error Probabilities Using Type–Enumerators

In order to continue, we need a few definitions. First, define the *average log–likelihood*

$$f(Q_{XZ}) = \frac{1}{n} \log W_2(\boldsymbol{z}|\boldsymbol{x}) = \sum_{(x,z) \in \mathcal{X} \times \mathcal{Z}} Q_{XZ}(x,z) \log W_2(z|x), \tag{113}$$

where Q_{XZ} is understood to be the joint empirical distribution of $(\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{X}^n \times \mathcal{Z}^n$, and let

$$E_0(R_y, Q_{UXZ}, t) \stackrel{\Delta}{=} \begin{cases} \big[R_y - I_Q(X;Z|U)\big] \cdot t & R_y \geq I_Q(X;Z|U) \\ \big[R_y - I_Q(X;Z|U)\big] & R_y < I_Q(X;Z|U). \end{cases} \tag{114}$$

Next, define

$$E_1(R_y, Q_{UZ}, t) \stackrel{\Delta}{=} \max_{Q_{X|UZ} \in \mathcal{S}(Q_{UZ})} \left[E_0(R_y, Q_{UXZ}, t) + t f(Q_{XZ})\right], \tag{115}$$

and furthermore, let

$$B(R_y, t, Q_{UU'}) \stackrel{\Delta}{=} \max_{Q_{Z|UU'}} \left[H_Q(Z|U,U') + E_1(R_y, Q_{UZ}, 1-t) + E_1(R_y, Q_{U'Z}, t)\right]. \tag{116}$$

For the strong user, let us denote

$$K(R_y, s, Q_{UU'}) \stackrel{\Delta}{=} \max_{Q_{XX'|UU'}} \big[R_y - I_Q(X;U'|U) - I_Q(UX;X'|U') - D_s(Q_{XX'})\big]. \tag{117}$$

For the weak user, we use the same upper bound that we have derived in (68):

$$P_{\mathrm{e}|m}(\mathcal{C}) \leq \frac{1}{M_y} \sum_{m' \neq m} \sum_{\boldsymbol{z}} \left[\left\{ \sum_{i=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{mi}) \right\}^{1-t} \left\{ \sum_{j=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{m'j}) \right\}^{t} \right]. \tag{118}$$

Let us further define

$$G(\mathcal{C}_m, \mathcal{C}_{m'}) \stackrel{\Delta}{=} \sum_{\boldsymbol{z}} \left[\left\{ \sum_{i=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{mi}) \right\}^{1-t} \left\{ \sum_{j=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{m'j}) \right\}^{t} \right], \tag{119}$$

such that we can write in short

$$P_{\mathrm{e}|m}(\mathcal{C}) \leq \frac{1}{M_y} \sum_{m' \neq m} G(\mathcal{C}_m, \mathcal{C}_{m'}), \tag{120}$$

to emphasize here the difference from Theorem 1. The "price" for not using the upper bound $(\sum_i a_i)^t \leq \sum_i a_i^t$ is that we need to keep the channel output $\boldsymbol{z}$. Next, for the strong user, we take the upper bound of (60)

$$P_{\mathrm{e}|mi}(\mathcal{C}) \leq \sum_{j \neq i} \exp\left\{-d_s(\boldsymbol{x}_{mi}, \boldsymbol{x}_{mj})\right\} + \sum_{m' \neq m} \sum_{j=0}^{M_y-1} \exp\left\{-d_s(\boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j})\right\} \tag{121}$$

$$= \sum_{Q_{UXX'} \in \mathcal{S}} N_{mi}^{\mathrm{IN}}(Q_{UXX'}, \mathcal{C}) \cdot e^{-nD_s(Q_{XX'})} + \sum_{m' \neq m} \sum_{j=0}^{M_y-1} \exp\left\{-d_s(\boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j})\right\}, \tag{122}$$

where $N_{mi}^{\mathrm{IN}}(Q_{UXX'}, \mathcal{C})$ has already been defined in (62).

6.3 Evaluation of Expected Values of Enumerators

We will prove in Section 6.4 that for every $\epsilon > 0$ and all sufficiently large n, there exists a code $\mathcal{C}'' \subseteq \mathcal{C}'$ with a rate pair (R_y, R_z), that satisfies, for every codeword (m, i) and every $Q_{UXX'}$,

$$\begin{aligned} & N_{mi}^{\mathrm{IN}}(Q_{UXX'}, \mathcal{C}'') \leq N_{\mathrm{IN}}^{*}(Q_{UXX'}) \\ & \stackrel{\Delta}{=} \begin{cases} \exp\Big\{n[R_y - I_Q(X; X'|U) + \epsilon]\Big\} & R_y \geq I_Q(X; X'|U) - \epsilon \\ 0 & R_y < I_Q(X; X'|U) - \epsilon, \end{cases} \end{aligned} \tag{123}$$

and

$$\sum_{m' \neq m} \sum_{j=0}^{M_y-1} \exp\left\{-d_s(\boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j})\right\} \leq \frac{1}{M_z'} \sum_{m=0}^{M_z'-1} \sum_{m' \neq m} \exp\Big\{n[K(R_y, s, \hat{P}_{\boldsymbol{u}_m \boldsymbol{u}_{m'}}) + \epsilon]\Big\}, \tag{124}$$

and furthermore, for every cloud m,

$$\sum_{m' \neq m} G(\mathcal{C}_m, \mathcal{C}_{m'}) \leq \frac{1}{M_z'} \sum_{m=0}^{M_z'-1} \sum_{m' \neq m} \exp\Big\{n[B(R_y, t, \hat{P}_{\boldsymbol{u}_m \boldsymbol{u}_{m'}}) + \epsilon]\Big\}. \tag{125}$$

In the meantime, we move to a rather technical part of calculating the conditional expectations (given the fixed set of cloud-centers) of the averages of the left-hand-sides of (123), (124) and (125). We start with the strong user. As in previous cases, we have that

$$\overline{N^{\rm IN}(Q_{UXX'})} \stackrel{\Delta}{=} \frac{1}{M_y M_z'} \sum_{m=0}^{M_z'-1} \sum_{i=0}^{M_y-1} \mathbb{E}\big\{N_{mi}^{\rm IN}(Q_{UXX'},\mathcal{C})\big\} \tag{126}$$

$$\doteq \exp\Big\{n[R_y - I_Q(X;X'|U)]\Big\}. \tag{127}$$

Preparing to the next step, let us define the enumerator $N_{mi}^{m'}(Q_{UU'XX'},\mathcal{C})$ to be the number of codewords $\boldsymbol{x}_{m'j}\in\mathcal{C}_{m'}$, such that the joint empirical distribution of $\boldsymbol{x}_{m'j}$ with $(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi})$ is $Q_{UU'XX'}$, where $(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi})\in\mathcal{T}(Q_{UU'X})$, that is

$$N_{mi}^{m'}(Q_{UU'XX'},\mathcal{C}) = \sum_{j=0}^{M_y-1} \mathcal{I}\Big\{(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j})\in\mathcal{T}(Q_{UU'XX'})\Big\}. \tag{128}$$

Also note that $N_{mi}^{\rm OUT}(Q_{UU'XX'},\mathcal{C}) = \sum_{m'\neq m} N_{mi}^{m'}(Q_{UU'XX'},\mathcal{C})$. By the method of types,

$$\mathbb{E}\left\{N_{mi}^{m'}(Q_{UU'XX'},\mathcal{C})\Big|\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi}\right\} \doteq \exp\Big\{n[R_y - I_Q(UX;X'|U')]\Big\}, \tag{129}$$

such that we can evaluate the second expression of the strong user

$$\mathbb{E}\left\{\sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\big\{-d_s(\boldsymbol{X}_{mi},\boldsymbol{X}_{m'j})\big\}\Big|\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\}$$

$$= \sum_{m'\neq m}\mathbb{E}\left\{\sum_{j=0}^{M_y-1}\exp\big\{-d_s(\boldsymbol{X}_{mi},\boldsymbol{X}_{m'j})\big\}\Big|\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \tag{130}$$

$$= \sum_{m'\neq m}\mathbb{E}\left[\mathbb{E}\left\{\sum_{j=0}^{M_y-1}\exp\big\{-d_s(\boldsymbol{x}_{mi},\boldsymbol{X}_{m'j})\big\}\Big|\boldsymbol{X}_{mi},\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\}\right] \tag{131}$$

$$= \sum_{m'\neq m}\sum_{Q_{X|UU'}}\sum_{\boldsymbol{x}_{mi}\in\mathcal{T}(Q_{X|UU'})}\frac{1}{|\mathcal{T}(P_{X|U})|}$$

$$\times\,\mathbb{E}\left\{\sum_{j=0}^{M_y-1}\exp\big\{-d_s(\boldsymbol{x}_{mi},\boldsymbol{X}_{m'j})\big\}\Big|\boldsymbol{x}_{mi}\in\mathcal{T}(Q_{X|UU'}),\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\}. \tag{132}$$

For the conditional expectation in the last expression, we have the following

$$\mathbb{E}\left\{\sum_{j=0}^{M_y-1}\exp\Big\{-d_s(\boldsymbol{x}_{mi},\boldsymbol{X}_{m'j})\Big\}\Big|\boldsymbol{x}_{mi}\in\mathcal{T}(Q_{X|UU'}),\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \tag{133}$$

$$= \mathbb{E}\left\{\sum_{Q_{X'|UU'X}} N_{mi}^{m'}(Q_{UU'XX'},\mathcal{C})\cdot e^{-nD_s(Q_{XX'})}\Big|\boldsymbol{x}_{mi}\in\mathcal{T}(Q_{X|UU'}),\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \tag{134}$$

$$= \sum_{Q_{X'|UU'X}} \mathbb{E}\left\{N_{mi}^{m'}(Q_{UU'XX'},\mathcal{C})\middle|\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi}\right\}\cdot e^{-nD_s(Q_{XX'})} \tag{135}$$

$$\doteq \sum_{Q_{X'|UU'X}} \exp\left\{n[R_y - I_Q(UX;X'|U')]\right\}\cdot e^{-nD_s(Q_{XX'})} \tag{136}$$

$$\doteq \exp\left\{n\cdot\max_{Q_{X'|XUU'}}[R_y - I_Q(UX;X'|U') - D_s(Q_{XX'})]\right\}, \tag{137}$$

and substituting it back into (132) gives

$$\mathbb{E}\left\{\sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\left\{-d_s(\boldsymbol{X}_{mi},\boldsymbol{X}_{m'j})\right\}\middle|\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \tag{138}$$

$$\doteq \sum_{m'\neq m}\sum_{Q_{X|UU'}}\sum_{\boldsymbol{x}_{mi}\in\mathcal{T}(Q_{X|UU'})}\frac{1}{|\mathcal{T}(P_{X|U})|}$$

$$\times\exp\left\{n\max_{Q_{X'|XUU'}}[R_y - I_Q(UX;X'|U') - D_s(Q_{XX'})]\right\} \tag{139}$$

$$= \sum_{m'\neq m}\sum_{Q_{X|UU'}}\frac{|\mathcal{T}(Q_{X|UU'})|}{|\mathcal{T}(P_{X|U})|}\cdot\exp\left\{n\cdot\max_{Q_{X'|XUU'}}[R_y - I_Q(UX;X'|U') - D_s(Q_{XX'})]\right\} \tag{140}$$

$$\doteq \sum_{m'\neq m}\max_{Q_{X|UU'}} e^{-nI_Q(X;U'|U)}\exp\left\{n\cdot\max_{Q_{X'|XUU'}}\left[R_y - I_Q(UX;X'|U') - D_s(Q_{XX'})\right]\right\} \tag{141}$$

$$= \sum_{m'\neq m}\exp\left\{nK(R_y,s,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})\right\}, \tag{142}$$

where $K(R_y,s,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})$ has been defined in (117). Hence,

$$\mathbb{E}\left\{\frac{1}{M_yM_z'}\sum_{m=0}^{M_z'-1}\sum_{i=0}^{M_y-1}\sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\left\{-d_s(\boldsymbol{X}_{mi},\boldsymbol{X}_{m'j})\right\}\middle|\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \tag{143}$$

$$= \frac{1}{M_yM_z'}\sum_{m=0}^{M_z'-1}\sum_{i=0}^{M_y-1}\mathbb{E}\left\{\sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\left\{-d_s(\boldsymbol{X}_{mi},\boldsymbol{X}_{m'j})\right\}\middle|\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \tag{144}$$

$$\doteq \frac{1}{M_yM_z'}\sum_{m=0}^{M_z'-1}\sum_{i=0}^{M_y-1}\sum_{m'\neq m}\exp\left\{nK(R_y,s,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})\right\} \tag{145}$$

$$= \frac{1}{M_z'}\sum_{m=0}^{M_z'-1}\sum_{m'\neq m}\exp\left\{nK(R_y,s,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})\right\}. \tag{146}$$

Next, we assess the expectation of $G(\mathcal{C}_m,\mathcal{C}_{m'})$. Let $N_{\boldsymbol{z},\boldsymbol{u}_m}(Q_{UXZ},\mathcal{C})$ denote the number of codewords $\boldsymbol{x}_{mi}\in\mathcal{C}_m$, such that the joint empirical distribution of $\boldsymbol{x}_{mi}$ with $(\boldsymbol{u}_m,\boldsymbol{z})$ is Q_{UXZ},

that is

$$N_{\boldsymbol{z},\boldsymbol{u}_m}(Q_{UXZ},\mathcal{C}) = \sum_{i=0}^{M_y-1} \mathcal{I}\Big\{(\boldsymbol{u}_m,\boldsymbol{x}_{mi},\boldsymbol{z}) \in \mathcal{T}(Q_{UXZ})\Big\}. \tag{147}$$

It follows from eq. (4) that

$$\begin{aligned}
&\mathbb{E}\left\{[N_{\boldsymbol{z},\boldsymbol{u}_m}(Q_{UXZ},\mathcal{C})]^{1-t}\right\} \\
&\doteq \begin{cases} \exp\Big\{n\big[R_y - I_Q(X;Z|U)\big](1-t)\Big\} & R_y \geq I_Q(X;Z|U) \\ \exp\Big\{n\big[R_y - I_Q(X;Z|U)\big]\Big\} & R_y < I_Q(X;Z|U) \end{cases} \qquad (148)\\
&= \exp\left\{nE_0(R_y,Q_{UXZ},1-t)\right\}. \qquad (149)
\end{aligned}$$

Hence, from the independence of codewords of different clouds, we get

$$\begin{aligned}
&\mathbb{E}\left\{\sum_{m'\neq m} G(\mathcal{C}_m,\mathcal{C}_{m'})\,\middle|\,\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \\
&= \mathbb{E}\left\{\sum_{m'\neq m}\sum_{\boldsymbol{z}}\left[\left\{\sum_{i=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right\}^{1-t}\left\{\sum_{j=0}^{M_y-1} W_2(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right\}^{t}\right]\,\middle|\,\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \qquad (150)\\
&= \mathbb{E}\left\{\sum_{m'\neq m}\sum_{\boldsymbol{z}}\left[\left\{\sum_{Q_{X|UZ}} N_{\boldsymbol{z},\boldsymbol{u}_m}(Q_{UXZ},\mathcal{C})e^{nf(Q_{XZ})}\right\}^{1-t}\right.\right. \\
&\qquad\qquad \left.\left.\times\left\{\sum_{Q_{X'|U'Z}} N_{\boldsymbol{z},\boldsymbol{u}_{m'}}(Q_{U'X'Z},\mathcal{C})e^{nf(Q_{X'Z})}\right\}^{t}\right]\,\middle|\,\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \qquad (151)\\
&\overset{\text{PD}}{\doteq} \mathbb{E}\left\{\sum_{m'\neq m}\sum_{\boldsymbol{z}}\left[\sum_{Q_{X|UZ}}[N_{\boldsymbol{z},\boldsymbol{u}_m}(Q_{UXZ},\mathcal{C})]^{1-t}\,e^{n(1-t)f(Q_{XZ})}\right.\right. \\
&\qquad\qquad \left.\left.\times\sum_{Q_{X'|U'Z}}\big[N_{\boldsymbol{z},\boldsymbol{u}_{m'}}(Q_{U'X'Z},\mathcal{C})\big]^{t}\,e^{ntf(Q_{X'Z})}\right]\,\middle|\,\{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \qquad (152)\\
&= \sum_{m'\neq m}\sum_{\boldsymbol{z}}\left[\sum_{Q_{X|UZ}}\mathbb{E}\Big\{\,[N_{\boldsymbol{z},\boldsymbol{u}_m}(Q_{UXZ},\mathcal{C})]^{1-t}\,\Big\}e^{n(1-t)f(Q_{XZ})}\right. \\
&\qquad\qquad \left.\times\sum_{Q_{X'|U'Z}}\mathbb{E}\Big\{\,\big[N_{\boldsymbol{z},\boldsymbol{u}_{m'}}(Q_{U'X'Z},\mathcal{C})\big]^{t}\,\Big\}e^{ntf(Q_{X'Z})}\right] \qquad (153)\\
&\doteq \sum_{m'\neq m}\sum_{\boldsymbol{z}}\left[\sum_{Q_{X|UZ}}\exp\Big\{nE_0(R_y,Q_{UXZ},1-t)\Big\}\cdot e^{n(1-t)f(Q_{XZ})}\right. \\
&\qquad\qquad \left.\times\sum_{Q_{X'|U'Z}}\exp\Big\{nE_0(R_y,Q_{U'X'Z},t)\Big\}\cdot e^{ntf(Q_{X'Z})}\right] \qquad (154)
\end{aligned}$$

$$\doteq \sum_{m'\neq m}\sum_{\boldsymbol{z}}\Bigg[\max_{Q_{X|UZ}}\exp\Big\{nE_0(R_y,Q_{UXZ},1-t)\Big\}\cdot e^{n(1-t)f(Q_{XZ})} \times \max_{Q_{X'|U'Z}}\exp\Big\{nE_0(R_y,Q_{U'X'Z},t)\Big\}\cdot e^{ntf(Q_{X'Z})}\Bigg] \tag{155}$$

$$= \sum_{m'\neq m}\sum_{Q_{Z|UU'}}\sum_{\boldsymbol{z}\in\mathcal{T}(Q_{Z|UU'}|\boldsymbol{u}_m,\boldsymbol{u}_{m'})}\Big[\exp\big\{nE_1(R_y,Q_{UZ},1-t)\big\} \times\exp\big\{nE_1(R_y,Q_{U'Z},t)\big\}\Big] \tag{156}$$

$$\doteq \sum_{m'\neq m}\sum_{Q_{Z|UU'}}\exp\Big\{n\big[H_Q(Z|U,U')+E_1(R_y,Q_{UZ},1-t)+E_1(R_y,Q_{U'Z},t)\big]\Big\} \tag{157}$$

$$\doteq \sum_{m'\neq m}\max_{Q_{Z|UU'}}\exp\Big\{n\big[H_Q(Z|U,U')+E_1(R_y,Q_{UZ},1-t)+E_1(R_y,Q_{U'Z},t)\big]\Big\} \tag{158}$$

$$= \sum_{m'\neq m}\exp\Big\{nB(R_y,t,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})\Big\}, \tag{159}$$

which means that

$$\mathbb{E}\left\{\frac{1}{M'_z}\sum_{m=0}^{M'_z-1}\sum_{m'\neq m}G(\mathcal{C}_m,\mathcal{C}_{m'})\,\middle|\,\{\boldsymbol{u}_l\}_{l=0}^{M'_z-1}\right\} = \frac{1}{M'_z}\sum_{m=0}^{M'_z-1}\sum_{m'\neq m}\exp\Big\{nB(R_y,t,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})\Big\}. \tag{160}$$

6.4 Proving the Existence of Codebooks with Specified Satellites Type–Enumerators

Define now the following three events. For the strong user, define

$$\mathcal{A}_1 \stackrel{\Delta}{=} \bigcup_{Q_{UXX'}}\left\{\mathcal{C}:\ \frac{1}{M_yM'_z}\sum_{m=0}^{M'_z-1}\sum_{i=0}^{M_y-1}N^{\text{IN}}_{mi}(Q_{UXX'},\mathcal{C})>e^{n[R_y-I_Q(X;X'|U)+\epsilon/2]}\right\}, \tag{161}$$

and

$$\mathcal{A}_2 \stackrel{\Delta}{=} \left\{\mathcal{C}:\ \frac{1}{M_yM'_z}\sum_{m=0}^{M'_z-1}\sum_{i=0}^{M_y-1}\sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\big\{-d_s(\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j})\big\} > \frac{1}{M'_z}\sum_{m=0}^{M'_z-1}\sum_{m'\neq m}e^{n[K(R_y,s,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})+\epsilon/2]}\right\}. \tag{162}$$

For the weak user, define

$$\mathcal{A}_3 \stackrel{\Delta}{=} \left\{\mathcal{C}:\ \frac{1}{M'_z}\sum_{m=0}^{M'_z-1}\sum_{m'\neq m}G(\mathcal{C}_m,\mathcal{C}_{m'})>\frac{1}{M'_z}\sum_{m=0}^{M'_z-1}\sum_{m'\neq m}e^{n[B(R_y,t,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})+\epsilon/2]}\right\}. \tag{163}$$

Notice that the expression in the right hand side inside each of these events equals the conditional expectation (given the set of cloud centers) of the expression in the left hand side, multiplied by $e^{n\epsilon/2}$. Hence, by Markov's inequality,

$$\Pr\left\{\mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3 \Big| \{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \tag{164}$$

$$\stackrel{\text{UB}}{\leq} \Pr\left\{\mathcal{A}_1 \Big| \{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} + \Pr\left\{\mathcal{A}_2 \Big| \{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} + \Pr\left\{\mathcal{A}_3 \Big| \{\boldsymbol{u}_l\}_{l=0}^{M_z'-1}\right\} \tag{165}$$

$$\stackrel{\text{MI}}{\leq} \sum_{Q_{UXX'}} e^{-n\epsilon/2} + e^{-n\epsilon/2} + e^{-n\epsilon/2} \tag{166}$$

$$\leq 3 \cdot (n+1)^{|\mathcal{U}|\cdot|\mathcal{X}|^2} \cdot e^{-n\epsilon/2} \to 0. \tag{167}$$

Without repeating on a similar proof as in Section 5.2, we conclude that upon eliminating exceptional clouds and exceptional codewords from the code, we end up with a code $\mathcal{C}''$, for which all clouds satisfy both (112) and

$$\sum_{m' \neq m} G_{mm'}(\mathcal{C}_m, \mathcal{C}_{m'}) \leq \frac{1}{M_z'} \sum_{m=0}^{M_z'-1} \sum_{m' \neq m} e^{n[B(R_y,t,\hat{P}_{\boldsymbol{u}_m \boldsymbol{u}_{m'}})+\epsilon]} \quad \forall m, \tag{168}$$

and all codewords satisfy

$$\sum_{m' \neq m} \sum_{j=0}^{M_y-1} \exp\left\{-d_s(\boldsymbol{x}_{mi}, \boldsymbol{x}_{m'j})\right\} \leq \frac{1}{M_z'} \sum_{m=0}^{M_z'-1} \sum_{m' \neq m} e^{n[K(R_y,s,\hat{P}_{\boldsymbol{u}_m \boldsymbol{u}_{m'}})+\epsilon]} \quad \forall (m,i) \tag{169}$$

and $N_{mi}^{\text{IN}}(Q_{UXX'}, \mathcal{C}'') \leq N_{\text{IN}}^*(Q_{UXX'})$ for all joint types $Q_{UXX'}$.

6.5 Plugging Back Type–Enumerators to the Bounds

As a consequence of the previous parts, we have seen the existence of a code $\mathcal{C}''$ for which

$$\max_m P_{\text{e}|m}(\mathcal{C}'')$$

$$\leq \frac{1}{M_y} \sum_{m' \neq m} G(\mathcal{C}_m, \mathcal{C}_{m'}) \tag{170}$$

$$\leq \frac{1}{M_y} \frac{1}{M_z'} \sum_{m=0}^{M_z'-1} \sum_{m' \neq m} \exp\left\{n[B(R_y,t,\hat{P}_{\boldsymbol{u}_m \boldsymbol{u}_{m'}}) + \epsilon]\right\} \tag{171}$$

$$= \frac{1}{M_y} \frac{1}{M_z'} \sum_{m=0}^{M_z'-1} \sum_{Q_{UU'} \in \mathcal{Q}} N_m(Q_{UU'}, \mathcal{C}') \cdot \exp\left\{n[B(R_y,t,Q_{UU'}) + \epsilon]\right\} \tag{172}$$

$$\leq \frac{1}{M_y} \sum_{\substack{Q_{UU'} \in \mathcal{Q} \\ I_Q(U;U') \leq R_z + \epsilon}} \exp\left\{n[R_z - I_Q(U;U') + \epsilon]\right\} \cdot \exp\left\{n[B(R_y,t,Q_{UU'}) + \epsilon]\right\} \tag{173}$$

$$= \sum_{\substack{Q_{UU'}\in\mathcal{Q}\\ I_Q(U;U')\le R_z+\epsilon}} \exp\Big\{n[R_z - I_Q(U;U') + B(R_y,t,Q_{UU'}) - R_y + 2\epsilon]\Big\}, \tag{174}$$

as well as

$$\max_{m,i} P_{\mathrm{e}|mi}(\mathcal{C}'')$$

$$\le \sum_{Q_{UXX'}\in\mathcal{S}} N_{mi}^{\mathrm{IN}}(Q_{UXX'},\mathcal{C}'')\cdot e^{-nD_s(Q_{XX'})} + \sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\big\{-d_s(\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j})\big\} \tag{175}$$

$$\le \sum_{Q_{UXX'}\in\mathcal{S}} N_{\mathrm{IN}}^{*}(Q_{UXX'})\cdot e^{-nD_s(Q_{XX'})} + \frac{1}{M_z'}\sum_{m=0}^{M_z'-1}\sum_{m'\neq m}\exp\Big\{n[K(R_y,s,\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}})+\epsilon]\Big\} \tag{176}$$

$$= \sum_{\substack{Q_{UXX'}\in\mathcal{S}\\ I_Q(X;X'|U)\le R_y+\epsilon}} \exp\Big\{n[R_y - I_Q(X;X'|U) - D_s(Q_{XX'})+\epsilon]\Big\}$$
$$+ \frac{1}{M_z'}\sum_{m=0}^{M_z'-1}\sum_{Q_{UU'}\in\mathcal{Q}} N_m(Q_{UU'},\mathcal{C}')\cdot\exp\Big\{n[K(R_y,s,Q_{UU'})+\epsilon]\Big\} \tag{177}$$

$$\le \sum_{\substack{Q_{UXX'}\in\mathcal{S}\\ I_Q(X;X'|U)\le R_y+\epsilon}} \exp\Big\{n[R_y - I_Q(X;X'|U) - D_s(Q_{XX'})+\epsilon]\Big\}$$
$$+ \sum_{\substack{Q_{UU'}\in\mathcal{Q}\\ I_Q(U;U')\le R_z+\epsilon}} \exp\Big\{n[R_z - I_Q(U;U')+\epsilon]\Big\}\cdot\exp\Big\{n[K(R_y,s,Q_{UU'})+\epsilon]\Big\} \tag{178}$$

and due to the arbitrariness of $\epsilon > 0$, this means that there exists a sequence of codes with a rate pair (R_y, R_z) for which (40) and (41) hold. As for the final expression for the weak user, one can easily move from the expressions of E_0, E_1 and B [defined in (114)–(116)] to the expression of $\hat{D}_t$ [defined in (34)] by using standard information–measures identities. This completes the proof of Theorem 2.

7 Proof of Theorem 3

Consider first the two expressions

$$\Phi_{m,i}(\boldsymbol{u}_m,\boldsymbol{y}) \stackrel{\Delta}{=} \sum_{j\neq i}\exp\{ng(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mj}\boldsymbol{y}})\}, \tag{179}$$

$$\Psi_m(\boldsymbol{y}) \stackrel{\Delta}{=} \sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\{ng(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}\boldsymbol{y}})\}, \tag{180}$$

where the sum in $\Phi_{m,i}(\boldsymbol{u}_m, \boldsymbol{y})$ is over all wrong codewords in the true cloud, while the sum in $\Psi_m(\boldsymbol{y})$ is over all codewords of the incorrect clouds. Let $\epsilon > 0$ be arbitrarily small, and for every $\boldsymbol{y} \in \mathcal{Y}^n$ and $\boldsymbol{u}_m \in \mathcal{U}^n$, define the events

$$\mathcal{S}_\epsilon(m, i, \boldsymbol{u}_m, \boldsymbol{y}) = \Big\{\mathcal{C} : \ \Phi_{m,i}(\boldsymbol{u}_m, \boldsymbol{y}) \le \exp\{n\phi(R_y - \epsilon, \hat{P}_{\boldsymbol{u}_m \boldsymbol{y}})\}\Big\}, \tag{181}$$

$$\mathcal{K}_\epsilon(m, \boldsymbol{y}) = \Big\{\mathcal{C} : \ \Psi_m(\boldsymbol{y}) \le \exp\{n\psi(R_y - \epsilon, R_z - \epsilon, \hat{P}_{\boldsymbol{y}})\}\Big\}, \tag{182}$$

where ϕ and ψ are defined in (44) and (45), respectively. In Appendix C we show that the vast majority of constant composition codes $\{\mathcal{C}\}$ are outside $\mathcal{K}_\epsilon(m, \boldsymbol{y})$, simultaneously for all m and all $\boldsymbol{y}$. Proving that the vast majority of constant composition codes are outside $\mathcal{S}_\epsilon(m, i, \boldsymbol{u}_m, \boldsymbol{y})$, simultaneously for all (m, i), all $\boldsymbol{u}_m$ and all $\boldsymbol{y}$, is only a slight modification of [13, Appendix B] and can be found in [18, Appendix C]. More precisely, it is shown that

$$\Pr\{\mathcal{S}_\epsilon(m, i, \boldsymbol{u}_m, \boldsymbol{y})\} \le \exp\left\{-e^{n\epsilon} + n\epsilon + 1\right\}, \tag{183}$$

$$\Pr\{\mathcal{K}_\epsilon(m, \boldsymbol{y})\} \le \exp\left\{-e^{n\epsilon} + n\epsilon + 1\right\}, \tag{184}$$

for every (m, i), $\boldsymbol{u}_m$ and $\boldsymbol{y}$, and hence,

$$\Pr\left\{\bigcup_m \bigcup_i \bigcup_{\boldsymbol{u}_m \in \mathcal{U}^n} \bigcup_{\boldsymbol{y} \in \mathcal{Y}^n} \left\{\mathcal{S}_\epsilon(m, i, \boldsymbol{u}_m, \boldsymbol{y}) \cup \mathcal{K}_\epsilon(m, \boldsymbol{y})\right\}\right\} \tag{185}$$

$$\stackrel{\Delta}{=} \Pr\left\{\mathcal{B}_\epsilon\right\} \tag{186}$$

$$\stackrel{\text{UB}}{\le} 2e^{n(R_z + R_y)} |\mathcal{U}|^n |\mathcal{Y}|^n \exp\left\{-e^{n\epsilon} + n\epsilon + 1\right\}, \tag{187}$$

which still decays double-exponentially. Thus, for all codes in $\mathcal{G}_\epsilon = \mathcal{B}_\epsilon^c$, which is the vast majority of codes, we have both $\Phi_{m,i}(\boldsymbol{u}_m, \boldsymbol{y}) \ge \exp\{n\phi(R_y - \epsilon, \hat{P}_{\boldsymbol{u}_m \boldsymbol{y}})\}$ and $\Psi_m(\boldsymbol{y}) \ge \exp\{n\psi(R_y - \epsilon, R_z - \epsilon, \hat{P}_{\boldsymbol{y}})\}$ simultaneously for all $m = 0, 1, \ldots, M_z - 1$, $i = 0, 1, \ldots, M_y - 1$, $\boldsymbol{y} \in \mathcal{Y}^n$ and $\boldsymbol{u}_m \in \mathcal{U}^n$.

7.1 Analysis for the weak user

For a given code $\mathcal{C}$, the average probability of error given that one of the messages from cloud m was transmitted is given by

$$P_{e|m}(\mathcal{C}) = \frac{1}{M_y} \sum_{i=0}^{M_y - 1} \sum_{m' \ne m} \sum_{\boldsymbol{z}} W_2(\boldsymbol{z} | \boldsymbol{x}_{mi}) \cdot \Pr\left\{\tilde{m} = m' | \boldsymbol{z}\right\} \tag{188}$$

$$= \frac{1}{M_y} \sum_{i=0}^{M_y - 1} \sum_{m' \ne m} \sum_{\boldsymbol{z}} W_2(\boldsymbol{z} | \boldsymbol{x}_{mi})$$

$$\times \frac{\sum_{j=0}^{M_y-1} \exp\{ng(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}\boldsymbol{z}})\}}{\sum_{l=0}^{M_y-1} \exp\{ng(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{ml}\boldsymbol{z}})\} + \sum_{\tilde{m}\neq m}\sum_{k=0}^{M_y-1} \exp\{ng(\hat{P}_{\boldsymbol{u}_{\tilde{m}}\boldsymbol{x}_{\tilde{m}k}\boldsymbol{z}})\}}. \tag{189}$$

Notice that for every codebook within the set $\mathcal{G}_\epsilon$, and a given index j [appearing in the nominator of the last term of (189)], we have

$$\sum_{l=0}^{M_y-1} \exp\{ng(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{ml}\boldsymbol{z}})\} + \sum_{\tilde{m}\neq m}\sum_{k=0}^{M_y-1} \exp\{ng(\hat{P}_{\boldsymbol{u}_{\tilde{m}}\boldsymbol{x}_{\tilde{m}k}\boldsymbol{z}})\} \tag{190}$$

$$= \exp\{ng(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{z}})\} + \sum_{l\neq i} \exp\{ng(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{ml}\boldsymbol{z}})\} + \exp\{ng(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}\boldsymbol{z}})\} + \sum_{\substack{\tilde{m}\neq m \\ (\tilde{m},k)\neq(m',j)}} \exp\{ng(\hat{P}_{\boldsymbol{u}_{\tilde{m}}\boldsymbol{x}_{\tilde{m}k}\boldsymbol{z}})\} \tag{191}$$

$$\geq \exp\{ng(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{z}})\} + \exp\{n\phi(R_y-\epsilon, \hat{P}_{\boldsymbol{u}_m\boldsymbol{z}})\} + \exp\{ng(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}\boldsymbol{z}})\} + \exp\{n\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{z}})\} \tag{192}$$

$$\doteq \exp\left\{n\max\left[g(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{z}}), \phi(R_y-\epsilon, \hat{P}_{\boldsymbol{u}_m\boldsymbol{z}}), g(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}\boldsymbol{z}}), \psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{z}})\right]\right\}. \tag{193}$$

For any $\rho \geq 1$, we take the expectation over the randomness of $\mathcal{C}^m = \mathcal{C} \setminus \{\boldsymbol{u}_m\}$, while $\boldsymbol{u}_m$ is kept fixed.

$$\mathbb{E}\left\{\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho}\middle|\boldsymbol{u}_m\right\}$$

$$= \sum_{\mathcal{C}^m} P(\mathcal{C}^m)\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho} \tag{194}$$

$$= \sum_{\mathcal{C}^m\in\mathcal{B}_\epsilon} P(\mathcal{C}^m)\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho} + \sum_{\mathcal{C}^m\in\mathcal{G}_\epsilon} P(\mathcal{C}^m)\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho} \tag{195}$$

$$= \sum_{\mathcal{C}^m\in\mathcal{B}_\epsilon} P(\mathcal{C}^m)\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho} + \sum_{\mathcal{C}^m\in\mathcal{G}_\epsilon} P(\mathcal{C}^m)\left(\frac{1}{M_y}\sum_{i=0}^{M_y-1}\sum_{m'\neq m}\sum_{\boldsymbol{z}} W(\boldsymbol{z}|\boldsymbol{x}_{mi}) \times \sum_{j=0}^{M_y-1}\frac{\exp\{ng(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}\boldsymbol{z}})\}}{\sum_{l=0}^{M_y-1}\exp\{ng(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{ml}\boldsymbol{z}})\} + \sum_{\tilde{m}\neq m}\sum_{k=0}^{M_y-1}\exp\{ng(\hat{P}_{\boldsymbol{u}_{\tilde{m}}\boldsymbol{x}_{\tilde{m}k}\boldsymbol{z}})\}}\right)^{1/\rho} \tag{196}$$

$$\leq \sum_{\mathcal{C}^m\in\mathcal{B}_\epsilon} P(\mathcal{C}^m)\cdot 1^{1/\rho} + \sum_{\mathcal{C}^m\in\mathcal{G}_\epsilon} P(\mathcal{C}^m)\left(\frac{1}{M_y}\sum_{i=0}^{M_y-1}\sum_{m'\neq m}\sum_{\boldsymbol{z}} W(\boldsymbol{z}|\boldsymbol{x}_{mi}) \times \sum_{j=0}^{M_y-1}\frac{\exp\{ng(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}\boldsymbol{z}})\}}{\exp\left\{n\max\left[g(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}\boldsymbol{z}}), \phi(R_y-\epsilon, \hat{P}_{\boldsymbol{u}_m\boldsymbol{z}}), g(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}\boldsymbol{z}}), \psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{z}})\right]\right\}}\right)^{1/\rho} \tag{197}$$

$$\dot{\leq} \sum_{\mathcal{C}^m} P(\mathcal{C}^m)\left(\frac{1}{M_y}\sum_{i=0}^{M_y-1}\sum_{m'\neq m}\sum_{j=0}^{M_y-1}\sum_{\boldsymbol{z}} W(\boldsymbol{z}|\boldsymbol{x}_{mi})\right.$$

$$\times \frac{\exp\{ng(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}}\boldsymbol{z})\}}{\exp\Big\{n\max\Big[g(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}}\boldsymbol{z}),\phi(R_y-\epsilon,\hat{P}_{\boldsymbol{u}_m}\boldsymbol{z}),g(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}}\boldsymbol{z}),\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{z}})\Big]\Big\}}\Bigg)^{1/\rho} \tag{198}$$

$$\stackrel{\text{(a)}}{\doteq} \mathbb{E}\left[\left(\frac{1}{M_y}\sum_{i=0}^{M_y-1}\sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\big\{-n\Omega(\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}\boldsymbol{x}_{mi}\boldsymbol{x}_{m'j}},R_y,R_z)\big\}\right)^{1/\rho}\Bigg|\boldsymbol{u}_m\right], \tag{199}$$

where the passage (a) is due to the following derivation:

$$\begin{aligned}
&\sum_{\boldsymbol{z}}W(\boldsymbol{z}|\boldsymbol{x}_{mi})\exp\Big\{-n\Big[\max\big\{g(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}}\boldsymbol{z}),\phi(R_y-\epsilon,\hat{P}_{\boldsymbol{u}_m}\boldsymbol{z}),\\
&\qquad\qquad\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{z}})\big\}-g(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}}\boldsymbol{z})\Big]_+\Big\}\\
&=\sum_{Q_{Z|UU'XX'}}\sum_{\boldsymbol{z}\in\mathcal{T}(Q_{Z|UU'XX'})}W(\boldsymbol{z}|\boldsymbol{x}_{mi})\exp\Big\{-n\Big[\max\big\{g(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}}\boldsymbol{z}),\phi(R_y-\epsilon,\hat{P}_{\boldsymbol{u}_m}\boldsymbol{z}),\\
&\qquad\qquad\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{z}})\big\}-g(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}}\boldsymbol{z})\Big]_+\Big\} \qquad (200)\\
&\doteq\max_{Q_{Z|UU'XX'}}\exp\Bigg\{n\Bigg(H_Q(Z|U,U',X,X')+\mathbb{E}_Q\log[W_2(Z|X)]\\
&\quad-\Big[\max\big\{g(Q_{UXZ}),\phi(R_y-\epsilon,Q_{UZ}),\psi(R_y-\epsilon,R_z-\epsilon,Q_Z)\big\}-g(Q_{U'X'Z})\Big]_+\Bigg)\Bigg\} \qquad (201)\\
&=\max_{Q_{Z|UU'XX'}}\exp\Bigg\{n\Bigg(H_Q(Z|U,U',X,X')-H_Q(Z|U,X)-D(Q_{Z|UX}\|W_{Z|X}|Q_{UX})\\
&\quad-\Big[\max\big\{g(Q_{UXZ}),\phi(R_y-\epsilon,Q_{UZ}),\psi(R_y-\epsilon,R_z-\epsilon,Q_Z)\big\}-g(Q_{U'X'Z})\Big]_+\Bigg)\Bigg\} \qquad (202)\\
&=\exp\Bigg\{-n\min_{Q_{Z|UU'XX'}}\Bigg(D(Q_{Z|UX}\|W_{Z|X}|Q_{UX})+I_Q(U'X';Z|UX)\\
&\quad+\Big[\max\big\{g(Q_{UXZ}),\phi(R_y-\epsilon,Q_{UZ}),\psi(R_y-\epsilon,R_z-\epsilon,Q_Z)\big\}-g(Q_{U'X'Z})\Big]_+\Bigg)\Bigg\} \qquad (203)\\
&=\exp\Big\{-n\Omega(Q_{UU'XX'},R_y-\epsilon,R_z-\epsilon)\Big\}. \qquad (204)
\end{aligned}$$

In order to continue, we have to define the following enumerators. Let $N_{\boldsymbol{u}_m}(Q)$ denote the number of triplets $(\boldsymbol{x}_{mi},\boldsymbol{u}_{m'},\boldsymbol{x}_{m'j})$, such that $(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j})\in\mathcal{T}(Q_{UU'XX'})$, that is

$$N_{\boldsymbol{u}_m}(Q)=\sum_{m'\neq m}\sum_{i=0}^{M_y-1}\sum_{j=0}^{M_y-1}\mathcal{I}\Big\{(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j})\in\mathcal{T}(Q_{UU'XX'})\Big\}. \tag{205}$$

Let $N_{\boldsymbol{u}_m,\boldsymbol{u}_{m'}}(Q)$ denote the number of pairs $(\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j})$, such that $(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j}) \in \mathcal{T}(Q_{UU'XX'})$, that is

$$N_{\boldsymbol{u}_m,\boldsymbol{u}_{m'}}(Q) = \sum_{i=0}^{M_y-1}\sum_{j=0}^{M_y-1} \mathcal{I}\Big\{(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j}) \in \mathcal{T}(Q_{UU'XX'})\Big\}. \tag{206}$$

Let $N_{\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi}}(Q)$ denote the number of codewords $\boldsymbol{x}_{m'j} \in \mathcal{C}_{m'}$, such that $(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j}) \in \mathcal{T}(Q_{UU'XX'})$, that is

$$N_{\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi}}(Q) = \sum_{j=0}^{M_y-1} \mathcal{I}\Big\{(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi},\boldsymbol{x}_{m'j}) \in \mathcal{T}(Q_{UU'XX'})\Big\}. \tag{207}$$

Let $\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})$ denote the set of cloud-centers $\boldsymbol{u}_{m'}$, such that the joint empirical distribution of $\boldsymbol{u}_{m'}$ with $\boldsymbol{u}_m$ is $Q_{UU'}$. In addition, let $\hat{N}_{\boldsymbol{u}_m}(Q_{UU'})$ be the total amount of them, that is

$$\hat{N}_{\boldsymbol{u}_m}(Q_{UU'}) = \sum_{m'\neq m} \mathcal{I}\Big\{(\boldsymbol{u}_m,\boldsymbol{u}_{m'}) \in \mathcal{T}(Q_{UU'})\Big\}. \tag{208}$$

Let $\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\boldsymbol{u}_{m'}}(Q_{UU'X})$ denote the set of codewords $\boldsymbol{x}_{mi} \in \mathcal{C}_m$, such that the joint empirical distribution of $\boldsymbol{x}_{mi}$ with $(\boldsymbol{u}_m,\boldsymbol{u}_{m'})$ is $Q_{UU'X}$. In addition, let $\tilde{N}_{\boldsymbol{u}_m,\boldsymbol{u}_{m'}}(Q_{UU'X})$ be the total amount of them, that is

$$\tilde{N}_{\boldsymbol{u}_m,\boldsymbol{u}_{m'}}(Q_{UU'X}) = \sum_{i=0}^{M_y-1} \mathcal{I}\Big\{(\boldsymbol{u}_m,\boldsymbol{u}_{m'},\boldsymbol{x}_{mi}) \in \mathcal{T}(Q_{UU'X})\Big\}. \tag{209}$$

Continuing with the conditional expectation of (199), we have

$$\mathbb{E}\left\{\left(\frac{1}{M_y}\sum_{i=0}^{M_y-1}\sum_{m'\neq m}\sum_{j=0}^{M_y-1}\exp\big\{-n\Omega(\hat{P}_{\boldsymbol{u}_m\boldsymbol{u}_{m'}\boldsymbol{x}_{mi}\boldsymbol{x}_{m'j}},R_y,R_z)\big\}\right)^{1/\rho}\middle|\boldsymbol{u}_m\right\} \tag{210}$$

$$= \mathbb{E}\left\{\left(e^{-nR_y}\sum_{Q_{UU'XX'}\in\mathcal{P}} N_{\boldsymbol{u}_m}(Q)e^{-n\Omega(Q,R_y,R_z)}\right)^{1/\rho}\middle|\boldsymbol{u}_m\right\} \tag{211}$$

$$= \mathbb{E}\left\{\left(\sum_{Q_{UU'XX'}\in\mathcal{P}} N_{\boldsymbol{u}_m}(Q)e^{-n[\Omega(Q,R_y,R_z)+R_y]}\right)^{1/\rho}\middle|\boldsymbol{u}_m\right\} \tag{212}$$

$$\stackrel{\text{PD}}{\doteq} \mathbb{E}\left\{\sum_{Q_{UU'XX'}\in\mathcal{P}} [N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\, e^{-n[\Omega(Q,R_y,R_z)+R_y]/\rho}\middle|\boldsymbol{u}_m\right\} \tag{213}$$

$$= \sum_{Q_{UU'XX'}\in\mathcal{P}} \mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\} e^{-n[\Omega(Q,R_y,R_z)+R_y]/\rho}. \tag{214}$$

In order to evaluate the moments $\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\}$, we condition on the set $\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})$ and assume that $\hat{N}_{\boldsymbol{u}_m}(Q_{UU'}) = e^{n\mu}$. For a given $\rho > 1$, let $s \in [1, \rho]$. then,

$$\begin{aligned}
&\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})\right\} \\
&= \mathbb{E}\left\{\left[\sum_{\tilde{\boldsymbol{u}}\in\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})} N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/\rho}\middle|\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})\right\} && (215)\\
&= \mathbb{E}\left\{\left(\left[\sum_{\tilde{\boldsymbol{u}}\in\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})} N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}\right)^{s/\rho}\middle|\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})\right\} && (216)\\
&\overset{\text{PD}}{\leq} \mathbb{E}\left\{\left(\sum_{\tilde{\boldsymbol{u}}\in\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})} \left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}\right)^{s/\rho}\middle|\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})\right\} && (217)\\
&\overset{\text{JI}}{\leq} \left(\mathbb{E}\left\{\sum_{\tilde{\boldsymbol{u}}\in\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})} \left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}\middle|\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})\right\}\right)^{s/\rho} && (218)\\
&= \left(\sum_{\tilde{\boldsymbol{u}}\in\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})} \mathbb{E}\left\{\left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}\right\}\right)^{s/\rho} && (219)
\end{aligned}$$

where JI stands for Jensen's inequality. In turn, in order to evaluate the moments $\mathbb{E}\left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}$, we condition on the set $\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})$ and assume that $\tilde{N}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X}) = e^{n\lambda}$. For a given $s \in [1, \rho]$, let $t \in [1, s]$. then,

$$\begin{aligned}
&\mathbb{E}\left\{\left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}\middle|\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})\right\} \\
&= \mathbb{E}\left\{\left[\sum_{\tilde{\boldsymbol{x}}\in\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})} N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}},\tilde{\boldsymbol{x}}}(Q)\right]^{1/s}\middle|\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})\right\} && (220)\\
&= \mathbb{E}\left\{\left(\left[\sum_{\tilde{\boldsymbol{x}}\in\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})} N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}},\tilde{\boldsymbol{x}}}(Q)\right]^{1/t}\right)^{t/s}\middle|\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})\right\} && (221)\\
&\overset{\text{PD}}{\leq} \mathbb{E}\left\{\left(\sum_{\tilde{\boldsymbol{x}}\in\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})} \left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}},\tilde{\boldsymbol{x}}}(Q)\right]^{1/t}\right)^{t/s}\middle|\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})\right\} && (222)
\end{aligned}$$

$$\overset{\text{JI}}{\leq} \left(\mathbb{E} \left\{ \sum_{\tilde{\boldsymbol{x}} \in \tilde{\mathcal{A}}_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}}(Q_{UU'X})} \left[N_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{x}}}(Q) \right]^{1/t} \middle| \tilde{\mathcal{A}}_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}}(Q_{UU'X}) \right\} \right)^{t/s} \tag{223}$$

$$= \left(\sum_{\tilde{\boldsymbol{x}} \in \tilde{\mathcal{A}}_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}}(Q_{UU'X})} \mathbb{E} \left\{ \left[N_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{x}}}(Q) \right]^{1/t} \right\} \right)^{t/s}. \tag{224}$$

According to eq. (4),

$$\mathbb{E} \left\{ \left[N_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{x}}}(Q) \right]^{1/t} \right\} \doteq \begin{cases} \exp\{n[R_y - I_Q(UX; X'|U')]/t\} & R_y > I_Q(UX; X'|U') \\ \exp\{n[R_y - I_Q(UX; X'|U')]\} & R_y \leq I_Q(UX; X'|U'), \end{cases} \tag{225}$$

and so,

$$\mathbb{E} \left\{ \left[N_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}}(Q) \right]^{1/s} \middle| \tilde{\mathcal{A}}_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}}(Q_{UU'X}) \right\}$$

$$\leq e^{n\lambda t/s} \cdot \left(\mathbb{E} \left\{ \left[N_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{x}}}(Q) \right]^{1/t} \right\} \right)^{t/s} \tag{226}$$

$$\doteq e^{n\lambda t/s} \cdot \begin{cases} \exp\{n[R_y - I_Q(UX; X'|U')]/s\} & R_y > I_Q(UX; X'|U') \\ \exp\{n[R_y - I_Q(UX; X'|U')]t/s\} & R_y \leq I_Q(UX; X'|U') \end{cases} \tag{227}$$

$$= \begin{cases} \exp\{n[\lambda t + R_y - I_Q(UX; X'|U')]/s\} & R_y > I_Q(UX; X'|U') \\ \exp\{n[\lambda + R_y - I_Q(UX; X'|U')]t/s\} & R_y \leq I_Q(UX; X'|U'). \end{cases} \tag{228}$$

After optimizing over t, we get

$$\frac{1}{n} \log \mathbb{E} \left\{ \left[N_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}}(Q) \right]^{1/s} \middle| \tilde{\mathcal{A}}_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}}(Q_{UU'X}) \right\}$$

$$\leq \min_{1 \leq t \leq s} \begin{cases} [\lambda t + R_y - I_Q(UX; X'|U')]/s & R_y > I_Q(UX; X'|U') \\ [\lambda + R_y - I_Q(UX; X'|U')]\, t/s & R_y \leq I_Q(UX; X'|U'), \lambda + R_y > I_Q(UX; X'|U') \\ [\lambda + R_y - I_Q(UX; X'|U')]\, t/s & R_y \leq I_Q(UX; X'|U'), \lambda + R_y \leq I_Q(UX; X'|U') \end{cases} \tag{229}$$

$$= \begin{cases} [\lambda + R_y - I_Q(UX; X'|U')]/s & R_y > I_Q(UX; X'|U') \\ [\lambda + R_y - I_Q(UX; X'|U')]/s & R_y \leq I_Q(UX; X'|U'), \lambda + R_y > I_Q(UX; X'|U') \\ [\lambda + R_y - I_Q(UX; X'|U')]\, s/s & R_y \leq I_Q(UX; X'|U'), \lambda + R_y \leq I_Q(UX; X'|U') \end{cases} \tag{230}$$

$$= \begin{cases} [\lambda + R_y - I_Q(UX; X'|U')]/s & \lambda + R_y > I_Q(UX; X'|U') \\ [\lambda + R_y - I_Q(UX; X'|U')] & \lambda + R_y \leq I_Q(UX; X'|U') \end{cases} \tag{231}$$

$$\triangleq \tilde{E}_1(R_y, Q_{UU'XX'}, s, \lambda). \tag{232}$$

Next, we take the expectation over the set $\tilde{\mathcal{A}}_{\boldsymbol{u}_m, \tilde{\boldsymbol{u}}}(Q_{UU'X})$. Since the expression of (231) is independent of the specific codewords, we only have to average it over the cardinality of

$\tilde{\mathcal{A}}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})$, namely, over $\tilde{N}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})$.

$$\begin{aligned}
&\mathbb{E}_{\tilde{N}}\left[\exp\left\{n\tilde{E}_1(R_y,Q_{UU'XX'},s,\lambda)\right\}\right]\\
&\doteq\sum_{i=0}^{R_y/\epsilon}\Pr\left\{e^{ni\epsilon}\le\tilde{N}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})\le e^{n(i+1)\epsilon}\right\}\cdot\exp\left\{n\tilde{E}_1(R_y,Q_{UU'XX'},s,i\epsilon)\right\}. \qquad (233)
\end{aligned}$$

Notice that the type–class enumerator $\tilde{N}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})$ is a binomial random variable, pertaining to e^{nR_y} trials and probability of success $e^{-nI_Q(X;U'|U)}$. On the one hand, we have [17]

$$\Pr\left\{\tilde{N}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})\le e^{nt}\right\}\doteq\mathcal{I}\left\{R_y\le I_Q(X;U'|U)+t\right\}, \qquad (234)$$

and on the other hand, we have [16, pp. 167–169]

$$\Pr\left\{\tilde{N}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})\ge e^{nt}\right\}\doteq e^{-n\bar{E}(Q_{UU'X},t)}, \qquad (235)$$

where,

$$\bar{E}(Q_{UU'X},t)=\begin{cases}\left[I_Q(X;U'|U)-R_y\right]_+ & \left[R_y-I_Q(X;U'|U)\right]_+\ge t\\ \infty & \left[R_y-I_Q(X;U'|U)\right]_+<t.\end{cases} \qquad (236)$$

First, assume that $R_y>I_Q(X;U'|U)$. We conclude that the event $\left\{e^{nt}\le\tilde{N}_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q_{UU'X})\le e^{n(t+\epsilon)}\right\}$ occurs with very high probability if and only if $\big(R_y-I_Q(X;U'|U)\big)\in[t,t+\epsilon)$, otherwise, its probability has a double exponential decay. Therefore, it turns out that the sum in (233) is dominated by one summand, the one for which $i=\big(R_y-I_Q(X;U'|U)\big)/\epsilon$. Otherwise, if $R_y\le I_Q(X;U'|U)$, then the sum is dominated by the term $i=0$, which has an exponent of $I_Q(X;U'|U)-R_y$. By using the fact that $\epsilon>0$ is arbitrarily small, we get that the expectation in question behaves like

$$\begin{aligned}
\mathbb{E}\left\{\left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}\right\}&\le\mathbb{E}_{\tilde{N}}\left[\exp\left\{n\tilde{E}_1(R_y,Q_{UU'XX'},s,\lambda)\right\}\right] \qquad (237)\\
&\doteq\exp\left\{n\cdot\left(\tilde{E}_1(R_y,Q_{UU'XX'},s,\lambda^*)-\left[I_Q(X;U'|U)-R_y\right]_+\right)\right\}, \qquad (238)
\end{aligned}$$

where $\lambda^*\stackrel{\Delta}{=}[R_y-I_Q(X;U'|U)]_+$. It can also be written as

$$\begin{aligned}
&\mathbb{E}\left\{\left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}\right\}\\
&\le\begin{cases}\exp\left\{n\left([R_y-I_Q(X;U'|U)]_++R_y-I_Q(UX;X'|U')\right)/s-n\left[I_Q(X;U'|U)-R_y\right]_+\right\}\\ \qquad\text{If}\quad [R_y-I_Q(X;U'|U)]_++R_y>I_Q(UX;X'|U')\\ \exp\left\{n\left([R_y-I_Q(X;U'|U)]_++R_y-I_Q(UX;X'|U')-\left[I_Q(X;U'|U)-R_y\right]_+\right)\right\}\\ \qquad\text{If}\quad [R_y-I_Q(X;U'|U)]_++R_y\le I_Q(UX;X'|U').\end{cases}
\end{aligned} \qquad (239)$$

We substitute it back into (219) and get that

$$\begin{aligned}
&\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})\right\} \\
&\leq \left(\sum_{\tilde{\boldsymbol{u}}\in\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})} \mathbb{E}\left\{\left[N_{\boldsymbol{u}_m,\tilde{\boldsymbol{u}}}(Q)\right]^{1/s}\right\}\right)^{s/\rho} \qquad (240)\\
&\leq \begin{cases} \exp\Big\{n\Big([R_y - I_Q(X;U'|U)]_+ + R_y - I_Q(UX;X'|U') \\ \qquad\qquad + s\left(\mu - [I_Q(X;U'|U) - R_y]_+\right)\Big)/\rho\Big\} \\ \qquad \text{If} \quad [R_y - I_Q(X;U'|U)]_+ + R_y > I_Q(UX;X'|U') \\ \exp\left\{n\left(2R_y - I_Q(X;U'|U) - I_Q(UX;X'|U') + \mu\right)s/\rho\right\} \\ \qquad \text{If} \quad [R_y - I_Q(X;U'|U)]_+ + R_y \leq I_Q(UX;X'|U'). \end{cases} \qquad (241)
\end{aligned}$$

Optimizing over $s \in [1,\rho]$, we get, after some algebraic manipulations (the full details can be found in [18])

$$\begin{aligned}
&\frac{1}{n}\log\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})\right\} \\
&\leq \begin{cases} \left[2R_y - I_Q(X;U'|U) - I_Q(UX;X'|U') + \mu\right]/\rho \\ \quad \text{If } 2R_y - I_Q(X;U'|U) - I_Q(UX;X'|U') + \mu > 0 \\ \left([R_y - I_Q(X;U'|U)]_+ + R_y - I_Q(UX;X'|U')\right)/\rho + \mu - [I_Q(X;U'|U) - R_y]_+ \\ \quad \text{If } [R_y - I_Q(X;U'|U)]_+ + R_y > I_Q(UX;X'|U'),\ \mu - [I_Q(X;U'|U) - R_y]_+ \leq 0 \\ \left[2R_y - I_Q(X;U'|U) - I_Q(UX;X'|U') + \mu\right] \\ \quad \text{If } [R_y - I_Q(X;U'|U)]_+ + R_y \leq I_Q(UX;X'|U'), \\ \qquad 2R_y - I_Q(X;U'|U) - I_Q(UX;X'|U') + \mu \leq 0 \end{cases} \qquad (242)\\
&\triangleq \tilde{E}_2(R_y, Q_{UU'XX'}, \rho, \mu). \qquad (243)
\end{aligned}$$

Next, we have to take the expectation over the set $\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})$. Since the expression of (242) is independent of the specific cloud-centers, we only have to average it over the size of the set $\hat{\mathcal{A}}_{\boldsymbol{u}_m}(Q_{UU'})$, namely, over $\hat{N}_{\boldsymbol{u}_m}(Q_{UU'})$. We assess the expectation as

$$\begin{aligned}
&\mathbb{E}_{\hat{N}}\left[\exp\left\{n\tilde{E}_2(R_y, Q_{UU'XX'}, \rho, \mu)\right\}\right] \\
&\doteq \sum_{i=0}^{R_z/\epsilon} \Pr\left\{e^{ni\epsilon} \leq \hat{N}_{\boldsymbol{u}_m}(Q_{UU'}) \leq e^{n(i+1)\epsilon}\right\} \cdot \exp\left\{n\tilde{E}_2(R_y, Q_{UU'XX'}, \rho, i\epsilon)\right\}. \qquad (244)
\end{aligned}$$

The type–class enumerator $\hat{N}_{\boldsymbol{u}_m}(Q_{UU'})$ is a binomial random variable, pertaining to e^{nR_z} trials and success rate of $e^{-nI_Q(U;U')}$. By making a similar large deviations analysis as we did for the sum (233), we get that the expectation in question behaves like

$$\begin{aligned}
\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\} &\leq \mathbb{E}_{\hat{N}}\left[\exp\left\{n\tilde{E}_2(R_y, Q_{UU'XX'}, \rho, \mu)\right\}\right] \qquad (245)\\
&\doteq \exp\left\{n\cdot\left(\tilde{E}_2(R_y, Q_{UU'XX'}, \rho, \mu^*) - \left[I_Q(U;U') - R_z\right]_+\right)\right\}, \qquad (246)
\end{aligned}$$

where $\mu^* \triangleq [R_z - I_Q(U;U')]_+$. Following a few algebraic steps (see [18]), we find that

$$\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\} \leq \begin{cases} e^{n\left(2R_y+R_z-I_Q(UX;U'X')\right)/\rho} \\ \quad \text{If } 2R_y+R_z-I_Q(UX;U'X')>0,\ R_z>I_Q(U;U') \\ e^{n\left(2R_y-I_Q(X;U'|U)-I_Q(UX;X'|U')\right)/\rho-n\left[I_Q(U;U')-R_z\right]} \\ \quad \text{If } 2R_y-I_Q(X;U'|U)-I_Q(UX;X'|U')>0,\ R_z\leq I_Q(U;U') \\ e^{n\left([R_y-I_Q(X;U'|U)]_+ +R_y-I_Q(UX;X'|U')+\rho\left(\mu^*-\left[I_Q(X;U'|U)-R_y\right]_+-\left[I_Q(U;U')-R_z\right]_+\right)\right)/\rho} \\ \quad \text{If } [R_y-I_Q(X;U'|U)]_+ +R_y>I_Q(UX;X'|U'),\ \mu^*-[I_Q(X;U'|U)-R_y]_+\leq 0 \\ e^{n\left(2R_y-I_Q(X;U'|U)-I_Q(UX;X'|U')+\mu^*-\left[I_Q(U;U')-R_z\right]_+\right)} \\ \quad \text{If } [R_y-I_Q(X;U'|U)]_+ +R_y\leq I_Q(UX;X'|U'), \\ \qquad 2R_y-I_Q(X;U'|U)-I_Q(UX;X'|U')+\mu^*\leq 0. \end{cases} \tag{247}$$

Notice that this upper bound on $\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\}$ is, in fact, independent of $\boldsymbol{u}_m$. Thus, by substituting it back into (214), we get an upper bound on $\mathbb{E}\left\{\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho}\middle|\boldsymbol{u}_m\right\}$, which is independent of $\boldsymbol{u}_m$.

7.2 Analysis for the strong user

Thanks to the similarity to the analysis for the weak user, we only provide an outline here, where the full derivation can be found in [18]. For a given code $\mathcal{C}$, the probability of error given that message (m,i) was transmitted is given by the following summation over all wrong codewords in the codebook

$$P_{\mathrm{e}|mi}(\mathcal{C}) = \sum_{(m',j)\neq(m,i)}\sum_{\boldsymbol{y}} W_1(\boldsymbol{y}|\boldsymbol{x}_{mi})\cdot\Pr\left\{\hat{m}=m',\hat{i}=j\middle|\boldsymbol{y}\right\} \tag{248}$$

$$= \sum_{(m',j)\neq(m,i)}\sum_{\boldsymbol{y}} W_1(\boldsymbol{y}|\boldsymbol{x}_{mi}) \times \frac{\exp\{ng(\hat{P}_{\boldsymbol{u}_{m'}\boldsymbol{x}_{m'j}}\boldsymbol{y})\}}{\exp\{ng(\hat{P}_{\boldsymbol{u}_m\boldsymbol{x}_{mi}}\boldsymbol{y})\}+\sum_{(\tilde{m},k)\neq(m,i)}\exp\{ng(\hat{P}_{\boldsymbol{u}_{\tilde{m}}\boldsymbol{x}_{\tilde{m}k}}\boldsymbol{y})\}}. \tag{249}$$

For any $\rho \geq 1$, we take the expectation over the randomness of the incorrect part of the codebook, $\mathcal{C}^{mi} = \mathcal{C}\setminus\{\boldsymbol{u}_m,\boldsymbol{x}_{mi}\}$, where $\boldsymbol{u}_m$ and $\boldsymbol{x}_{mi}$ are kept fixed for now. Then,

$$\begin{aligned}&\mathbb{E}\left\{\left[P_{\mathrm{e}|mi}(\mathcal{C})\right]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right\} \\ &\leq \sum_{Q_{UXX'}\in\mathcal{S}}\mathbb{E}\left(\left[N_{mi}^{\mathrm{IN}}(Q_{UXX'},\mathcal{C})\right]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right)\exp\left\{-n\Upsilon(Q_{UXX'},R_y,R_z)/\rho\right\} \\ &\quad+\sum_{Q_{UU'XX'}\in\mathcal{P}}\mathbb{E}\left(\left[N_{mi}^{\mathrm{OUT}}(Q_{UU'XX'},\mathcal{C})\right]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right)\exp\left\{-n\Omega(Q_{UU'XX'},R_y,R_z)/\rho\right\},\end{aligned} \tag{250}$$

where $N_{mi}^{\text{IN}}(Q_{UXX'},\mathcal{C})$ and $N_{mi}^{\text{OUT}}(Q_{UU'XX'},\mathcal{C})$ are defined in (62) and (63), respectively. Now, it only remains to assess the conditional expectations of (250). For the left one, we have

$$\begin{aligned}&\mathbb{E}\left\{[N_{mi}^{\text{IN}}(Q_{UXX'},\mathcal{C})]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right\}\\&\doteq\begin{cases}\exp\{n[R_y-I_Q(X;X'|U)]/\rho\} & R_y>I_Q(X;X'|U)\\ \exp\{n[R_y-I_Q(X;X'|U)]\} & R_y\le I_Q(X;X'|U),\end{cases}\end{aligned}\tag{251}$$

and for the right conditional expectation,

$$\begin{aligned}&\frac{1}{n}\log\mathbb{E}\left\{[N_{mi}^{\text{OUT}}(Q_{UU'XX'},\mathcal{C})]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right\}\\&\le\begin{cases}\left(R_z+R_y-I_Q(UX;U'X')\right)/\rho\\ \qquad\text{If}\quad R_z+R_y>I_Q(UX;U'X'),\ R_z>I_Q(UX;U')\\ \left(R_y-I_Q(UX;X'|U')\right)/\rho-[I_Q(UX;U')-R_z]\\ \qquad\text{If}\quad R_y>I_Q(UX;X'|U'),\ R_z\le I_Q(UX;U')\\ \left([R_z-I_Q(UX;U')]_+ +R_y-I_Q(UX;X'|U')-[I_Q(UX;U')-R_z]_+\right)\\ \qquad\text{If}\quad [R_z-I_Q(UX;U')]_+ +R_y\le I_Q(UX;X'|U').\end{cases}\end{aligned}\tag{252}$$

Note that the expressions of (251) and (252) are, in fact, independent of $(\boldsymbol{u}_m,\boldsymbol{x}_{mi})$. Substituting them back into (250) provides an upper bound on $\mathbb{E}\left\{\left[P_{\text{e}|mi}(\mathcal{C})\right]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right\}$, which is independent of $(\boldsymbol{u}_m,\boldsymbol{x}_{mi})$.

7.3 Wrapping up

Since the bounds of (250) and (214) are independent of $\boldsymbol{u}_m$ and $\boldsymbol{x}_{mi}$, they also hold for the unconditional expectations, $\mathbb{E}\left[P_{\text{e}|m}(\mathcal{C})\right]^{1/\rho}$ and $\mathbb{E}\left[P_{\text{e}|mi}(\mathcal{C})\right]^{1/\rho}$, i.e.,

$$\mathbb{E}\left\{\left[P_{\text{e}|m}(\mathcal{C})\right]^{1/\rho}\right\}\le\sum_{Q_{UU'XX'}\in\mathcal{P}}\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\}e^{-n[\Omega(Q,R_y,R_z)+R_y]/\rho}\triangleq\boldsymbol{\Delta}_{\text{w}},$$

and

$$\begin{aligned}&\mathbb{E}\left\{\left[P_{\text{e}|mi}(\mathcal{C})\right]^{1/\rho}\right\}\\&\le\sum_{Q_{UXX'}\in\mathcal{S}}\mathbb{E}\left([N_{mi}^{\text{IN}}(Q_{UXX'},\mathcal{C})]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right)\exp\left\{-n\Upsilon(Q_{UXX'},R_y,R_z)/\rho\right\}\\&\quad+\sum_{Q_{UU'XX'}\in\mathcal{P}}\mathbb{E}\left([N_{mi}^{\text{OUT}}(Q_{UU'XX'},\mathcal{C})]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right)\exp\left\{-n\Omega(Q_{UU'XX'},R_y,R_z)/\rho\right\}\triangleq\boldsymbol{\Delta}_{\text{s}}.\end{aligned}$$

According to MI, we get that

$$\Pr\left[\left\{\frac{1}{M_zM_y}\sum_{m=0}^{M_z-1}\sum_{i=0}^{M_y-1}\left[P_{\mathrm{e}|mi}(\mathcal{C})\right]^{1/\rho}>4\cdot\boldsymbol{\Delta}_{\mathrm{s}}\right\}\bigcup\left\{\frac{1}{M_z}\sum_{m=0}^{M_z-1}\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho}>4\cdot\boldsymbol{\Delta}_{\mathrm{w}}\right\}\right]$$

$$\overset{\text{UB}}{\leq}\Pr\left\{\frac{1}{M_zM_y}\sum_{m=0}^{M_z-1}\sum_{i=0}^{M_y-1}\left[P_{\mathrm{e}|mi}(\mathcal{C})\right]^{1/\rho}>4\cdot\boldsymbol{\Delta}_{\mathrm{s}}\right\}+\Pr\left\{\frac{1}{M_z}\sum_{m=0}^{M_z-1}\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho}>4\cdot\boldsymbol{\Delta}_{\mathrm{w}}\right\} \tag{253}$$

$$\overset{\text{MI}}{\leq}\frac{1}{4}+\frac{1}{4}, \tag{254}$$

which means that there exists a code with both

$$\begin{aligned}\frac{1}{M_zM_y}\sum_{m=0}^{M_z-1}\sum_{i=0}^{M_y-1}\left[P_{\mathrm{e}|mi}(\mathcal{C})\right]^{1/\rho}&\leq 4\cdot\boldsymbol{\Delta}_{\mathrm{s}},\\ \frac{1}{M_z}\sum_{m=0}^{M_z-1}\left[P_{\mathrm{e}|m}(\mathcal{C})\right]^{1/\rho}&\leq 4\cdot\boldsymbol{\Delta}_{\mathrm{w}}.\end{aligned} \tag{255}$$

Now, for a given code $\mathcal{C}$, index the message pairs $\{(m,i)\}$ according to decreasing order of $\{P_{\mathrm{e}|mi}(\mathcal{C})\}$, and furthermore, index the clouds $\{m\}$ according to decreasing order of $\{P_{\mathrm{e}|m}(\mathcal{C})\}$. Expurgate half of the pairs that have the highest scores. Then, among the remaining codewords one can find at least $M_z/2$ different clouds, such that each one of them still contains at least $M_y/2$ codewords. Afterwards, expurgate from the original codebook quarter of the clouds that have the highest scores. Intersecting those two remaining sets, at least $M_z/4$ values of m are in common. Denote this good eighth by $\mathcal{C}'$. Then, for this sub–code

$$\begin{aligned}\left[P_{\mathrm{e}|mi}(\mathcal{C}')\right]^{1/\rho}&\leq 16\cdot\boldsymbol{\Delta}_{\mathrm{s}}\quad\forall(m,i),\\ \left[P_{\mathrm{e}|m}(\mathcal{C}')\right]^{1/\rho}&\leq 16\cdot\boldsymbol{\Delta}_{\mathrm{w}}\quad\forall m.\end{aligned} \tag{256}$$

For the weak user, we have that

$$\max_m P_{\mathrm{e}|m}(\mathcal{C}')\leq\left(\sum_{Q_{UU'XX'}\in\mathcal{P}}\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\Big|\boldsymbol{u}_m\right\}e^{-n[\Omega(Q,R_y,R_z)+R_y]/\rho}\right)^{\rho}. \tag{257}$$

Since it holds for every $\rho\geq 1$, the maximal probability of error can be bounded as

$$\max_m P_{\mathrm{e}|m}(\mathcal{C}')\leq\inf_{\rho>1}\left(\sum_{Q_{UU'XX'}\in\mathcal{P}}\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\Big|\boldsymbol{u}_m\right\}e^{-n[\Omega(Q,R_y,R_z)+R_y]/\rho}\right)^{\rho} \tag{258}$$

$$\doteq\inf_{\rho>1}\left(\max_{Q_{UU'XX'}\in\mathcal{P}}\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\Big|\boldsymbol{u}_m\right\}e^{-n[\Omega(Q,R_y,R_z)+R_y]/\rho}\right)^{\rho} \tag{259}$$

$$= \inf_{\rho>1} \max_{Q_{UU'XX'}\in\mathcal{P}} \left(\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\}\right)^{\rho} e^{-n[\Omega(Q,R_y,R_z)+R_y]} \tag{260}$$

$$\leq \lim_{\rho\to\infty} \max_{Q_{UU'XX'}\in\mathcal{P}} \left(\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\}\right)^{\rho} e^{-n[\Omega(Q,R_y,R_z)+R_y]} \tag{261}$$

$$= \max_{Q_{UU'XX'}\in\mathcal{P}} \lim_{\rho\to\infty} \left(\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\middle|\boldsymbol{u}_m\right\}\right)^{\rho} e^{-n[\Omega(Q,R_y,R_z)+R_y]}. \tag{262}$$

Finally,

$$\lim_{\rho\to\infty} \left(\mathbb{E}\left\{[N_{\boldsymbol{u}_m}(Q)]^{1/\rho}\right\}\middle|\boldsymbol{u}_m\right)^{\rho}$$

$$\leq \lim_{\rho\to\infty} \begin{cases} e^{n\left(2R_y+R_z-I_Q(UX;U'X')\right)} \\ \quad \text{If } 2R_y+R_z-I_Q(UX;U'X')>0,\ R_z>I_Q(U;U') \\ e^{n\left(2R_y-I_Q(X;U'|U)-I_Q(UX;X'|U')\right)-n\left[I_Q(U;U')-R_z\right]\cdot\rho} \\ \quad \text{If } 2R_y-I_Q(X;U'|U)-I_Q(UX;X'|U')>0,\ R_z\leq I_Q(U;U') \\ e^{n\left([R_y-I_Q(X;U'|U)]_+ +R_y-I_Q(UX;X'|U')+\rho\left(\mu^*-\left[I_Q(X;U'|U)-R_y\right]_+-\left[I_Q(U;U')-R_z\right]_+\right)\right)} \\ \quad \text{If } [R_y-I_Q(X;U'|U)]_+ +R_y>I_Q(UX;X'|U'),\ \mu^*-[I_Q(X;U'|U)-R_y]_+\leq 0 \\ e^{n\left(2R_y-I_Q(X;U'|U)-I_Q(UX;X'|U')+\mu^*-\left[I_Q(U;U')-R_z\right]_+\right)\cdot\rho} \\ \quad \text{If } [R_y-I_Q(X;U'|U)]_+ +R_y\leq I_Q(UX;X'|U'), \\ \quad\quad 2R_y-I_Q(X;U'|U)-I_Q(UX;X'|U')+\mu^*\leq 0 \end{cases} \tag{263}$$

$$= \begin{cases} e^{n\left(2R_y+R_z-I_Q(UX;U'X')\right)} & I_Q(UX;U'X')<2R_y+R_z,\ I_Q(U;U')<R_z \\ 0 & \text{elsewhere,} \end{cases} \tag{264}$$

and so,

$$\max_m P_{\mathrm{e}|m}(\mathcal{C}')$$

$$\leq \max_{Q_{UU'XX'}\in\mathcal{P}:\ I_Q(UX;U'X')<2R_y+R_z, I_Q(U;U')<R_z} e^{-n[I_Q(UX;U'X')+\Omega(Q,R_y,R_z)-R_y-R_z]} \tag{265}$$

$$= \exp\left\{-n \min_{\substack{Q_{UU'XX'}\in\mathcal{P} \\ I_Q(UX;U'X')<2R_y+R_z, \\ I_Q(U;U')<R_z}} [I_Q(UX;U'X')+\Omega(Q,R_y,R_z)-R_y-R_z]\right\}, \tag{266}$$

which is (48). For the strong user,

$$\max_{m,i} P_{\mathrm{e}|mi}(\mathcal{C}')$$

$$\leq \max\left[\max_{Q_{UXX'}\in\mathcal{S}} \lim_{\rho\to\infty} \left\{\mathbb{E}\left([N_{mi}^{\mathrm{IN}}(Q_{UXX'},\mathcal{C})]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right)\right\}^{\rho} \exp\left\{-n\Upsilon(Q_{UXX'},R_y,R_z)\right\},\right.$$

$$\left.\max_{Q_{UU'XX'}\in\mathcal{P}} \lim_{\rho\to\infty} \left\{\mathbb{E}\left([N_{mi}^{\mathrm{OUT}}(Q_{UU'XX'},\mathcal{C})]^{1/\rho}\middle|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right)\right\}^{\rho} \exp\left\{-n\Omega(Q_{UU'XX'},R_y,R_z)\right\}\right]. \tag{267}$$

It is easy to see that

$$\lim_{\rho\to\infty}\left\{\mathbb{E}\left([N_{mi}^{\text{IN}}(Q_{UXX'},\mathcal{C})]^{1/\rho}\Big|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right)\right\}^{\rho} \tag{268}$$

$$=\lim_{\rho\to\infty}\begin{cases}\exp\{n[R_y-I_Q(X;X'|U)]\} & R_y>I_Q(X;X'|U)\\ \exp\{n[R_y-I_Q(X;X'|U)]\cdot\rho\} & R_y\le I_Q(X;X'|U)\end{cases} \tag{269}$$

$$=\begin{cases}\exp\{n[R_y-I_Q(X;X'|U)]\} & R_y>I_Q(X;X'|U)\\ 0 & R_y\le I_Q(X;X'|U),\end{cases} \tag{270}$$

and similarly,

$$\lim_{\rho\to\infty}\left\{\mathbb{E}\left([N_{mi}^{\text{OUT}}(Q_{UU'XX'},\mathcal{C})]^{1/\rho}\Big|\boldsymbol{u}_m,\boldsymbol{x}_{mi}\right)\right\}^{\rho} \tag{271}$$

$$\le\lim_{\rho\to\infty}\begin{cases}\exp\left\{n\left(R_z+R_y-I_Q(UX;U'X')\right)\right\}\\ \qquad\text{If}\quad R_z+R_y>I_Q(UX;U'X'),\ R_z>I_Q(UX;U')\\ \exp\left\{n\left(R_y-I_Q(UX;X'|U')\right)-n\left[I_Q(UX;U')-R_z\right]\cdot\rho\right\}\\ \qquad\text{If}\quad R_y>I_Q(UX;X'|U'),\ R_z\le I_Q(UX;U')\\ \exp\left\{n\left([R_z-I_Q(UX;U')]_+ +R_y-I_Q(UX;X'|U')-[I_Q(UX;U')-R_z]_+\right)\cdot\rho\right\}\\ \qquad\text{If}\quad [R_z-I_Q(UX;U')]_+ +R_y\le I_Q(UX;X'|U')\end{cases} \tag{272}$$

$$=\begin{cases}\exp\left\{n\left(R_z+R_y-I_Q(UX;U'X')\right)\right\} & R_z+R_y>I_Q(UX;U'X'),\ R_z>I_Q(UX;U')\\ 0 & \text{elsewhere.}\end{cases} \tag{273}$$

Substituting (270) and (273) back into (267) gives the desired result (49)–(51). The proof of Theorem 3 is now complete.

Appendix A

Comparison Between $E_{\mathbf{wu}}^{\mathbf{GLD}}(R_y,R_z)$ and $E_{\mathbf{wu}}^{\mathbf{ML1}}(R_y,R_z)$

In order to compare between the expurgated exponents, let us first compare between $D_t(Q_{XX'})$ [defined in (27)] and $\Omega(Q_{UU'XX'},R_y,R_z)$ [defined in (47)]. Recall that

$$\exp\Big\{-n\Omega(Q_{UU'XX'},R_y,R_z)\Big\} \tag{A.1}$$

$$\doteq\sum_{\boldsymbol{z}}W(\boldsymbol{z}|\boldsymbol{x})\exp\left\{-n\Big[\max\left\{g(\hat{P}_{\boldsymbol{uxz}}),\phi(R_y,\hat{P}_{\boldsymbol{uz}}),\psi(R_y,R_z,\hat{P}_{\boldsymbol{z}})\right\}-g(\hat{P}_{\boldsymbol{u'x'z}})\Big]_+\right\}$$

$$\doteq\sum_{\boldsymbol{z}}W(\boldsymbol{z}|\boldsymbol{x})\cdot\min\left\{1,\frac{\exp\{ng(\hat{P}_{\boldsymbol{u'x'z}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{uxz}})\}+\exp\{n\phi(R_y,\hat{P}_{\boldsymbol{uz}})\}+\exp\{n\psi(R_y,R_z,\hat{P}_{\boldsymbol{z}})\}}\right\}, \tag{A.2}$$

and,

$$\exp\{-nD_t(Q_{XX'})\}\doteq\sum_{\boldsymbol{z}}W_2(\boldsymbol{z}|\boldsymbol{x})\cdot\left[\frac{W_2(\boldsymbol{z}|\boldsymbol{x}')}{W_2(\boldsymbol{z}|\boldsymbol{x})}\right]^t,\qquad t\in[0,1], \tag{A.3}$$

which can also be written using the same notation as

$$\exp\{-nD_t(Q_{XX'})\} \doteq \sum_{\boldsymbol{z}} W_2(\boldsymbol{z}|\boldsymbol{x}) \cdot \left[\frac{e^{ng(\hat{P}_{\boldsymbol{u'x'z}})}}{e^{ng(\hat{P}_{\boldsymbol{uxz}})}}\right]^t, \qquad t \in [0,1]. \tag{A.4}$$

It is easy to argue that for $t \in [0,1]$:

$$\begin{aligned}
&\min\left\{1, \frac{\exp\{ng(\hat{P}_{\boldsymbol{u'x'z}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{uxz}})\} + \exp\{n\phi(R_y, \hat{P}_{\boldsymbol{uz}})\} + \exp\{n\psi(R_y, R_z, \hat{P}_{\boldsymbol{z}})\}}\right\} \\
&\leq \left[\frac{e^{ng(\hat{P}_{\boldsymbol{u'x'z}})}}{e^{ng(\hat{P}_{\boldsymbol{uxz}})}}\right]^t.
\end{aligned} \tag{A.5}$$

To see why this is true, let us distinguish between the cases $g(\hat{P}_{\boldsymbol{u'x'z}}) \leq g(\hat{P}_{\boldsymbol{uxz}})$ and $g(\hat{P}_{\boldsymbol{u'x'z}}) > g(\hat{P}_{\boldsymbol{uxz}})$. In the former case,

$$\begin{aligned}
&\min\left\{1, \frac{\exp\{ng(\hat{P}_{\boldsymbol{u'x'z}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{uxz}})\} + \exp\{n\phi(R_y, \hat{P}_{\boldsymbol{uz}})\} + \exp\{n\psi(R_y, R_z, \hat{P}_{\boldsymbol{z}})\}}\right\} \\
&\leq \frac{\exp\{ng(\hat{P}_{\boldsymbol{u'x'z}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{uxz}})\}} \qquad (A.6) \\
&\leq \left[\frac{\exp\{ng(\hat{P}_{\boldsymbol{u'x'z}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{uxz}})\}}\right]^t. \qquad (A.7)
\end{aligned}$$

In the latter case, the right-hand side of (A.5) exceeds unity, whereas the left-hand side is always less then unity. The conclusion from this observation is that at least for the choice $g(Q) = \sum_{u,x,z} Q(u,x,z) \log W_2(z|x)$, the inequality $\Omega(Q_{UU'XX'}, R_y, R_z) \geq D_t(Q_{XX'})$ holds for any $t \in [0,1]$. Now, we have the following

$$\begin{aligned}
&E_{\mathrm{wu}}^{\mathrm{ML1}}(R_y, R_z) \\
&= \max_{0\leq t\leq 1} \min_{\substack{Q_{UU'XX'}\in\mathcal{P} \\ I_Q(UX;U'X')<2R_y+R_z}} [I_Q(UX;U'X') + D_t(Q_{XX'})] - R_y - R_z \qquad (A.8) \\
&\leq \min_{\substack{Q_{UU'XX'}\in\mathcal{P} \\ I_Q(UX;U'X')<2R_y+R_z}} [I_Q(UX;U'X') + \Omega(Q_{UU'XX'}, R_y, R_z)] - R_y - R_z \qquad (A.9) \\
&\leq \min_{\substack{Q_{UU'XX'}\in\mathcal{P} \\ I_Q(UX;U'X')<2R_y+R_z \\ I_Q(U;U')<R_z}} [I_Q(UX;U'X') + \Omega(Q_{UU'XX'}, R_y, R_z)] - R_y - R_z \qquad (A.10) \\
&= E_{\mathrm{wu}}^{\mathrm{GLD}}(R_y, R_z), \qquad (A.11)
\end{aligned}$$

hence, the GLD–based expurgated bound, $E_{\mathrm{wu}}^{\mathrm{GLD}}(R_y, R_z)$, is at least as tight as the ML–based expurgated bound of Theorem 1.

Appendix B

Comparison Between $E_{\mathbf{wu}}^{\mathbf{GLD}}(R_y, R_z)$ and $E_{\mathbf{wu}}^{\mathbf{RC}}(R_y, R_z)$

The expurgated exponent of the weak user under the GLD is given by

$$E_{\text{wu}}^{\text{GLD}}(R_y, R_z) = \min_{Q_{UU'XX'} \in \tilde{\mathcal{P}}} [I_Q(UX; U'X') + \Omega(Q, R_y, R_z)] - R_y - R_z, \tag{B.1}$$

where,

$$\tilde{\mathcal{P}} \stackrel{\Delta}{=} \{Q_{UU'XX'} : Q \in \mathcal{P},\ I_Q(UX; U'X') < 2R_y + R_z,\ I_Q(U; U') < R_z\}. \tag{B.2}$$

Let us choose $g(Q) = I_Q(UX; Z)$. We then get

$$\phi(R_y, Q_{UZ}) = \max_{\{Q_{X|UZ}:\ I_Q(X;Z|U) \le R_y\}} [g(Q) - I_Q(X; Z|U)] + R_y = I_Q(U; Z) + R_y, \tag{B.3}$$

and

$$\psi(R_y, R_z, Q_Z) = \max_{\{Q_{UX|Z}:\ I_Q(U;Z) \le R_z,\ I_Q(UX;Z) \le R_z + R_y\}} [g(Q) - I_Q(UX; Z)] + R_z + R_y \tag{B.4}$$

$$= R_z + R_y, \tag{B.5}$$

and furthermore

$$\Omega(Q_{UU'XX'}, R_y, R_z) = \min_{Q_{Z|UU'XX'}} \Big(D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + I_Q(U'X'; Z|UX) + \Big[\max\big\{ g(Q_{UXZ}), \phi(R_y, Q_{UZ}), \psi(R_y, R_z, Q_Z) \big\} - g(Q_{U'X'Z}) \Big]_+ \Big) \tag{B.6}$$

$$= \min_{Q_{Z|UU'XX'}} \Big(D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + I_Q(U'X'; Z|UX) + \Big[\max\big\{ I_Q(UX; Z), I_Q(U; Z) + R_y, R_z + R_y \big\} - I_Q(U'X'; Z) \Big]_+ \Big). \tag{B.7}$$

Define

$$\tilde{\mathcal{P}}_z \stackrel{\Delta}{=} \{Q_{UU'XX'Z} : Q \in \mathcal{P},\ I_Q(UX; U'X') < 2R_y + R_z, I_Q(U; U') < R_z\}, \tag{B.8}$$

$$\hat{\mathcal{P}}_z \stackrel{\Delta}{=} \{Q_{U'X'|UXZ} : Q \in \mathcal{P},\ I_Q(UX; U'X') < 2R_y + R_z,\ I_Q(U; U') < R_z\}. \tag{B.9}$$

Substituting into (B.1) gives

$$
\begin{aligned}
&E_{\mathrm{wu}}^{\mathrm{GLD}}(R_y, R_z) \\
&\geq \min_{Q_{UU'XX'} \in \tilde{\mathcal{P}}} \Bigg\{ \min_{Q_{Z|UU'XX'}} \Bigg(D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + I_Q(U'X'; Z|UX) \\
&\qquad + \Big[\max\big\{ I_Q(UX;Z), I_Q(U;Z) + R_y, R_z + R_y \big\} - I_Q(U'X';Z) \Big]_+ \Bigg) + I_Q(UX; U'X') \Bigg\} \\
&\qquad\qquad - R_y - R_z \qquad \text{(B.10)} \\
&= \min_{Q_{UU'XX'Z} \in \tilde{\mathcal{P}}_z} \Bigg\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + I_Q(U'X'; UXZ) \\
&\qquad + \Big[\max\big\{ I_Q(UX;Z), I_Q(U;Z) + R_y, R_z + R_y \big\} - I_Q(U'X';Z) \Big]_+ \Bigg\} - R_y - R_z \qquad \text{(B.11)} \\
&= \min_{Q_{UU'XX'Z} \in \tilde{\mathcal{P}}_z} \Bigg\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + I_Q(U'X'; UX|Z) + I_Q(U'X'; Z) \\
&\qquad + \Big[\max\big\{ I_Q(UX;Z), I_Q(U;Z) + R_y, R_z + R_y \big\} - I_Q(U'X';Z) \Big]_+ \Bigg\} - R_y - R_z \qquad \text{(B.12)} \\
&= \min_{Q_{UU'XX'Z} \in \tilde{\mathcal{P}}_z} \Bigg\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + I_Q(U'X'; UX|Z) \\
&\qquad + \max\big\{ I_Q(UX;Z), I_Q(U'X';Z), I_Q(U;Z) + R_y, R_z + R_y \big\} \Bigg\} - R_y - R_z \qquad \text{(B.13)} \\
&= \min_{Q_{Z|UX}} \min_{Q_{U'X'|UXZ} \in \hat{\mathcal{P}}_z} \Bigg\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + I_Q(U'X'; UX|Z) \\
&\qquad + \max\big\{ I_Q(UX;Z), I_Q(U'X';Z), I_Q(U;Z) + R_y, R_z + R_y \big\} \Bigg\} - R_y - R_z \qquad \text{(B.14)} \\
&= \min_{Q_{Z|UX}} \Bigg\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + \min_{Q_{U'X'|UXZ} \in \hat{\mathcal{P}}_z} \Bigg(I_Q(U'X'; UX|Z) \\
&\qquad + \max\big\{ I_Q(UX;Z), I_Q(U'X';Z), I_Q(U;Z) + R_y, R_z + R_y \big\} \Bigg) \Bigg\} - R_y - R_z \qquad \text{(B.15)} \\
&= \min_{Q_{Z|UX}} \Bigg\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) \\
&\qquad + \max\big\{ I_Q(U;Z) + I_Q(X;Z|U), I_Q(U;Z) + R_y, R_z + R_y \big\} \Bigg\} - R_y - R_z \qquad \text{(B.16)} \\
&= \min_{Q_{Z|UX}} \Big\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) \\
&\qquad + \max\big\{ I_Q(U;Z) + I_Q(X;Z|U) - R_y - R_z, I_Q(U;Z) - R_z, 0 \big\} \Big\} \qquad \text{(B.17)}
\end{aligned}
$$

$$= \min_{Q_{Z|UX}} \Big\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + \big[\max\big\{ I_Q(U;Z) + I_Q(X;Z|U) - R_y - R_z, I_Q(U;Z) - R_z \big\} \big]_+ \Big\} \tag{B.18}$$

$$= \min_{Q_{Z|UX}} \Big\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + \big[I_Q(U;Z) + \max\big\{ I_Q(X;Z|U) - R_y, 0 \big\} - R_z \big]_+ \Big\} \tag{B.19}$$

$$= \min_{Q_{Z|UX}} \Big\{ D(Q_{Z|UX} \| W_{Z|X} | Q_{UX}) + \Big[I_Q(U;Z) + [I_Q(X;Z|U) - R_y]_+ - R_z \Big]_+ \Big\} \tag{B.20}$$

$$= E_{\rm wu}^{\rm RC}(R_y, R_z), \tag{B.21}$$

where the first inequality is because the metric $g(Q) = I_Q(UX;Z)$ may be sub–optimal, and where the inner minimum of (B.15) is attained for a random variable (U', X') which is independent of (U, X, Z), for which both $I_Q(U'X'; UX|Z) = 0$ and $I_Q(U'X'; Z) = 0$, and the constraints $I_Q(U;U') \le R_z$ and $I_Q(UX; U'X') \le 2R_y + R_z$ are obviously satisfied since $I_Q(U;U') = 0$ and $I_Q(UX; U'X') = 0$.

Appendix C

Proving that $\Pr\{\mathcal{K}_\epsilon(m, \boldsymbol{y})\}$ decays double-exponentially fast

Let $\tilde{N}_{\boldsymbol{y}}(Q_{UY})$ denote the number of cloud-centers $\boldsymbol{u}_{m'}$, $m' \neq m$, such that the joint empirical distribution of $\boldsymbol{u}_{m'}$ with $\boldsymbol{y}$ is Q_{UY}, that is

$$\tilde{N}_{\boldsymbol{y}}(Q_{UY}) = \sum_{m' \neq m} \mathcal{I}\left\{ (\boldsymbol{y}, \boldsymbol{u}_{m'}) \in \mathcal{T}(Q_{UY}) \right\}. \tag{C.1}$$

In addition, let $N_{\boldsymbol{y}}(Q)$ denote the number of codewords $\boldsymbol{x}_{m'j} \in \mathcal{C}_{m'}$, for any $m' \neq m$ for which $(\boldsymbol{y}, \boldsymbol{u}_{m'}) \in \mathcal{T}(Q_{UY})$, such that the joint empirical distribution of $\boldsymbol{x}_{m'j}$ with $\boldsymbol{u}_{m'}$ and $\boldsymbol{y}$ is Q_{UXY}, that is

$$N_{\boldsymbol{y}}(Q) = \sum_{m' \neq m} \sum_{j=0}^{M_y - 1} \mathcal{I}\left\{ (\boldsymbol{y}, \boldsymbol{u}_{m'}, \boldsymbol{x}_{m'j}) \in \mathcal{T}(Q_{UXY}) \right\}. \tag{C.2}$$

As before, we start by observing that

$$\Psi_m(\boldsymbol{y}) = \sum_{m' \neq m} \sum_{j=0}^{M_y - 1} \exp\{ng(\hat{P}_{\boldsymbol{u}_{m'} \boldsymbol{x}_{m'j} \boldsymbol{y}})\} = \sum_{Q_{U|Y}} \sum_{Q_{X|UY}} N_{\boldsymbol{y}}(Q) e^{ng(Q)}. \tag{C.3}$$

Thus, considering the randomness of the cloud–centers $\{\boldsymbol{U}_{m'}\}$ and the codewords $\{\boldsymbol{X}_{m'j}\}$,

$$\Pr\Big\{\Psi_m(\boldsymbol{y}) \le \exp\{n\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})\}\Big\} \tag{C.4}$$

$$= \Pr\Bigg\{\sum_{Q_{U|Y}}\sum_{Q_{X|UY}} N_{\boldsymbol{y}}(Q)e^{ng(Q)} \le \exp\{n\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})\}\Bigg\} \tag{C.5}$$

$$\le \Pr\Bigg\{\max_{Q_{U|Y}}\max_{Q_{X|UY}} N_{\boldsymbol{y}}(Q)e^{ng(Q)} \le \exp\{n\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})\}\Bigg\} \tag{C.6}$$

$$= \Pr\bigcap_{Q_{U|Y}}\bigcap_{Q_{X|UY}}\Big\{N_{\boldsymbol{y}}(Q)e^{ng(Q)} \le \exp\{n\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})\}\Big\} \tag{C.7}$$

$$= \Pr\bigcap_{Q_{U|Y}}\bigcap_{Q_{X|UY}}\Big\{N_{\boldsymbol{y}}(Q) \le \exp\{n[\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}}) - g(Q)]\}\Big\}. \tag{C.8}$$

We now show that by the very definition of the function $\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})$, there must exist some conditional distributions $Q^*_{U|Y}$ and $Q^*_{X|UY}$ such that for $Q^* = \hat{P}_{\boldsymbol{y}} \times Q^*_{U|Y} \times Q^*_{X|UY}$, the three inequalities $I_{Q^*}(U;Y) \le R_z-\epsilon$, $I_{Q^*}(U;Y)+I_{Q^*}(X;Y|U) \le R_z+R_y-2\epsilon$ and $R_z+R_y-I_{Q^*}(U;Y)-I_{Q^*}(X;Y|U)-2\epsilon \ge \psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}}) - g(Q^*)$ hold. Assume conversely, that for every $Q_{U|Y}$ and every $Q_{X|UY}$, which defines $Q = \hat{P}_{\boldsymbol{y}} \times Q_{U|Y} \times Q_{X|UY}$, either $I_Q(U;Y) > R_z-\epsilon$ or $I_Q(U;Y)+I_Q(X;Y|U) > R_z+R_y-2\epsilon$ or $R_z+R_y-I_Q(U;Y)-I_Q(X;Y|U)-2\epsilon < \psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}}) - g(Q)$, which means that for every Q

$$R_z-\epsilon \tag{C.9}$$

$$< \max\Big\{I_Q(U;Y), I_Q(U;Y)+I_Q(X;Y|U)-R_y+\epsilon,$$
$$I_Q(U;Y)+I_Q(X;Y|U)-R_y+\epsilon+\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})-g(Q)\Big\} \tag{C.10}$$

$$= I_Q(U;Y) + \max\Big\{0, I_Q(X;Y|U)-R_y+\epsilon,$$
$$I_Q(X;Y|U)-R_y+\epsilon+\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})-g(Q)\Big\} \tag{C.11}$$

$$= I_Q(U;Y) + \Big[\max\Big\{I_Q(X;Y|U)-R_y+\epsilon,$$
$$I_Q(X;Y|U)-R_y+\epsilon+\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})-g(Q)\Big\}\Big]_+ \tag{C.12}$$

$$= I_Q(U;Y) + \Big[I_Q(X;Y|U)-R_y+\epsilon+\max\Big\{0, \psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})-g(Q)\Big\}\Big]_+ \tag{C.13}$$

$$= I_Q(U;Y) + \Big[I_Q(X;Y|U)-R_y+\epsilon+\big[\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})-g(Q)\big]_+\Big]_+ \tag{C.14}$$

or, in other words, that for every $Q_{U|Y}$ and $Q_{X|UY}$ there exists $(t,s) \in [0,1]^2$ such that

$$R_z-\epsilon < I_Q(U;Y) + t\Big[I_Q(X;Y|U)-R_y+\epsilon+s\big[\psi(R_y-\epsilon, R_z-\epsilon, \hat{P}_{\boldsymbol{y}})-g(Q)\big]\Big], \tag{C.15}$$

or equivalently,

$$\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}}) \tag{C.16}$$

$$> \max_{Q_{U|Y}}\max_{Q_{X|UY}}\min_{(t,s)\in[0,1]^2} g(Q)+\frac{1}{s}\left(\frac{R_z-I_Q(U;Y)-\epsilon}{t}+R_y-I_Q(X;Y|U)-\epsilon\right) \tag{C.17}$$

$$= \max_{\{Q_{U|Y}:\ I_Q(U;Y)\le R_z-\epsilon\}}\max_{Q_{X|UY}}\min_{s\in[0,1]} g(Q)+\frac{1}{s}\Big(R_z+R_y-I_Q(UX;Y)-2\epsilon\Big) \tag{C.18}$$

$$= \max_{\{Q_{U|Y}:\ I_Q(U;Y)\le R_z-\epsilon\}}\ \max_{\{Q_{X|UY}:\ I_Q(UX;Y)\le R_z+R_y-2\epsilon\}}[g(Q)-I_Q(UX;Y)]+R_z+R_y-2\epsilon \tag{C.19}$$

$$= \max_{\{Q_{UX|Y}:\ I_Q(U;Y)\le R_z-\epsilon,\ I_Q(UX;Y)\le R_z+R_y-2\epsilon\}}[g(Q)-I_Q(UX;Y)]+R_z+R_y-2\epsilon \tag{C.20}$$

$$\equiv \psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}}), \tag{C.21}$$

which is a contradiction. Let then $Q^*_{U|Y}$ and $Q^*_{X|UY}$ be two conditional distributions as defined above. Then,

$$\Pr\bigcap_{Q_{U|Y}}\bigcap_{Q_{X|UY}}\Big\{N_{\boldsymbol{y}}(Q)\le\exp\{n[\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}})-g(Q)]\}\Big\} \tag{C.22}$$

$$\le \Pr\Big\{N_{\boldsymbol{y}}(Q^*)\le\exp\{n[\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}})-g(Q^*)]\}\Big\}. \tag{C.23}$$

Now, remember that $I_{Q^*}(U;Y)\le R_z-\epsilon$, $I_{Q^*}(UX;Y)\le R_z+R_y-2\epsilon$ and $R_z+R_y-I_{Q^*}(UX;Y)-2\epsilon\ge\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}})-g(Q^*)$. Next, given that $\tilde{N}_{\boldsymbol{y}}(Q^*_{UY})=e^{n\lambda}$, $N_{\boldsymbol{y}}(Q^*)$ is a binomial random variable with $e^{n(R_y+\lambda)}$ trials and success rate of the exponential order of $e^{-nI_{Q^*}(X;Y|U)}$. By the Chernoff bound, the probability of (C.23) is upper bounded by

$$\Pr\Big\{N_{\boldsymbol{y}}(Q^*)\le\exp\{n[\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}})-g(Q^*)]\}\Big|\tilde{N}_{\boldsymbol{y}}(Q^*_{UY})=e^{n\lambda}\Big\} \tag{C.24}$$

$$\le\begin{cases}\exp\Big\{-e^{n(R_y+\lambda)}D(e^{-an}\|e^{-bn})\Big\} & \lambda\in\mathcal{G}\\ 1 & \lambda\in\mathcal{G}^c,\end{cases} \tag{C.25}$$

where $a=R_y+\lambda+g(Q^*)-\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}})$, $b=I_{Q^*}(X;Y|U)$,

$$\mathcal{G}=\Big\{\lambda\ \Big|\ I_{Q^*}(X;Y|U)\le R_y-\epsilon+\lambda+g(Q^*)-\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}})\Big\}, \tag{C.26}$$

and where $D(\alpha\|\beta)$, for $\alpha,\beta\in[0,1]$, is the binary divergence function. Noting that

$$a-b=R_y+\lambda-I_{Q^*}(X;Y|U)+g(Q^*)-\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}}) \tag{C.27}$$

$$=R_z+R_y-I_{Q^*}(UX;Y)+g(Q^*)-\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}})+\lambda+I_{Q^*}(U;Y)-R_z \tag{C.28}$$

$$\ge 2\epsilon+\lambda+I_{Q^*}(U;Y)-R_z \tag{C.29}$$

$$\triangleq\Theta(\lambda,R_z), \tag{C.30}$$

the binary divergence can be lower bounded as:

$$D(e^{-an}\|e^{-bn}) \geq e^{-bn}\{1-e^{-(a-b)n}[1+n(a-b)]\} \tag{C.31}$$

$$\geq e^{-nI_{Q^*}(X;Y|U)}\Big[1-e^{-n\Theta(\lambda,R_z)}(1+n\Theta(\lambda,R_z))\Big], \tag{C.32}$$

where in the second inequality, we have used the decreasing monotonicity of the function $f(t) = (1+t)e^{-t}$ for $t \geq 0$. Thus,

$$\exp\Big\{-e^{n(R_y+\lambda)}D(e^{-an}\|e^{-bn})\Big\} \tag{C.33}$$

$$\leq \exp\Big\{-e^{n(R_y+\lambda)}\cdot e^{-nI_{Q^*}(X;Y|U)}\Big[1-e^{-n\Theta(\lambda,R_z)}(1+n\Theta(\lambda,R_z))\Big]\Big\} \tag{C.34}$$

$$= \exp\Big\{-e^{n(R_z+R_y-I_{Q^*}(UX;Y)+I_{Q^*}(U;Y)-R_z+\lambda)}\cdot\Big[1-e^{-n\Theta(\lambda,R_z)}(1+n\Theta(\lambda,R_z))\Big]\Big\} \tag{C.35}$$

$$\leq \exp\Big\{-e^{n(2\epsilon+I_{Q^*}(U;Y)-R_z+\lambda)}\cdot\Big[1-e^{-n\Theta(\lambda,R_z)}(1+n\Theta(\lambda,R_z))\Big]\Big\} \tag{C.36}$$

$$= \exp\Big\{-e^{n\Theta(\lambda,R_z)}\cdot\Big[1-e^{-n\Theta(\lambda,R_z)}(1+n\Theta(\lambda,R_z))\Big]\Big\} \tag{C.37}$$

$$= \exp\Big\{-e^{n\Theta(\lambda,R_z)}+n\Theta(\lambda,R_z)+1\Big\}. \tag{C.38}$$

Let us continue from (C.25) and upper bound as follows

$$\Pr\Big\{N_{\boldsymbol{y}}(Q^*) \leq \exp\{n[\psi(R_y-\epsilon,R_z-\epsilon,\hat{P}_{\boldsymbol{y}})-g(Q^*)]\}\Big|\tilde{N}_{\boldsymbol{y}}(Q^*_{UY}) = e^{n\lambda}\Big\} \tag{C.39}$$

$$\leq \exp\Big\{-e^{n(R_y+\lambda)}D(e^{-an}\|e^{-bn})\Big\}\cdot\mathcal{I}\{\mathcal{G}\}+\mathcal{I}\{\mathcal{G}^c\} \tag{C.40}$$

$$\leq \exp\Big\{-e^{n\Theta(\lambda,R_z)}+n\Theta(\lambda,R_z)+1\Big\}+\mathcal{I}\{\mathcal{G}^c\}. \tag{C.41}$$

Then, the next step will be to evaluate the expectation over the enumerator $\tilde{N}_{\boldsymbol{y}}(Q^*_{UY})$.

$$\mathbb{E}\Big[\exp\Big\{-e^{n\Theta(\lambda,R_z)}+n\Theta(\lambda,R_z)+1\Big\}+\mathcal{I}\{\mathcal{G}^c\}\Big] \tag{C.42}$$

$$\doteq \sum_{i=0}^{R_z/\delta}\Pr\Big\{e^{ni\delta} \leq \tilde{N}_{\boldsymbol{y}}(Q^*_{UY}) \leq e^{n(i+1)\delta}\Big\}$$

$$\times\Big[\exp\Big\{-e^{n\Theta(i\delta,R_z)}+n\Theta(i\delta,R_z)+1\Big\}+\mathcal{I}\{\mathcal{G}^c\}\Big]. \tag{C.43}$$

Moving forward, notice that the type class enumerator $\tilde{N}_{\boldsymbol{y}}(Q^*_{UY})$ is a sum of e^{nR_z} independent binary random variables, each one with an expectation of $e^{-nI_{Q^*}(U;Y)}$. On the one hand, we have

$$\Pr\Big\{\tilde{N}_{\boldsymbol{y}}(Q^*_{UY}) \leq e^{nt}\Big\} \doteq \mathcal{I}\Big\{R_z \leq I_{Q^*}(U;Y)+t\Big\}, \tag{C.44}$$

and on the other hand, we have

$$\Pr\Big\{\tilde{N}_{\boldsymbol{y}}(Q^*_{UY}) \geq e^{nt}\Big\} \doteq e^{-n\hat{E}(Q^*_{UY},t)} \tag{C.45}$$

where,

$$\hat{E}(Q^*_{UY},t) = \begin{cases} \big[I_{Q^*}(U;Y) - R_z\big]_+ & \big[R_z - I_{Q^*}(U;Y)\big]_+ \geq t \\ \infty & \big[R_z - I_{Q^*}(U;Y)\big]_+ < t. \end{cases} \tag{C.46}$$

Since $R_z - \epsilon \geq I_{Q^*}(U;Y)$, we conclude that the event $\big\{e^{nt} \leq \tilde{N}_{\boldsymbol{y}}(Q^*_{UY}) \leq e^{n(t+\delta)}\big\}$ occurs with very high probability if and only if $\big(R_z - \epsilon - I_{Q^*}(U;Y)\big) \in [t, t+\delta)$, otherwise, its probability has a double exponential decay. Therefore, it turns out that the sum in (C.43) is dominated by one summand only, the one for which $i = \big(R_z - \epsilon - I_{Q^*}(U;Y)\big)/\delta$. For this value of i, we have that

$$\Theta(i\delta, R_z) = 2\epsilon + i\delta + I_{Q^*}(U;Y) - R_z = \epsilon, \tag{C.47}$$

and, in addition,

$$\mathcal{I}\big\{\mathcal{G}^c\big\} = \mathcal{I}\big\{I_{Q^*}(X;Y|U) > R_y - \epsilon + i\delta + g(Q^*) - \psi(R_y - \epsilon, R_z - \epsilon, \hat{P}_{\boldsymbol{y}})\big\} \tag{C.48}$$

$$= \mathcal{I}\big\{I_{Q^*}(UX;Y) > R_y + R_z - 2\epsilon + g(Q^*) - \psi(R_y - \epsilon, R_z - \epsilon, \hat{P}_{\boldsymbol{y}})\big\} \tag{C.49}$$

$$= 0, \tag{C.50}$$

thanks to the fact that $R_z + R_y - I_{Q^*}(UX;Y) - 2\epsilon \geq \psi(R_y - \epsilon, R_z - \epsilon, \hat{P}_{\boldsymbol{y}}) - g(Q^*)$. By using the fact that $\delta > 0$ is arbitrarily small, we get that the expectation in question behaves like

$$\mathbb{E}\bigg[\exp\Big\{-e^{n\Theta(\lambda,R_z)} + n\Theta(\lambda,R_z) + 1\Big\} + \mathcal{I}\big\{\mathcal{G}^c\big\}\bigg] \doteq \exp\big\{-e^{n\epsilon} + n\epsilon + 1\big\}. \tag{C.51}$$

References

[1] T. M. Cover, "Broadcast channels," *IEEE Trans. on Inform. Theory*, vol. 18, no. 1, pp. 2–14, January 1972.

[2] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. on Inform. Theory*, vol. IT-19, pp. 197-207, March 1973.

[3] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. on Inform. Theory*, vol. IT-23, pp. 60-64, January 1977.

[4] J. Körner and A. Sgarro, "Universally attainable error exponents for broadcast channels with degraded message sets," *IEEE Trans. on Inform. Theory*, vol. 26, no.6, pp. 670-679, November 1980.

[5] Y. Kaspi and N. Merhav, "Error exponents for broadcast channels with degraded message sets," *IEEE Trans. on Inform. Theory*, vol. 57, no. 1, pp. 101–123, January 2011.

[6] R. Averbuch and N. Merhav, "Exact random coding exponents and universal decoders for the asymmetric broadcast channel," submitted for publication, February 2017. Available on–line at `https://arxiv.org/pdf/1702.08003.pdf`

[7] I. Csiszár, J. Körner, and K. Marton, "A new look at the error exponent of discrete memoryless channels," in *Proc. ISIT '77*, p. 107 (abstract), Cornell University, Ithaca, New York, 1977.

[8] J. Scarlett, L. Peng, N. Merhav, A. Martinèz, and A. Guillén i Fàbregas, "Expurgated random–coding ensembles: exponents, refinements and connections," *IEEE Trans. on Inform. Theory*, vol. 60, no. 8, pp. 4449–4462, August 2014.

[9] N. Merhav, "List decoding – random coding exponents and expurgated exponents," *IEEE Trans. on Inform. Theory*, vol. 60, no. 11, pp. 6749–6759, November 2014.

[10] L. Weng, S. S. Pradhan, and A. Anastasopoulos, "Error exponent regions for gaussian broadcast and multiple–access channels," *IEEE Trans. on Inform. Theory*, vol. 54, no. 7, pp. 2919–2942, July 2008.

[11] A. Nazari, A. Anastasopoulos, and S. S. Pradhan, "Error exponent for multiple–access channels: lower bounds," *IEEE Trans. on Inform. Theory*, vol. 60, no. 9, pp. 5095–5115, September 2014.

[12] J. Scarlett, A. Martinèz, and A. Guillén i Fàbregas, "The likelihood decoder: error exponents and mismatch," *Proc. 2015 IEEE International Symposium on Information Theory (ISIT 2015)*, pp. 86–90, Hong Kong, June 2015.

[13] N. Merhav, "The generalized stochastic likelihood decoder: random coding and expurgated bounds," *IEEE Trans. on Inform. Theory*, vol. 63, no. 8, pp. 5039–5051, August

2017. See also a correction at *IEEE Trans. on Inform. Theory*, vol. 63, no. 10, pp. 6827–6829, October 2017.

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

[15] R. G. Gallager, *Information Theory and Reliable Communication*, New York, Wiley 1968.

[16] N. Merhav,"Statistical physics and information theory," *Foundations and Trends in Communications and Information Theory,* vol. 6, nos. 12, pp. 1–212, 2009.

[17] N. Merhav, "Exact random coding error exponents of optimal bin index decoding," *IEEE Trans. on Inform. Theory*, vol. 60, no. 10, pp. 6024–6031, October 2014.

[18] R. Averbuch, N. Weinberger, and N. Merhav, "Expurgated bounds for the asymmetric broadcast channel," available on–line at `https://arxiv.org/pdf/1711.10299.pdf`

[19] I. Csiszár and J. Körner, "Graph decomposition: a new key to coding theorems," *IEEE Trans. on Inform. Theory*, vol. 27, no. 1, pp. 5–12, January 1981.