Foundations and Trends® in Communications and Information Theory

Information-Theoretic Foundations of Mismatched Decoding

Suggested Citation: Jonathan Scarlett, Albert Guillén i Fàbregas, Anelia Somekh-Baruch and Alfonso Martinez (2020), "Information-Theoretic Foundations of Mismatched Decoding", Foundations and Trends® in Communications and Information Theory: Vol. 17, No. 2–3, pp 149–400. DOI: 10.1561/0100000101.

Jonathan Scarlett
National University of Singapore, Singapore
scarlett@comp.nus.edu.sg

Albert Guillén i Fàbregas
ICREA, Universitat Pompeu Fabra, Spain
University of Cambridge, UK
guillen@ieee.org

Anelia Somekh-Baruch
Bar-Ilan University, Israel
somekha@biu.ac.il

Alfonso Martinez
Universitat Pompeu Fabra, Spain
alfonso.martinez@ieee.org

This article may be used only for the purpose of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval.

now
the essence of knowledge
Boston — Delft

Contents

Notation **151**

1 Introduction **154**
1.1 Problem Setup . . . 155
1.2 Applications of Mismatched Decoding . . . 158
1.3 Other Channel Uncertainty Models . . . 164
1.4 Outline of the Monograph . . . 166

2 Discrete Memoryless Channels **170**
2.1 Introduction . . . 170
2.2 Properties of the Mismatch Capacity . . . 171
2.3 Achievable Rates: GMI and LM Rate . . . 173
2.4 Examples . . . 180
2.5 Properties of the GMI and LM Rate . . . 188
2.6 Proofs of Achievable Rates . . . 200
2.7 Multi-Letter Improvements . . . 208

3 Continuous-Alphabet Memoryless Channels **210**
3.1 Introduction . . . 210
3.2 Problem Setup . . . 211
3.3 i.i.d. Random Coding and the GMI . . . 212
3.4 Cost-Constrained Random Coding and the LM Rate . . . 214

3.5 Cost-Constrained Random Coding with Multiple Auxiliary Costs . . . 216
3.6 Proofs of Achievable Rates . . . 218
3.7 Ensemble Tightness Results . . . 220
3.8 Examples . . . 224

4 Mismatch in Rate Distortion Theory 238
4.1 Introduction . . . 238
4.2 Problem Setup . . . 239
4.3 Achievability Result . . . 242
4.4 Examples . . . 244
4.5 Achievability Proof . . . 247
4.6 Multi-Letter Improvement and Converse . . . 251
4.7 Mismatched Random Codebooks with Optimal Encoding . . . 253
4.8 Gaussian Compression for Non-Gaussian Sources . . . 262

5 Multiple-Access Channels 266
5.1 Introduction . . . 266
5.2 Problem Setup . . . 267
5.3 Achievable Rate Region . . . 268
5.4 Examples . . . 271
5.5 Proof of Achievable Rate Region . . . 274
5.6 Ensemble Tightness . . . 279
5.7 Dual Expressions and Continuous Alphabets . . . 284

6 Multi-User Coding Techniques for Single-User Channels 288
6.1 Introduction . . . 288
6.2 Expurgated Parallel Coding . . . 289
6.3 Superposition Coding . . . 293
6.4 Refined Superposition Coding . . . 304

7 Error Exponents 317
7.1 Introduction . . . 317
7.2 Random-Coding Exponents for DMCs . . . 318
7.3 Random-Coding Exponents with Continuous Alphabets . . . 323
7.4 Expurgated Exponents . . . 326

7.5 Attaining the Best of Both Exponents . . . 328
7.6 Numerical Example . . . 333
7.7 Error Exponents for Multi-User Coding Techniques . . . 334

8 Upper Bounds on the Mismatch Capacity 338
8.1 Introduction . . . 338
8.2 Initial Results . . . 339
8.3 The Csiszár–Narayan Conjecture . . . 341
8.4 Other Decoding Rules . . . 351
8.5 Further Multi-Letter Upper Bounds . . . 355
8.6 A Single-Letter Upper Bound . . . 358

9 Overview of Other Topics 365
9.1 Alternative Achievable Rates and Capacity Notions . . . 365
9.2 Continuous and Fading Channels . . . 366
9.3 Refined Asymptotics . . . 367
9.4 Modified Channel Coding Settings . . . 368
9.5 Other Decoders . . . 369
9.6 Mismatch in Practical Codes . . . 371

10 Conclusion 372

Appendix 377

A Appendix 378
A.1 Continuity Arguments . . . 378
A.2 Other Omitted Proofs . . . 386

Acknowledgments 391

References 392

Information-Theoretic Foundations of Mismatched Decoding

Jonathan Scarlett[1], Albert Guillén i Fàbregas[2], Anelia Somekh-Baruch[3] and Alfonso Martinez[4]

[1] *National University of Singapore, Singapore; scarlett@comp.nus.edu.sg*
[2] *ICREA, Universitat Pompeu Fabra, Spain, and University of Cambridge, UK; guillen@ieee.org*
[3] *Bar-Ilan University, Israel; somekha@biu.ac.il*
[4] *Universitat Pompeu Fabra, Spain; alfonso.martinez@ieee.org*

ABSTRACT

Shannon's channel coding theorem characterizes the maximal rate of information that can be reliably transmitted over a communication channel when optimal encoding and decoding strategies are used. In many scenarios, however, practical considerations such as channel uncertainty and implementation constraints rule out the use of an optimal decoder. The mismatched decoding problem addresses such scenarios by considering the case that the decoder cannot be optimized, but is instead fixed as part of the problem statement. This problem is not only of direct interest in its own right, but also has close connections with other long-standing theoretical problems in information theory.

In this monograph, we survey both classical literature and recent developments on the mismatched decoding problem, with an emphasis on achievable random-coding rates for memoryless channels. We present two widely-considered

Jonathan Scarlett, Albert Guillén i Fàbregas, Anelia Somekh-Baruch and Alfonso Martinez (2020), "Information-Theoretic Foundations of Mismatched Decoding", Foundations and Trends® in Communications and Information Theory: Vol. 17, No. 2–3, pp 149–400. DOI: 10.1561/0100000101.

achievable rates known as the generalized mutual information (GMI) and the LM rate, and overview their derivations and properties. In addition, we survey several improved rates via multi-user coding techniques, as well as recent developments and challenges in establishing upper bounds on the mismatch capacity, and an analogous mismatched encoding problem in rate-distortion theory. Throughout the monograph, we highlight a variety of applications and connections with other prominent information theory problems.

Notation

(Introduced in Section 1)

$\mathcal{X}$, $\mathcal{Y}$	Input and output alphabets
m, $\hat{m}$	Message and its estimate
M	Number of codewords
n	Block length
R	Coding rate
W, W^n	Channel transition law and its n-letter extension
q, q^n	Decoding metric and its n-letter extension
$\mathcal{C}$	Codebook
$\boldsymbol{x}^{(m)}$	m-th codeword
p_{e}	Average error probability
$p_{\mathrm{e,max}}$	Maximal error probability
C_{M}	Mismatch capacity
C	Matched capacity

(Introduced in Section 2)

H_2	Binary entropy function
$\boldsymbol{X}$, $\boldsymbol{Y}$	Transmitted codeword and received sequence
$\overline{\boldsymbol{X}}$	Non-transmitted codeword
$P_{\boldsymbol{X}}$	Codeword distribution for random coding
$\widehat{P}_{\boldsymbol{x}}$	Empirical distribution of $\boldsymbol{x}$
I_{GMI}, I_{LM}	Generalized mutual information (GMI) and LM rate
C_{GMI}, C_{LM}	Rates with optimized input distributions
Q_X	Input distribution
$\mathcal{P}$	Set of all distributions on a given alphabet
$\mathcal{P}_n$	Set of all empirical distributions on a given alphabet
$\widetilde{P}_{XY}$	Auxiliary distribution in primal rates
s, a	Auxiliary parameters in dual rates
b	Auxiliary function on the output alphabet
$\mathcal{T}^n$	Type class
$Q_{X,n}$	Type approximating Q_X
$\bar{p}_{\mathrm{e}}$	Random-coding error probability
i_s^n	Information density quantity used in proofs
$C_{\mathrm{GMI}}^{(k)}$, $C_{\mathrm{LM}}^{(k)}$	Multi-letter achievable rates
$C_{\mathrm{GMI}}^{(\infty)}$, $C_{\mathrm{LM}}^{(\infty)}$	Limiting multi-letter achievable rates

(Introduced in Section 3)

c, Γ	System cost function and threshold
a, a_l	Auxiliary cost functions
c^n, a^n, a_l^n	Additive multi-letter extensions of cost functions
ϕ_c, ϕ_a, ϕ_l	Means of cost functions
Ω_n	Normalizing constant for cost-constrained ensemble
$\mathcal{D}_n$	Constraint set for cost-constrained ensemble
I'_{LM}	Fixed-cost LM rate
r_l, $\bar{r}_l$	Auxiliary parameters for cost-constrained ensemble
L	Number of auxiliary costs
μ, σ^2	Noise mean and variance in additive channels

(Introduced in Section 4)

Π_X, Π_X^n	Source distribution and n-letter extension
$\widehat{\boldsymbol{x}}^{(m)}$	Codeword in rate-distortion theory
d_0, d_1	Encoding metric and true distortion measure
d_0^n, d_1^n	Additive n-letter extensions of distortion functions
D_1, D_1^*	Distortion threshold and distortion-rate function
$Q_{\widehat{X}}$	Auxiliary distribution in rate-distortion theory
$\overline{D}_1$, $\overline{D}_1^{(k)}$	Achievable distortion and multi-letter version
$\widetilde{\mathcal{P}}$	Constraint set in achievable distortion expression
d, D	Distortion measure and level when $d_0 = d_1$
R^*_{Matched}, D^*_{Matched}	Matched rate-distortion function and distortion-rate function
$D_{\min}$, D_{prod}	Extreme values of distortion level
$\overline{R}_{\text{iid}}$, $\overline{R}_{\text{cc}}$	Random coding rate-distortion functions
σ^2	Source variance

(Introduced in Section 5)

M_1, M_2	Multiple-access channel codebook sizes
R_1, R_2	Multiple-access channel rates
Q_1, Q_2	Multiple-access channel input distributions
$P_{\boldsymbol{X}_1}$, $P_{\boldsymbol{X}_2}$	Multiple-access channel codeword distributions

(Introduced in Section 6)

M_0, M_1, $\{M_{1u}\}$	Superposition coding codebook sizes
R_0, R_1, $\{R_{1u}\}$	Superposition coding rates
$\{n_u\}$	Refined superposition coding sub-block lengths
Q_{UX}	Superposition coding input distribution
$P_{\boldsymbol{U}\boldsymbol{X}}$	Superposition coding codeword distribution
$\mathcal{U}$	Auxiliary alphabet for superposition coding
$\boldsymbol{U}$	Auxiliary codeword for superposition coding
$p_{\text{e},\nu}$, $\bar{p}_{\text{e},\nu}$	Multi-user coding error probabilities

(Introduced in Section 7)

$E_{\mathrm{r}}^{\mathrm{iid}}$, E_0^{iid}	i.i.d. exponent and Gallager function
$E_{\mathrm{r}}^{\mathrm{cc}}$, E_0^{cc}	Constant-composition exponent and Gallager function
$E_{\mathrm{r}}^{\mathrm{cost}}$, E_0^{cost}	Cost-constrained exponent and Gallager function
$E_{\mathrm{ex}}^{\mathrm{cc}}$, $E_{\mathrm{x}}^{\mathrm{cc}}$	Constant-composition expurgated exponent and Gallager function
ρ	Dual error exponent parameter
E_{CK}	Csiszár-Körner exponent
E_{RGV}	Random Gilbert-Varshamov exponent
d, Δ	RGV distance function and parameter

(Introduced in Section 8)

ϵ_k	Error probability with block length k
η_k	Minimal difference of k-letter log-metric values
$q_{\max}$	Maximal metric value
$q_{\max}^k(y^k)$	Maximal metric value for fixed output sequence
B	Upper bound on $\lvert \log q(x,y) \rvert$
$\mathcal{A}_k$	Output vectors with a unique metric maximizer
Φ_k	Conditional probability of random codeword exceeding metric value
$\mathcal{X}_q^*$	Set of inputs with maximal metric difference
$\mathcal{M}_{\max}$, $\mathcal{M}_{\max,n}$	Set of maximal conditional joint distributions and type-based variant
$\overline{R}$	Single-letter mismatch capacity upper bound
$\mathcal{G}_{\boldsymbol{x}}$	Bipartite graph associated with two channels
$\mathcal{E}_{\boldsymbol{x}}$	Edge set associated with two channels

1

Introduction

Shannon's channel coding theorem [87] characterizes the conditions under which reliable communication is possible over a noisy channel. This groundbreaking result spawned decades of research on the theory and practice of communication, and still continues to shape the development of practical communication systems.

One of the key assumptions of the channel coding theorem is that the encoder and decoder can be optimized for the specific channel model under consideration. In particular, the achievability part is usually proved used decoding methods such as maximum-likelihood decoding or joint typicality decoding, both of which directly depend on the channel transition law.

In many practical scenarios, however, one does not have accurate knowledge of the channel. In addition, even if the channel model is assumed to be known, the implementation of its corresponding optimal decoding rule may be hindered by practical issues such as computational limitations. As a result, there is substantial motivation to develop a theory of communication that explicitly accounts for channel uncertainty and decoding constraints.

This monograph surveys the topic of *mismatched decoding*, in which the decoder is fixed and possibly suboptimal, and only the codebook can be optimized. As well as addressing pertinent practical issues such as channel uncertainty and complexity limitations, this problem has strong connections with other fundamental problems in information theory, including zero-error communication.

In this introductory section, we first formally introduce the problem of reliable communication under mismatched decoding in Section 1.1. In Section 1.2, we outline several applications in which mismatched decoding is encountered, as well as formalizing the connections to other theoretical problems in information theory. In Section 1.3, we briefly overview other prominent approaches to addressing model uncertainty in information theory, including universal decoding, channels with a state, and adversarial channel models, and we discuss their connections and differences to the mismatched decoding perspective. Section 1.4 outlines the remainder of the monograph.

1.1 Problem Setup

In this subsection, we formally introduce the problem of point-to-point channel coding (see Figure 1.1) with mismatched decoding, which will be studied throughout the monograph.

We consider a communication channel defined on an input alphabet $\mathcal{X}$ and an output alphabet $\mathcal{Y}$. For a given use of the channel, if the input is given by $x \in \mathcal{X}$, then the output $y \in \mathcal{Y}$ is randomly generated according to some channel transition law $W(y|x)$. It is useful to initially think of $\mathcal{X}$ and $\mathcal{Y}$ being finite, in which case $W(y|x)$ is a conditional probability mass function. However, when we turn to continuous-alphabet channels, we will use the same notation to represent a probability density function.

We focus our attention on *memoryless channels* used without feedback, meaning that different uses of the channel are independent: if we input $\boldsymbol{x} = (x_1, \ldots, x_n)$ in n channel uses, then the output sequence $\boldsymbol{y} = (y_1, \ldots, y_n)$ is generated according to $W^n(\boldsymbol{y}|\boldsymbol{x}) \triangleq \prod_{i=1}^n W(y_i|x_i)$.

The goal is to reliably transmit a message $m \in \{1, \ldots, M\}$ over the channel in n uses. We assume that m is uniformly random over

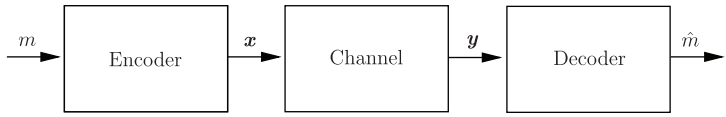


Figure 1.1: Illustration of channel coding with message m, channel input $\boldsymbol{x} \in \mathcal{X}^n$, channel output $\boldsymbol{y} \in \mathcal{Y}^n$, and estimate $\hat{m}$.

the M possibilities. The encoding scheme is represented by a *codebook* $\mathcal{C} = \{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(M)}\}$; when the message is m, the encoder transmits the corresponding codeword $\boldsymbol{x}^{(m)} \in \mathcal{X}^n$.

Naturally, the defining feature of the mismatched decoding problem is at the decoder. In the regular (i.e., matched) channel coding problem, one may consider an optimized decoding rule for producing the estimate $\hat{m}$ as a function of $\boldsymbol{y}$ and $\mathcal{C}$. In the *mismatched decoding* problem, however, the decoder is fixed as follows:

$$\hat{m} = \operatorname*{arg\,max}_{j=1,\ldots,M} q^n(\boldsymbol{x}^{(j)}, \boldsymbol{y}), \quad \text{with} \quad q^n(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} q(x_i, y_i), \tag{1.1}$$

where $q(x, y)$ is a non-negative function called the *decoding metric*. As a result, our only freedom is in designing the codebook $\mathcal{C}$. We will shortly discuss the issue of tie-breaking (i.e., the possibility of a non-unique maximum in (1.1)); see Remark 1.1 below.

Given the message estimate as in (1.1), the *error probability* associated with the codebook $\mathcal{C}$ is given by

$$p_{\mathrm{e}}(\mathcal{C}) = \mathbb{P}[\hat{m} \neq m], \tag{1.2}$$

where the probability is with respect to the randomness of the message and the channel. We seek to characterize the trade-off between the error probability p_{e}, the coding rate $R = \frac{1}{n} \log M$, and the block length n.

We briefly mention that (1.1) can be generalized to the maximization of a metric $q_n(\boldsymbol{x}^{(j)}, \boldsymbol{y})$ that need not factorize according to a product over the symbols. The consideration of (1.1) is analogous to the *memorylessness* assumption of the channel, namely, $W^n(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{n} W(y_i|x_i)$. The consideration of memoryless channels and product-wise decoding metrics permits analytical tractability, while already coming with many interesting challenges and having several applications (see Section 1.2).

Nevertheless, the study of mismatched decoding is by no means limited to this setting, e.g., see [8, 47, 57].

For any given codebook, the decoding rule minimizing p_{e} is *maximum-likelihood* (ML) decoding, corresponding to (1.1) with $q(x, y) = W(y|x)$. Accordingly, we will refer to this case as *matched decoding*. The standard Shannon capacity (henceforth called the *matched capacity*), denoted by $C(W)$, is defined to be the supremum of all rates such that one can achieve arbitrarily small error probability under matched decoding, and the following definition generalizes this notion to the mismatched case.

Definition 1.1 (Mismatch Capacity)**.** For a mismatched memoryless channel described by (W, q), we say that R is an *achievable rate* if, for all $\delta > 0$, there exists a sequence of codebooks $\mathcal{C}_n$ with $M \geq e^{n(R-\delta)}$ codewords such that $p_{\mathrm{e}}(\mathcal{C}_n) \to 0$ under the decoding rule (1.1). The *mismatch capacity* $C_{\mathrm{M}} = C_{\mathrm{M}}(W, q)$ is defined to be the supremum of all achievable rates.

We consider $M \geq e^{n(R-\delta)}$ instead of $M \geq e^{nR}$ in this definition in order to ensure that the limit of any achievable rate is also achievable. This will allow to safely use the compact terminology "achievable rate" in place of "lower bound on the mismatch capacity".

Remark 1.1 (Tie Breaking)**.** In general, it is possible for multiple messages to simultaneously achieve the maximum in (1.1); in such cases, we adopt a pessimistic approach and assume that an error always occurs when there is a tie. However, the mismatch capacity is unchanged regardless of the tie-breaking rule, since breaking ties as errors increases the error probability by at most a factor of two [31].

Remark 1.2 (Maximal vs. Average Error Probability)**.** While we consider the average error probability in (1.2), the mismatch capacity is also unchanged when we consider the maximal error probability, defined as

$$p_{\mathrm{e,max}}(\mathcal{C}) = \max_{m=1,\ldots,M} \mathbb{P}[\hat{m} \neq m | \boldsymbol{x}^{(m)} \text{ sent}]. \tag{1.3}$$

This claim follows via a standard expurgation argument [38, Sec. 5.6]: For any codebook $\mathcal{C}$ of size M satisfying $p_{\mathrm{e}}(\mathcal{C}) \leq \epsilon$ for some $\epsilon > 0$,

we can consider a sub-codebook $\mathcal{C}'$ of size $\frac{M}{2}$ whose codewords have the smallest conditional error probability. Each remaining codeword must satisfy $\mathbb{P}[\hat{m} \neq m | \boldsymbol{x}^{(m)} \text{ sent}] \leq 2\epsilon$ when the mismatched decoding rule (1.1) is applied using $\mathcal{C}$. Then, the same follows for the mismatched decoder that uses $\mathcal{C}'$, since for any $\boldsymbol{x} \in \mathcal{C}'$, there are only fewer incorrect codewords $\overline{\boldsymbol{x}} \in \mathcal{C}'$ (compared to $\overline{\boldsymbol{x}} \in \mathcal{C}$) that could lead to the error event $q^n(\overline{\boldsymbol{x}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})$. Hence, $p_{\text{e,max}}(\mathcal{C}') \leq 2\epsilon$, and since the rate of $\mathcal{C}'$ is asymptotically the same as that of $\mathcal{C}$, the desired claim follows.

In the following subsection, we overview several applications of the mismatched decoding, highlighting some cases in which the seemingly innocuous definition of the mismatch capacity C_{M} should be interpreted with care.

1.2 Applications of Mismatched Decoding

1.2.1 Channel Uncertainty

An immediate application of mismatched decoding is that in which the decoder has an incorrect channel estimate $\widehat{W}(y|x)$ that is used as if it were correct, corresponding to (1.1) with $q(x, y) = \widehat{W}(y|x)$. Such channel uncertainty can arise in many different forms, including incorrect parameter estimates [26], amplitude mismatch in real channels [65], phase offsets in complex channels [65], and incorrect models of additive noise distributions [53].

While this application of mismatched decoding is seemingly natural, the notion of *mismatch capacity* often needs to be interpreted with care. In particular, Definition 1.1 assumes that the codebook can be optimized, but performing such an optimization may implicitly require knowledge of both W and q. Hence, if W is unknown, then a potentially more meaningful quantity is the rate achieved when a codebook designed for $\widehat{W}$ is applied to W, and used with the above-mentioned decoding metric $q(x, y) = \widehat{W}(y|x)$.

Fortunately, the achievable rates presented in this monograph will account for both scenarios. In particular, we will generally provide performance bounds for a *given random-coding ensemble*. We can optimize over the given class of random-coding ensembles to get the best possible

lower bound on the mismatch capacity, or in accordance with the above discussion, we can consider the rate for a fixed ensemble (e.g., one which would be capacity-achieving if the true channel were $\widehat{W}$) as the appropriate performance measure.

1.2.2 Fading Channels

In wireless communication scenarios, the channel typically exhibits unknown fluctuations resulting from the signal arriving at the receiver via multiple paths in a dynamic environment. A widely-adopted complex-valued channel model in this context is the following:

$$Y_i = H_i X_i + Z_i, \tag{1.4}$$

where $H_i \in \mathbb{C}$ is a *fading coefficient*, and $Z_i \in \mathbb{C}$ is additive complex Gaussian noise. The use of a complex-valued model arises naturally when transmitting signals modulated by a high-frequency carrier; see for example [102].

Since the coefficients H_i are determined by a complex dynamic environment, one cannot expect to know their values precisely. As a result, one typically performs *channel estimation* to obtain estimates $\{\widehat{H}_i\}_{i=1}^n$ of the coefficients $\{H_i\}_{i=1}^n$, and then uses the estimates as if they were the true coefficients. In particular, the mismatched version of the maximum-likelihood rule, which would be optimal if the estimates were correct, is given by

$$\hat{m} = \arg\min_{j=1,\ldots,M} \sum_{i=1}^{n} |Y_i - \widehat{H}_i X_i^{(j)}|^2, \tag{1.5}$$

where $(X_1^{(j)}, \ldots, X_n^{(j)})$ is the j-th codeword. As a result, we have a *mismatched decoding rule* as in (1.1), with decoding metric $q(x, (\widehat{h}, y)) = e^{-|y-\widehat{h}x|^2}$. Note that $\widehat{h}$ is treated as part of the channel output here, since it is known at the decoder. In this application, one typically considers random coding techniques with a Gaussian input distribution, rather than (unrealistically) adopting the best codebook as per Definition 1.1.

It should be noted that the "channel" dictating the conditional distribution of $(\widehat{\boldsymbol{H}}, \boldsymbol{Y})$ given $\boldsymbol{X}$ is only memoryless under strong assumptions. However, even in the presence of memory, the mismatched

decoding perspective has led to a variety of powerful theoretical results, including both achievable rates and provable limitations of random coding imposed by the mismatch [57, 105].

1.2.3 Reduced Decoding Complexity

Even in cases that the channel W is known, its associated optimal decoding rule may be too complex to implement or even approximate. In such scenarios, it may be preferable or even essential to adopt a decoding rule that is simpler but suboptimal.

For instance, while the maximum-likelihood decoder under additive Gaussian noise is the nearest-neighbor decoder, a non-Gaussian noise distribution may lead to a much more complex maximum-likelihood rule. To circumvent this difficulty, one may simply revert to the nearest-neighbor decoder (or a similar rule) despite its suboptimality, and this becomes a problem of mismatched decoding. By characterizing the resulting performance, we can understand the extent to which simple decoding rules remain useful beyond the specific channels that they were designed for.

A more detailed example of reducing complexity at the expense of suboptimal decoding is discussed in the following subsection.

1.2.4 Bit-Interleaved Coded Modulation

The technique of bit-interleaved coded modulation (BICM) [44] is motivated by the question of how binary codes can be employed even when the communication channel induced by a modulation scheme is non-binary. BICM has been widely adopted in practical settings, including a number of standards.

If the transmitted symbol x belongs to a signal constellation set $\mathcal{X}$ with 2^ℓ elements, we may represent x using a label of ℓ bits, $\boldsymbol{b} = (b_1, \ldots, b_\ell)$. We make the dependence explicit by writing the i-th bit in the binary mapping of x as $b_i(x)$. The key element of BICM is the observation that the overall non-binary code $\mathcal{C}$ may also be decoded at bit level. More formally, in place of a potentially complicated symbol-wise decoding metric, one can decode according to a product of bit-wise

metrics:

$$q(x,y) = \prod_{i=1}^{\ell} q_i(b_i(x), y), \tag{1.6}$$

where the bit decoding metrics $q_i(b,y)$ (for $i = 1, \ldots, \ell$) are given as

$$q_i(b,y) = \sum_{x \in \mathcal{X}} P_{X|B_i}(x|b) W(y|x), \tag{1.7}$$

and where $P_{X|B_i}$ is the conditional distribution induced by the channel input distribution. The decoding metric (1.6) corresponds to treating the ℓ bits labeling a symbol as if they were independent, and the metric (1.7) for a given bit corresponds to treating all other bits as noise.

A particularly relevant case in practice is that in which the input distribution is uniform (i.e., all constellation symbols occur with equal probability), in which case the conditional probability $P_{X|B_i}(x|b_i)$ is zero if $b_i(x) \neq b$, and $2^{-(\ell-1)}$ if $b_i(x) = b$. In this case, decoding can proceed using standard binary decoding techniques, e.g., forming ℓ parallel bit decoding metrics for every channel output.

BICM was first explicitly cast as a mismatched decoding problem in [62], permitting the application of general mismatched decoding achievable rates specialized to the decoding metric given in (1.6).

1.2.5 Finite-Precision Arithmetic

Another interesting application of mismatched decoding is that of finite-precision arithmetic [15, 75]. We illustrate the general idea through the specific example of an additive white Gaussian noise (AWGN) channel $Y = X + Z$, with $Z \sim \mathrm{N}(0, \sigma^2)$.

The optimal decoding metric, corresponding to maximum-likelihood decoding, is $\log q(x,y) = -(y-x)^2$, corresponding to nearest-neighbor decoding. However, such a decoder may be ruled out when hardware constraints limit the precision of the arithmetic. In such cases, a more feasible decoding rule is given by

$$\log q(x,y) = -(\Phi_Y(y) - \Phi_X(x))^2, \tag{1.8}$$

where Φ_X and Φ_Y are scalar quantization functions. This is again a mismatched decoding problem.

We note that this setup is not equivalent to applying the scalar quantizer Φ_X prior to transmission, as we still allow the transmitted codeword to take arbitrary values; it is only the decoder that is required to work with the quantized values. In this application, it is often reasonable to assume that the codebook designer knows both W and q, which justifies the optimization over the codebook in Definition 1.1.

1.2.6 Optical Communication

Optical communication is of considerable practical importance, but the development of precise information-theoretic rates is a notoriously hard problem. While optical fiber channels with low signal powers are typically well-approximated by an additive white Gaussian noise (AWGN) channel, severe non-linearities hinder the performance at higher powers [8, 85]. In addition, the underlying channel can be far from memoryless, due to phenomena such as inter-symbol interference.

Despite the complexity of the channel, simpler achievable rates can be attained by considering suboptimal decoding rules (see [85] for a recent survey), such as the AWGN decoding metric that ignores channel memory. In this sense, the corresponding optical communication problem can be viewed through the lens of mismatched decoding [40].

The technique of BICM, outlined in Section 1.2.4, is also widespread in the field of optical communication, and a mismatched decoding perspective can again be adopted [8]. A distinction here is that the resulting achievable rates are typically difficult to compute exactly due to the presence of high-dimensional integrals. Nevertheless, one can accurately approximate the rates using techniques such as Gauss-Hermite quadrature or Monte Carlo integration, as is often done even for standard Gaussian channels when discrete input constellations are employed.

1.2.7 Zero Undetected Error Capacity

In studies of channel coding, it is often useful to let the decoder explicitly output an error when it is not confident enough to make a decision on the transmitted message. When this is done, one can distinguish between a *detected error*, where the decoder chooses to declare an error, and an

undetected error, where the decoder outputs an incorrect estimate of the message. It is then of interest to understand the trade-off between the two, particularly in cases where undetected errors are considered much more costly.

In the special case of DMCs, an interesting extreme case of this trade-off is the *zero-undetected error capacity*, which is the highest achievable rate when the probability of an undetected error must be exactly zero (i.e., the error probability equals the detected error probability) [2, 26]. For this to occur, the decoder must only output a message estimate if there is a unique codeword for which the output sequence has positive probability; if multiple codewords are feasible (i.e., consistent with the output), an error must be declared. For this setting to be meaningful, some of the channel transition probabilities $W(y|x)$ must be zero.

The zero-undetected error capacity is a special case of the mismatch capacity, corresponding to the decoding metric [26]

$$q(x, y) = \mathbb{1}\{W(y|x) > 0\}, \tag{1.9}$$

often referred to as the *erasures-only* metric. To see this, note that $q^n(\boldsymbol{x}, \boldsymbol{y}) = 1$ whenever $\boldsymbol{x}$ is feasible, and $q^n(\boldsymbol{x}, \boldsymbol{y}) = 0$ otherwise. Hence, assuming ties are broken as errors, the mismatched decoding rule leads to an error if and only if multiple codewords are feasible, as desired.

1.2.8 Zero-Error Capacity

Another interesting special case of the mismatch capacity is the *zero-error capacity* [51, 88] of a DMC, which is the highest communication rate such that the probability of *any* error is exactly zero. The zero-error capacity is known to depend on the channel $W(y|x)$ only through the $|\mathcal{X}| \times |\mathcal{X}|$ matrix $\boldsymbol{A}_W$ with (x, x')-th entry

$$A_W(x, x') = \mathbb{1}\{W(y|x)W(y|x') > 0 \text{ for some } y \in \mathcal{Y}\}, \tag{1.10}$$

where we assume without loss of generality that $\mathcal{X} = \{1, \ldots, |\mathcal{X}|\}$. The matrix $\boldsymbol{A}_W$ is the *adjacency matrix of the channel graph*, formed by connecting two inputs if and only if they share a common output according to W.

With this definition, the zero-error capacity of W is equal to the mismatch capacity of the following channel-metric pair, whose input

and output alphabets are both $\mathcal{X}$ [26]:

$$W_0(x'|x) = \mathbb{1}\{x = x'\}, \quad q_0(x, x') = A_W(x, x'). \tag{1.11}$$

That is, the channel is noiseless, and the metric is the indicator function of (x, x') sharing a common output in W, sometimes referred to as being *adjacent* or *confusable*.

To see that $C_{\mathrm{M}}(W_0, q_0)$ equals the zero-error capacity of W, we first recall from Remark 1.2 that the mismatch capacity is unchanged when the requirement of p_{e} being arbitrarily small is replaced by the requirement that $p_{\mathrm{e,max}}$ (see (1.3)) is arbitrarily small. Hence, it suffices to show that any given sequence of codebooks $\{\mathcal{C}_n\}_{n\geq 1}$ attains zero error probability on W (i.e., any transmitted codeword can be uniquely decoded with probability one) if and only if $p_{\mathrm{e,max}}(\mathcal{C}_n) \to 0$ for the mismatched DMC (W_0, q_0).

The "only if" statement follows from the fact that any zero-error codebook $\mathcal{C}_n$ for W yields $q_0^n(\boldsymbol{x}, \boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \mathcal{C}$, and $q_0^n(\boldsymbol{x}, \overline{\boldsymbol{x}}) = 0$ for all $\boldsymbol{x}, \overline{\boldsymbol{x}} \in \mathcal{C}_n$ with $\boldsymbol{x} \neq \overline{\boldsymbol{x}}$, due to the choice of q_0 in (1.11). Thus, $\mathcal{C}_n$ satisfies $p_{\mathrm{e,max}}(\mathcal{C}) = 0$ under the mismatched DMC (W_0, q_0), in which the channel output deterministically equals the input. To establish the other direction, we note that since both the channel W_0 and the decoder (with ties broken as errors) are deterministic, the quantities $\mathbb{P}[\hat{m} \neq m | \boldsymbol{x}^{(m)} \text{ sent}]$ defining $p_{\mathrm{e,max}}$ in (1.3) can only be zero or one. Thus, the only way to attain $p_{\mathrm{e,max}}(\mathcal{C}_n) \to 0$ is to have $p_{\mathrm{e,max}}(\mathcal{C}_n) = 0$ for all sufficiently large n. By the choice of q_0, one can only have $p_{\mathrm{e,max}}(\mathcal{C}_n) = 0$ when all codewords have no common outputs, which implies that $\mathcal{C}_n$ is a zero-error codebook for W.

A related example can be found in [26], showing that a graph-theoretic quantity called the *Sperner capacity* is also a special case of the mismatch capacity.

1.3 Other Channel Uncertainty Models

The mismatched decoding problem provides one of several perspectives on the role of channel uncertainty in information theory. An excellent survey comparing the various perspectives can be found in [56], so we provide only a brief discussion here.

It is well-known that *universal decoding rules* exist that are able to achieve the capacity of any DMC when the codebook is optimized. A prominent example is the maximum empirical mutual information decoder [24, Ex. 6.20]. Adopting such a decoder under channel uncertainty would typically lead to strictly better achievable rates compared to a mismatched decoding metric. However, such decoding rules are *multi-letter* in nature, and computationally efficient techniques for implementing or approximating them remain elusive. In contrast, the decoding rule in (1.1) is *single-letter*, in the sense that it is written as a product over individual symbols, and many practical decoding methods can be viewed as approximating this rule (despite not necessarily implementing it exactly).

Channels with a state, in which the transition law takes the form $W(y|x,s)$ for some unknown state s, also provide a powerful means to address channel uncertainty [30, Ch. 7]. In particular, the capacity varies depending on whether the state is known at the encoder and/or decoder, and whether it is known causally or non-causally. A prominent example is fading channels, where s represents a fading coefficient. In fact, as evidenced in the previous subsection, the study of channels with a state is by no means disjoint from the mismatched decoding problem. Rather, channels with a state are simply another class of channels where the decoding rule may be either matched or mismatched; e.g., see [57] for mismatched fading channels, and [33] for mismatched DMCs with a state.

Another widely-adopted approach is to consider *adversarial* channel models, taking the form $W_n(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s})$ for some sequence $\boldsymbol{s} = (s_1, \ldots, s_n)$ that can be viewed as being chosen by an adversary (possibly subject to some constraints) in order to hinder the communication. Hence, one requires small error probability simultaneously for all possible sequences $\boldsymbol{s}$. Under the constraint $s_1 = \cdots = s_n$ this is referred to as a *compound channel* [16], whereas the case of arbitrary length-n sequences is referred to as the *arbitrarily varying channel* (AVC) [17]. Adversarial channel models are of particular interest for communication in the presence of jamming.

To our knowledge, the study of AVCs has had minimal interaction with the mismatched decoding perspective in the existing literature.

This is because bounds on the capacity of the AVC typically adopt rather complicated multi-letter decoding rules, e.g., not simply maximizing a metric on $(\boldsymbol{x}, \boldsymbol{y})$, but instead considering various combinations of triplets $(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{y})$ [25]. In addition, the study of AVCs often leads to very different insights and challenges compared to the mismatched decoding problem, such as the fact that the capacity may differ for the average error probability vs. maximum error probability criteria, and the consideration of randomized encoding. See [56] for a detailed overview.

Despite these differences, we briefly mention two examples of connections between adversarial channel models and mismatched decoding: (i) For certain compound channels, the capacity can be achieved by a single-letter mismatched decoding rule [99]; we present this example in Section 2.4.4. (ii) It is shown in [26] that the erasures-only decoding metric (see (1.9)) can achieve the capacity of certain AVCs under deterministic coding and the average-error criterion.

1.4 Outline of the Monograph

Through the monograph, we focus primarily on achievable rates (i.e., lower bounds on the mismatch capacity) via random coding. This has been the primary focus of the existing literature, with upper bounds on the mismatch capacity generally remaining elusive. The study of achievable rates in itself comes with many interesting challenges and differences compared to the matched setting. To name just one example, we will see that both constant-composition codes and multi-user coding techniques can lead to better achievable rates than standard i.i.d. random coding, in stark contrast with the matched setting in which i.i.d. random codes are capacity-achieving.

In Section 2, we present achievable rates for discrete memoryless channels, introducing the *generalized mutual information* (GMI) for i.i.d. coding, and the *LM rate* for constant-composition random coding. We present several key properties of both rates, including conditions for positivity, conditions under which the matched capacity is achieved, and ensemble tightness results. In addition, we present a variety of representative examples from the literature.

In Section 3, we present a generalization of the GMI and LM rate to continuous-alphabet channels, using the idea of random coding with auxiliary costs. We again provide several representative examples focusing on additive channels and variants of mismatched nearest-neighbor decoding. In particular, these examples include non-Gaussian noise channels and fading channels, both of which have been among the most widespread applications of mismatched decoding in the literature.

In Section 4, we momentarily depart from the channel coding problem and survey the role of mismatch in rate-distortion theory. In this case, the problem becomes one of *mismatched encoding*, and we again present achievable rates that are proved via random coding, as well as a multi-letter converse result. In addition, we overview a different type of mismatch in which the optimal encoder is adopted but a suboptimal random coding distribution is used, highlighting the important special case of Gaussian compression techniques for non-Gaussian sources.

In Section 5, we study the mismatched multiple-access channel (MAC), presenting an ensemble-tight achievable rate region for constant-composition random coding, as well as an extension to continuous alphabets. In Section 6, we survey further results demonstrating that multi-user coding techniques, including both MAC coding methods and superposition coding, can lead to improved rates for single-user channels. In particular, we present a refined version of superposition coding that provides the best known achievable rate at the time of writing.

In Section 7, we move beyond achievable rates and present an overview of results on error exponents, which provide refined characterizations of the asymptotic behavior of the error probability at fixed rates. In particular, we discuss the role of cost-constrained random coding with multiple auxiliary costs for attaining the best known achievable error exponents, highlighting the differences compared to the study of achievable rates alone.

In Section 8, we overview recent developments and challenges in obtaining upper bounds on the mismatch capacity. In accordance with the literature on this topic, we focus primarily on multi-letter bounds that are non-computable but can still provide interesting insight on the problem. In addition, we present a recent single-letter upper bound.

In Section 9, we briefly outline several topics in mismatched decoding that were omitted from the main sections, including refined asymptotics, modified channel coding settings, further results on continuous and fading channels, and mismatch in practical coding methods. For each of these, we give several pointers to the relevant literature. Conclusions are drawn in Section 10.

Unless it is explicitly stated otherwise, it should be understood that the results surveyed in this monograph come from the existing literature. We highlight the main references at the beginning of each section.

Notes on the Presentation and Assumed Background. We make an effort to provide self-contained proofs of most results, but for certain parts that are highly technical and/or less central to the survey, we instead refer the reader to the relevant literature.

We assume that the reader is familiar with the classical theorems of information theory (e.g., source coding, channel coding, and rate-distortion theory), as well as the associated information measures (e.g., entropy, mutual information, and KL divergence) and their properties (e.g., chain rule, data processing inequality). As for slightly less standard prior knowledge, the reader is encouraged to learn the basics of the method of types [24, Ch. 2] and Gallager's proof of the channel coding theorem [38, Ch. 5] if either of these is unfamiliar.

Notation. We will introduce our notation throughout the monograph, but for reference, we also summarize some recurring notation here. Definitions of the most commonly-used mathematical symbols and abbreviations are listed prior to Section 1.

The set of all probability distributions on an alphabet $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{X})$, and the set of all conditional distributions on $\mathcal{Y}$ given $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{Y}|\mathcal{X})$. More precisely, in the case that $\mathcal{X}$ is finite, we write $P_X \in \mathcal{P}(\mathcal{X})$ when P_X is a probability mass function (PMF) on $\mathcal{X}$, and similarly for conditional PMFs. When we consider continuous alphabets, we use the same notation for probability density functions (PDFs). The set of all empirical distributions on a vector in $\mathcal{X}^n$ (i.e., types [24, Sec. 2], [37]) is denoted by $\mathcal{P}_n(\mathcal{X})$, and the set of all conditional empirical

distributions (i.e., conditional types) on $\mathcal{Y}^n$ given $\mathcal{X}^n$ is denoted by $\mathcal{P}_n(\mathcal{Y}|\mathcal{X})$. Given $P_X \in \mathcal{P}_n(\mathcal{X})$, the type class $\mathcal{T}^n(P_X)$ is defined to be the set of all sequences in $\mathcal{X}^n$ with type P_X.

We use bold symbols for vectors (e.g., $\boldsymbol{x}$), and we denote the corresponding i-th entry using a subscript (e.g., x_i). In the case of random vectors, we use capital letters (e.g., $\boldsymbol{X}$ and X_i). The probability of an event is denoted by $\mathbb{P}[\cdot]$, and the symbol $\sim$ means "distributed as". The marginals of a joint distribution $P_{XY}(x, y)$ are denoted by $P_X(x)$ and $P_Y(y)$. Similarly, $P_{Y|X}(y|x)$ denotes the conditional distribution induced by $P_{XY}(x, y)$. Expectation with respect to a joint distribution $P_{XY}(x, y)$ is denoted by $\mathbb{E}_P[\cdot]$. When the probability distribution is understood from the context, we simply write $\mathbb{E}[\cdot]$. Given a distribution $Q_X(x)$ and a conditional distribution $W(y|x)$, we write $Q_X \times W$ to denote the joint distribution $Q_X(x)W(y|x)$, and similarly when there are more than two distributions.

We use standard notation for the entropy $H(X)$, conditional entropy $H(X|Y)$, mutual information $I(X;Y)$, conditional mutual information $I(X;Y|Z)$, KL divergence $D(P\|Q)$ and conditional KL divergence $D(P_{Y|X}\|Q_{Y|X}|P_X)$ [22]. We often explicitly state the underlying distribution of the random variables with a subscript (e.g., $I_P(X;Y)$, $I_{\widetilde{P}}(X;Y)$). The set of real numbers is denoted by $\mathbb{R}$, and the set of complex numbers is denoted by $\mathbb{C}$. We define $[\alpha]^+ = \max\{0, \alpha\}$, and we denote the indicator function of an event by $\mathbb{1}\{\cdot\}$. The floor function is denoted by $\lfloor \cdot \rfloor$, and the ceiling function by $\lceil \cdot \rceil$. All logarithms have base e, and all rates are in units of nats except in the examples, where bits are used.

2

Discrete Memoryless Channels

2.1 Introduction

Throughout the monograph, it will be useful to distinguish between channels with finite alphabets and continuous alphabets. In this section, we focus on the finite-alphabet setting, which will allow us to provide a more elementary introduction using widely-known techniques from the literature. In Section 3, we will consider the continuous-alphabet setting.

We consider the problem setup as described in Section 1.1, in the special case that $\mathcal{X}$ and $\mathcal{Y}$ are finite, and hence $W(y|x)$ represents a conditional probability mass function. Recall that a non-negative decoding metric $q(x, y)$ dictates the decision rule in (1.1). Collectively, we refer to the pair (W, q) as a *mismatched DMC*.

The achievable rates presented this section are based on *random coding*, a ubiquitous tool in information theory in which one proves the existence of good codebooks by bounding the average error probability of a randomly-generated codebook [87]. A given random construction, characterized by the joint distribution of the codewords, is called a *random-coding ensemble*. In this section, we assume that the codewords

are drawn independently from a common distribution $P_{\boldsymbol{X}}$:

$$\{\boldsymbol{X}^{(m)}\}_{m=1}^{M} \sim \prod_{i=1}^{n} P_{\boldsymbol{X}}(\boldsymbol{x}^{(m)}). \tag{2.1}$$

In Section 6, we will turn to alternative ensembles in which some form of dependence is introduced among the codewords.

This section is predominantly based on the works of Hui [46], Csiszár and Körner [23], Csiszár and Narayan [26], and Merhav *et al.* [65].

2.2 Properties of the Mismatch Capacity

In this subsection, we provide some straightforward properties of the error probability and mismatch capacity that will be useful for motivating the random coding ensembles and interpreting their achievable rates.

We begin with the following proposition regarding the equivalence of certain decoding metrics, which in fact applies not only to mismatched DMCs, but also general mismatched memoryless channels.

Proposition 2.1 (Equivalence of Metrics)**.** For any memoryless channel $W(y|x)$ and codebook $\mathcal{C} = (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(M)})$, if two decoding metrics $q, \widetilde{q}$ are related according to

$$\log \widetilde{q}(x, y) = b(y) + s \log q(x, y) \tag{2.2}$$

for some $b(y)$ and $s > 0$, then we have for all $\boldsymbol{y} \in \mathcal{Y}^n$ that

$$\operatorname*{arg\,max}_{j=1,\ldots,M} \widetilde{q}^n(\boldsymbol{x}^{(j)}, \boldsymbol{y}) = \operatorname*{arg\,max}_{j=1,\ldots,M} q^n(\boldsymbol{x}^{(j)}, \boldsymbol{y}), \tag{2.3}$$

i.e., the two decoding metrics are equivalent.

Recalling that $q^n(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} q(x_i, y_i)$, this claim follows by simply writing $\log \widetilde{q}^n(\boldsymbol{x}, \boldsymbol{y}) = b^n(\boldsymbol{y}) + s \log q^n(\boldsymbol{x}, \boldsymbol{y})$, where $b^n(\boldsymbol{y}) = \sum_{i=1}^{n} b(y_i)$. Since $s > 0$, this implies that the two maximization problems in (2.3) are equivalent.

We can strengthen Proposition 2.1 for a useful class of codebooks called *constant-composition codebooks*. The *composition* of a sequence $\boldsymbol{x}$

(also referred to as its *type* or *empirical distribution*) is defined as

$$\widehat{P}_{\boldsymbol{x}}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{x_i = x\}, \tag{2.4}$$

and a constant-composition codebook is one in which all codewords have the same composition. Thus, every sequence has the same number of occurrences of any given symbol, and the codewords are permutations of one another.

Proposition 2.2 (Equivalence of Metrics for Constant-Composition Codebooks)**.** For any memoryless channel $W(y|x)$ and constant-composition codebook $\mathcal{C} = (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(M)})$, if two decoding metrics $q, \widetilde{q}$ are related according to

$$\log \widetilde{q}(x, y) = a(x) + b(y) + s \log q(x, y) \tag{2.5}$$

for some $a(x)$, $b(y)$, and $s > 0$, then we have for all $\boldsymbol{y} \in \mathcal{Y}^n$ that

$$\operatorname*{arg\,max}_{j=1,\ldots,M} \widetilde{q}^n(\boldsymbol{x}^{(j)}, \boldsymbol{y}) = \operatorname*{arg\,max}_{j=1,\ldots,M} q^n(\boldsymbol{x}^{(j)}, \boldsymbol{y}), \tag{2.6}$$

i.e., the two decoding metrics are equivalent for constant-composition codebooks.

The proof is similar to that of Proposition 2.1: We have $\log \widetilde{q}^n(\boldsymbol{x}, \boldsymbol{y}) = a^n(\boldsymbol{x}) + b^n(\boldsymbol{y}) + s \log q^n(\boldsymbol{x}, \boldsymbol{y})$, and for any constant-composition codebook, the value $a^n(\boldsymbol{x}) = \sum_{i=1}^n a(x_i)$ is the same for every codeword, as the codewords are all permutations of one another.

The constant-composition assumption may seem restrictive, but in fact, *any* codebook can be reduced to a constant-composition codebook with a negligible loss in the rate. To see this, note that $\widehat{P}_{\boldsymbol{x}}$ in (2.4) takes the form $(\frac{n_1}{n}, \ldots, \frac{n_{|\mathcal{X}|}}{n})$ with $n_x \in \{0, 1, \ldots, n\}$ for all $x \in \mathcal{X}$, and hence, the total number of different compositions is at most $(n+1)^{|\mathcal{X}|}$. As a result, given a codebook with $M = e^{nR}$ codewords, among all possible sub-codebooks with codewords having the same composition, there must exist one whose size is at least

$$M' \geq \frac{M}{(n+1)^{|\mathcal{X}|}}. \tag{2.7}$$

If $\mathcal{X}$ is finite, as is the case for a DMC, then it holds for any fixed rate R that $\frac{1}{n}\log M' \to R$ as $n \to \infty$. In addition, the maximal error probability in the reduced codebook is no higher than that of the original codebook, since for any given transmitted codeword $\boldsymbol{x}$, there are only fewer incorrect codewords $\overline{\boldsymbol{x}}$ remaining that could lead to the error event $q^n(\overline{\boldsymbol{x}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})$. Since the mismatch capacity is always the same under the average and maximal error criteria (see Remark 1.2), we have the following corollary.

Corollary 2.1 (Metrics with Equal Mismatch Capacities)**.** For any memoryless channel W and decoding metrics q and $\widetilde{q}$, we have the following:

(i) The mismatch capacities of (W, q) and $(W, \widetilde{q})$ are identical provided that q and $\widetilde{q}$ are related according to (2.2) for some $b(y)$ and $s > 0$.

(ii) If the input alphabet $\mathcal{X}$ is finite, then the mismatch capacities of (W, q) and $(W, \widetilde{q})$ are identical provided that $q, \widetilde{q}$ are related according to (2.5) for some $a(x)$, $b(y)$, and $s > 0$.

2.3 Achievable Rates: GMI and LM Rate

In this subsection, we introduce the two most well-known achievable rates in the mismatched decoding literature: The *generalized mutual information* (GMI) and the *LM rate*.[1] These rates are derived using two common random coding techniques known as *i.i.d. random coding* and *constant-composition random coding*, which we introduce below. While the achievable rate for the latter of these was proved first in the literature [23, 46], we find it instructive to start with the former.

Both the GMI and LM rate will be written in two equivalent forms which, at first glance, might appear to be completely unrelated.

[1] We adopt these names because they have become widespread in the literature, but in our opinion, neither of them is ideal. While the GMI does "generalize" mutual information in the sense of recovering it in the special case $q(x, y) = W(y|x)$, the same could be said of any achievable rate that is tight when specialized to the matched case. The name *LM rate* appears to have arisen following [65], with the acronym implicitly standing for "lower [bound on the] mismatch [capacity]". This would again be equally suitable for *any* achievable rate.

The equivalence of the two (stated in Lemma 2.10 below) is proved using the method of *Lagrange duality* from the theory of convex optimization [18, Ch. 5]. Accordingly, we will refer to the two as *primal expressions* and *dual expressions*, and we will adopt similar terminology throughout the monograph whenever a similar equivalence holds.

We proceed by introducing the random coding ensembles (i.e., choices of the codeword distribution $P_{\boldsymbol{X}}$ in (2.1)) and their achievable rates, and then turn to a common discussion of the two in which we interpret both the primal and dual forms.

2.3.1 i.i.d. Random Coding and the GMI

By far the most widely-studied random-coding ensemble in the information theory literature is *i.i.d. random coding*, defined as follows. Here and subsequently, $\mathcal{P}(\mathcal{X})$ denotes the set of all probability distributions on the alphabet $\mathcal{X}$.

Definition 2.1 (i.i.d. Random Coding)**.** Under the *i.i.d. random coding ensemble* with *input distribution* $Q_X \in \mathcal{P}(\mathcal{X})$,[2] the codewords are independently drawn from the codeword distribution

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{n} Q_X(x_i). \tag{2.8}$$

That is, every symbol of every codeword is independently drawn from Q_X.

We will use i.i.d. random coding to prove the achievability of the *generalized mutual information* (GMI), which can equivalently be written in the primal form

$$I_{\mathrm{GMI}}(Q_X) = \min_{\substack{\widetilde{P}_{XY} \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_Y = P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} D(\widetilde{P}_{XY} \| Q_X \times P_Y) \tag{2.9}$$

with $P_{XY} = Q_X \times W$, or in the dual form,[3]

$$I_{\mathrm{GMI}}(Q_X) = \sup_{s \geq 0} \sum_{x,y} Q_X(x) W(y|x) \log \frac{q(x,y)^s}{\sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x},y)^s}. \tag{2.10}$$

[2] Not to be confused with the decoding metric $q(x,y)$.

[3] A supremum is used for $s \geq 0$ instead of a maximum since it may not be attained by a finite value of s; see (2.53) for an example.

These expressions are discussed in Sections 2.3.3–2.3.4 below. The following theorem provides a formal statement of achievability, which will be proved in Section 2.6.

Theorem 2.2 (Achievability of the GMI)**.** For any mismatched DMC (W, q) and a given input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, the rate $I_{\text{GMI}}(Q_X)$ is achievable via i.i.d. random coding. Consequently,

$$C_{\text{M}} \geq C_{\text{GMI}} \triangleq \max_{Q_X \in \mathcal{P}(\mathcal{X})} I_{\text{GMI}}(Q_X). \tag{2.11}$$

Of course, $I_{\text{GMI}}(Q_X)$ depends not only on Q_X, but also on (W, q). For the most part, we leave this dependence implicit, as we will be considering a fixed channel and decoding metric. However, in some cases, we will make the dependence explicit by writing $I_{\text{GMI}}(Q_X, W)$ or $I_{\text{GMI}}(Q_X, W, q)$.

2.3.2 Constant-Composition Random Coding and the LM Rate

Another well-known random coding method for channel coding is *constant-composition random coding*, which introduces some dependence between the symbols of a given codeword. In contrast to the matched setting [24, 37], we will see that this ensemble can attain better higher rates than the i.i.d. ensemble in the presence of mismatch (see Section 2.4).

Before defining the ensemble, we introduce some terminology. Recalling the notion of a *type* in (2.4), we let $\mathcal{P}_n(\mathcal{X}) \subset \mathcal{P}(\mathcal{X})$ be the subset of probability distributions that correspond to the type of some length-n sequence, i.e., each probability equals an integer multiple of $\frac{1}{n}$. For a given type $P_X \in \mathcal{P}_n(\mathcal{X})$, we define the *type class*

$$\mathcal{T}^n(P_X) = \{\boldsymbol{x} \in \mathcal{X}^n \colon \widehat{P}_{\boldsymbol{x}} = P_X\}, \tag{2.12}$$

where $\widehat{P}_{\boldsymbol{x}}$ is defined in (2.4). In words, $\mathcal{T}^n(P_X)$ contains all length-n sequences having empirical distribution P_X.

As with i.i.d. random coding, the constant-composition codeword distribution is specified by an input distribution $Q_X \in \mathcal{P}(\mathcal{X})$. Given Q_X, we define $Q_{X,n} \in \mathcal{P}_n(\mathcal{X})$ to be an arbitrary type having the same

support as Q_X, and satisfying the following:

$$\|Q_X - Q_{X,n}\|_\infty \triangleq \max_{x \in \mathcal{X}} |Q_X(x) - Q_{X,n}(x)| \le \frac{1}{n}. \tag{2.13}$$

This essentially amounts to rounding each value of $Q_X(x)$ up or down so that it is an integer multiple of $\frac{1}{n}$.

Definition 2.2 (Constant-Composition Random Coding)**.** Under the *constant-composition random coding ensemble* with input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, the codewords are independently drawn from the codeword distribution

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{|\mathcal{T}^n(Q_{X,n})|} \mathbb{1}\{\boldsymbol{x} \in \mathcal{T}^n(Q_{X,n})\}, \tag{2.14}$$

where $Q_{X,n} \in \mathcal{P}_n(\mathcal{X})$ satisfies (2.13). That is, each codeword is equiprobable on the set of sequences with empirical distribution $Q_{X,n}$.

We will use constant-composition random coding to prove the achievability of the *LM rate*, which can equivalently be written in the primal form

$$I_{\mathrm{LM}}(Q_X) = \min_{\substack{\widetilde{P}_{XY} \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_X = Q_X, \widetilde{P}_Y = P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \ge \mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y) \tag{2.15}$$

with $P_{XY} = Q_X \times W$, or in the dual form,

$$I_{\mathrm{LM}}(Q_X) = \sup_{s \ge 0, a(\cdot)} \sum_{x,y} Q_X(x) W(y|x) \log \frac{q(x,y)^s e^{a(x)}}{\sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x},y)^s e^{a(\overline{x})}}. \tag{2.16}$$

These expressions are discussed in Sections 2.3.3–2.3.4, and the following formal statement of achievability is proved in Section 2.6.

Theorem 2.3 (Achievability of the LM Rate)**.** For any mismatched DMC (W, q) and a given input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, the rate $I_{\mathrm{LM}}(Q_X)$ is achievable via constant-composition random coding. Consequently,

$$C_{\mathrm{M}} \ge C_{\mathrm{LM}} \triangleq \max_{Q_X \in \mathcal{P}(\mathcal{X})} I_{\mathrm{LM}}(Q_X). \tag{2.17}$$

As with the GMI, we will sometimes write $I_{\mathrm{LM}}(Q_X, W)$ or $I_{\mathrm{LM}}(Q_X, W, q)$ to make the dependence on the channel and/or decoding metric explicit.

2.3.3 Interpretations of the Rates

We proceed by giving some interpretations of the achievable rates, and some intuition as to how they arise from the random-coding analysis. At this point, it is worth noting that the LM rate is always at least as high as the GMI:

$$I_{\mathrm{GMI}}(Q_X) \leq I_{\mathrm{LM}}(Q_X). \tag{2.18}$$

This is easily seen from the fact that the minimization in the primal expression (2.15) is more constrained than that in (2.9), or from the fact that we can lower bound the dual expression (2.16) via the specific choice $a(x) = 0$ in order to obtain (2.10).

Although the GMI is weaker than the LM rate, it is still interesting to understand it separately, and to contrast it with the LM rate in order to better understand the differences between the corresponding random-coding ensembles. Moreover, in several applications of mismatched decoding, the GMI admits a simple analytical expression but the LM rate does not.

Primal Expressions. The idea behind the proof of Shannon's (matched) channel coding theorem is that if $R < I(X;Y)$, then with high probability under random coding, no incorrect codeword $\overline{\boldsymbol{x}}$ will be such that the pair $(\overline{\boldsymbol{x}}, \boldsymbol{y})$ "looks like" it was generated according to $Q_X \times W$ (e.g., in terms of joint typicality, or having a sufficiently high likelihood). In the mismatched setting, such codewords are not the only problematic ones; errors also occur when $(\overline{\boldsymbol{x}}, \boldsymbol{y})$ "looks like" it was drawn from *any* $\widetilde{P}_{XY}$ with a higher decoding metric than $Q_X \times W$. In both (2.9) and (2.15), the achievable rate is a minimum over all such $\widetilde{P}_{XY}$. The Y-marginal constraint arises since $\widehat{P}_{\boldsymbol{y}} \to P_Y$ by the law of large numbers, and the X-marginal constraint in the LM rate arises from the constant-composition construction.

As noted in [65], the primal expressions also closely resemble *rate-distortion functions* in lossy source coding; for instance, in the absence of the constraint $\widetilde{P}_Y = P_Y$ in (2.15), we would precisely recover the rate distortion function of the "source" X with "reconstruction" Y, distortion measure $d(x,y) = -\log q(x,y)$, and distortion threshold $-\mathbb{E}_P[\log q(X,Y)]$. See Section 4.7 for further discussion.

Dual Expressions. Gallager's proof of the channel coding theorem [36] is based on maximum-likelihood decoding, and is particularly amenable to a generalization to maximum-metric decoding. If this is done naively, one attains a rate of the form (2.10) with $s = 1$, which is rather limited (e.g., such a rate may be negative) [34]. However, the achievability of the GMI immediately follows from the equivalence of the decoding metrics $q(x, y)$ and $q(x, y)^s$ for $s > 0$, stated in Proposition 2.1 (the case $s = 0$ can also trivially be included). Similarly, the LM rate can be deduced using the equivalence of $q(x, y)$ and $q(x, y)^s e^{a(x)}$ for constant-composition codes, stated in Proposition 2.2. Essentially, the improvement of constant-composition random coding over i.i.d. random coding comes from the additional codebook structure that permits the introduction of $a(\cdot)$ into the bound.

To provide another interpretation, we note that the LM rate can be written as follows (see Lemma 2.8 in Section 2.5):

$$I_{\mathrm{LM}}(Q_X, W) = I(Q_X, W) - \inf_{V_{Y|X} \in \mathcal{V}_q} D(W \| V_{Y|X} | Q_X), \tag{2.19}$$

where $I(Q_X, W) = I(X; Y)$ with $(X, Y) \sim Q_X \times W$, and

$$\mathcal{V}_q = \Big\{ V_{Y|X} \in \mathcal{P}(\mathcal{Y}|\mathcal{X}) \colon V_{Y|X}(y|x) = q(x, y)^s e^{a(x)} e^{b(y)} \\ \text{for some } s \geq 0, a(x), b(y) \Big\}, \tag{2.20}$$

where we recall that $\mathcal{P}(\mathcal{Y}|\mathcal{X})$ is the set of all conditional distributions on $\mathcal{Y}$ given $\mathcal{X}$.

To understand the form given in (2.19), consider the case that $q(x, y) = \widehat{W}(y|x)$ is a conditional distribution. In analogy with source coding methods (e.g., Huffman coding) in which one pays a KL divergence penalty for assuming the wrong distribution [22, Ch. 5], one might expect to pay a penalty of $D(W \| \widehat{W} | Q_X)$ here compared to optimal decoding. The above formulation reveals that this is an *upper bound* on the loss, and that it can always be reduced to the infimum of $D(W \| V_{Y|X} | Q_X)$ over all conditional distributions $V_{Y|X}$ yielding an equivalent decoding metric in the sense of Proposition 2.2.

2.3.4 Advantages of the Primal and Dual Forms

At this point, it may be unclear why it is beneficial to consider two equivalent formulations of each achievable rate. The advantages will become increasingly clear throughout the section and the monograph, but we briefly outline some important examples here:

- As we already saw above, the two forms can admit interesting alternative interpretations.
- The most important properties of the rates (see Section 2.5) are often easier to prove using one particular form, though some can be proved using either of the two.
- As a notable example, when it comes to establishing *ensemble tightness* (roughly, showing that the rates cannot be improved under the random coding ensembles considered), the most elementary proofs for DMCs are based on the primal expression.
- The dual forms have the notable advantage that substituting *any* choices of $s \geq 0$ and/or $a(\cdot)$ yields a valid achievable rate. In contrast, the primal expressions are only guaranteed to provide a valid achievable rate under the minimizing joint distribution $\widetilde{P}_{XY}$.
- Another advantage of the dual forms is that they come with direct derivations that readily extend to continuous-alphabet channels (see Section 3).
- When it comes to computing the achievable rates numerically, it can be useful to compute both forms and verify that they coincide, as a check for robustness against numerical precision issues.

For reasonable-sized alphabets $\mathcal{X}$ and $\mathcal{Y}$, neither the primal nor dual forms pose significant computational challenges *for a given input distribution* Q_X: The primal expressions are convex minimization problems, and the dual expressions are concave maximization problems, and hence both can be evaluated using off-the-shelf solvers. In contrast, Lemma 2.12 below suggests that the optimization over Q_X can be difficult due to non-concavity.

2.4 Examples

In this subsection, we provide analytical and numerical evaluations of the GMI and LM rate for a variety of specific DMCs.

2.4.1 Binary Channels

Consider the case that both the channel inputs and outputs are binary, i.e., $|\mathcal{X}| = |\mathcal{Y}| = 2$. We assume without loss of generality that $\mathcal{X} = \mathcal{Y} = \{0, 1\}$. We will first show that the LM rate is either equal to the matched capacity or zero [26], and then discuss how the GMI does not exhibit such a dichotomy.

Evaluating the LM Rate. Recall that the LM rate is based on constant-composition codes, and that for any such code, two decoding metrics $q, \widetilde{q}$ are equivalent if

$$\log \widetilde{q}(x, y) = a(x) + b(y) + s \log q(x, y) \tag{2.21}$$

for some constant $s > 0$ and functions $a(x), b(y)$ (see Proposition 2.2). To study such equivalences, we introduce the notation

$$\boldsymbol{\ell}_q = \begin{bmatrix} \log q(0,0) & \log q(0,1) \\ \log q(1,0) & \log q(0,1) \end{bmatrix} \triangleq \begin{bmatrix} \ell_{00} & \ell_{01} \\ \ell_{10} & \ell_{11} \end{bmatrix}, \tag{2.22}$$

and we write $\boldsymbol{\ell}_q \equiv \boldsymbol{\ell}_{\widetilde{q}}$ if (2.21) holds.

Notice that choosing $a(x)$ (respectively, $b(y)$) in (2.21) amounts to adding or subtracting a constant from each row (respectively, column) of $\boldsymbol{\ell}_q$. By choosing these constants to equate certain entries to zero, we obtain

$$\boldsymbol{\ell}_q \equiv \begin{bmatrix} 0 & 0 \\ \ell_{10} - \ell_{00} & \ell_{11} - \ell_{01} \end{bmatrix} \equiv \begin{bmatrix} 0 & 0 \\ 0 & \ell_{11} - \ell_{01} - \ell_{10} + \ell_{00} \end{bmatrix}. \tag{2.23}$$

By a symmetric argument, we can obtain the same final equivalence with the top-left and bottom-right entries swapped, and in addition, we can sum the two matrices together to obtain

$$\boldsymbol{\ell}_q \equiv \begin{bmatrix} \ell_{00} + \ell_{11} - \ell_{01} - \ell_{10} & 0 \\ 0 & \ell_{00} + \ell_{11} - \ell_{01} - \ell_{10} \end{bmatrix}. \tag{2.24}$$

Since multiplying by $s > 0$ in (2.21) amounts to scaling the entire matrix, we deduce that as long as $\ell_{11} - \ell_{01} - \ell_{10} + \ell_{00} \neq 0$, one of the following must hold:

$$\boldsymbol{\ell}_q \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{or} \quad \boldsymbol{\ell}_q \equiv \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{2.25}$$

with the non-zero entries equaling $\operatorname{sign}(\ell_{00} + \ell_{11} - \ell_{01} - \ell_{10})$.

Now, letting δ_1 and δ_2 denote the transition probabilities $W(1|0)$ and $W(0|1)$ respectively, we observe that upon substituting $\ell^*_{xy} = \log W(y|x)$, it holds that $\ell^*_{00} + \ell^*_{11} - \ell^*_{01} - \ell^*_{10} > 0$ if and only if $\frac{(1-\delta_1)(1-\delta_2)}{\delta_1\delta_2} > 1$, which in turn is equivalent to $\delta_1 + \delta_2 < 1$. Hence, if $\delta_1 + \delta_2 < 1$ then the ML rule minimizes the Hamming distance, and otherwise, it maximizes the Hamming distance.

If the sign of $\ell_{00} + \ell_{11} - \ell_{01} - \ell_{10} = \log \frac{q(0,0)q(1,1)}{q(0,1)q(1,0)}$ for the decoding metric matches that of the ML rule, then clearly we achieve the matched capacity, as the corresponding decoding rules are equivalent. If not, then the decoder is minimizing a quantity that is meant to be maximized (or vice versa), and hence we attain a rate of zero.[4]

Evaluating the GMI. The GMI is based on i.i.d. random coding, which can produce general codebooks for which two decoding rules $q, \widetilde{q}$ are only equivalent if

$$\log \widetilde{q}(x,y) = b(y) + s \log q(x,y) \tag{2.26}$$

for some $s > 0$ and $b(y)$ (see Proposition 2.1). That is, we no longer have the freedom to choose $a(x)$ as in (2.21). Re-using the notation $\boldsymbol{\ell}_q \equiv \boldsymbol{\ell}_{\widetilde{q}}$ to mean equivalence according to (2.26), we have the following whenever $\ell_{00} \neq \ell_{10}$:

$$\boldsymbol{\ell}_q \equiv \begin{bmatrix} \ell_{00} - \ell_{10} & 0 \\ 0 & \ell_{11} - \ell_{01} \end{bmatrix} \equiv \begin{bmatrix} \pm 1 & 0 \\ 0 & \lambda \end{bmatrix}, \tag{2.27}$$

where the ± 1 entry equals the sign of $\ell_{00} - \ell_{10}$, and $\lambda = \pm(\frac{\ell_{11} - \ell_{01}}{\ell_{00} - \ell_{10}})$ may be positive or negative. In this form, the decoder can be interpreted as considering a *weighted* Hamming distance.

[4] This claim can be verified, for example, via Lemma 2.6 below.

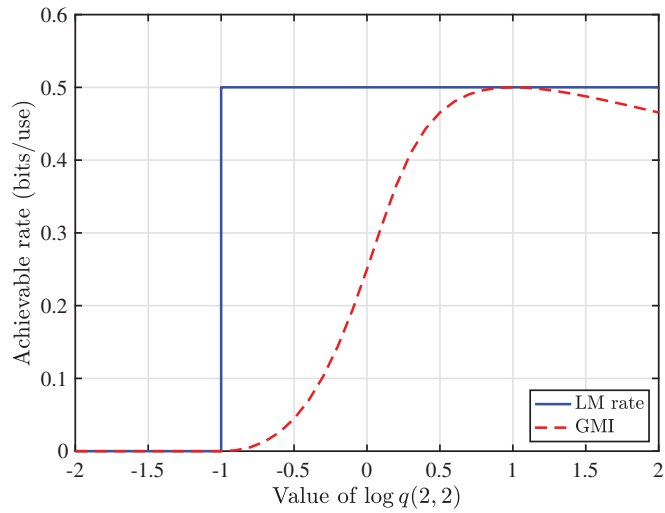


Figure 2.1: Binary example: GMI and LM rate as a function of $\lambda = \log q(2,2)$ with $\delta = 0.11$, under equiprobable inputs and a log-decoding metric of the form (2.27).

If we reduce both $\log W(y|x)$ and $\log q(x,y)$ to the form in (2.27), then even if the signs match, the achievable rate can degrade due to the mismatch. We demonstrate this in Figure 2.1, where we plot the GMI (as well as the LM rate) in the following setting:

- The channel is a BSC with crossover probability $\delta = 0.11$, hence having a matched capacity of roughly 0.5 bits/use.
- The (log-)decoding metric takes the form on the right-hand side of (2.27) with first entry $+1$, and with $\lambda = \log q(1,1)$ taking some value in the range $[-2, 2]$. The choice $\lambda = 1$ recovers optimal minimum Hamming distance decoding.
- The input distribution is $Q_X = (\frac{1}{2}, \frac{1}{2})$, which achieves the matched capacity under matched decoding.

In accordance with the above findings, the LM rate is zero for $\lambda < -1$ and equals the matched capacity for $\lambda > -1$. On the other hand, the GMI only achieves the matched capacity for $\lambda = 1$.

2.4.2 Symmetric Channels and Metrics

In the previous example, we saw that the gap between the GMI and LM rate grows when more asymmetry is *incorrectly* introduced into the decoding metric. In this example, we complement this finding via the following claim: Under matching output-symmetry and an equiprobable input distribution Q_X, the GMI and LM rate are identical.

The notion of matching output-symmetry needs some clarification; we adopt the following definition of Gallager [38, Sec. 4.5] regarding W, and extend it to the metric q in the natural way.

Definition 2.3 (Output-Symmetric Channel and Metric)**.** A DMC $W(y|x)$ is said to be output-symmetric if there exists a partition $\mathcal{Y}_1, \ldots, \mathcal{Y}_k$ of $\mathcal{Y}$ such that the following holds: Given the $|\mathcal{X}| \times |\mathcal{Y}|$ transition matrix corresponding to W, the $|\mathcal{X}| \times |\mathcal{Y}_j|$ sub-matrix corresponding to each $\mathcal{Y}_j$ $(j = 1, \ldots, k)$ is such that all rows are permutations of each other, and all columns are permutations of each other. The notion of output-symmetry for a decoding metric $q(x, y)$ is defined similarly.

A mismatched DMC (W, q) is said to exhibit matching output-symmetry if both W and q are output-symmetric with the same partition $\mathcal{Y}_1, \ldots, \mathcal{Y}_k$.

We briefly give two concrete examples of output-symmetry: For the binary symmetric channel one adopts the trivial partition $\mathcal{Y}_1 = \mathcal{Y} = \{0, 1\}$, whereas for the binary erasure channel, the two partitions are $\mathcal{Y}_1 = \{0, 1\}$ and $\mathcal{Y}_2 = \{e\}$ (the erasure symbol).

We can now formally state the following result, which, to the best of our knowledge, has not appeared previously.

Lemma 2.4 (GMI and LM Rate Under Output Symmetry)**.** For any pair (W, q) exhibiting matching output symmetry, if Q_X is the equiprobable distribution on $\mathcal{X}$, then

$$I_{\mathrm{GMI}}(Q_X) = I_{\mathrm{LM}}(Q_X). \tag{2.28}$$

The proof amounts to showing that the choice $a(x) = 0$ in (2.16) satisfies necessary and sufficient optimality conditions, implying that the rate coincides with (2.10). The details are given in Appendix A.2.

In the following example, we show that even when the pair (W, q) satisfies a stronger symmetry assumption (namely, Definition 2.3 restricted to $k = 1$), the achievable rates $I_{\mathrm{GMI}}(Q_X)$ and $I_{\mathrm{LM}}(Q_X)$ may differ when Q_X is non-equiprobable, and the LM rate can be strictly higher when Q_X is optimized.

2.4.3 Zero-Undetected Error Capacity

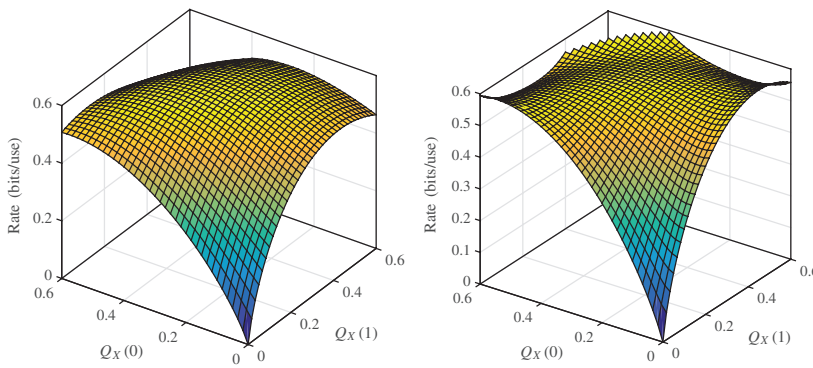
As we saw in Section 1.2.7, this choice $q(x, y) = \mathbb{1}\{W(y|x) > 0\}$ corresponds to the *zero-undetected error capacity*. Here we consider a specific example from [2], in which $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$, and the channel and decoding metric are given by

$$\boldsymbol{W} = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.75 & 0.25 \\ 0.25 & 0 & 0.75 \end{bmatrix}, \qquad \boldsymbol{q} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \tag{2.29}$$

where x indexes the rows and y indexes the columns.

In Figure 2.2, we plot the LM rate as a function of $(Q_X(0), Q_X(1))$, with the remaining entry $Q_X(2)$ determined so that the three sum to one. The optimal choices of Q_X are given as follows:

- For the GMI, $Q_X = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is optimal. Under this choice, $I_{\mathrm{GMI}}(Q_X) = I_{\mathrm{LM}}(Q_X) = 0.585$ bits/use; the two rates are identical in accordance with Lemma 2.4.



(a) GMI **(b) LM rate**

Figure 2.2: Zero-undetected error capacity example: GMI and LM rate for the channel and decoding metric in (2.29), as a function of $Q_X(0)$ and $Q_X(1)$.

- For the LM rate, the choice $Q_X = (0.449, 0.551, 0)$ (or any cyclic shift thereof) is optimal. Under this choice, $I_{\text{GMI}}(Q_X) = 0.502$ bits/use and $I_{\text{LM}}(Q_X) = 0.596$ bits/use.

We see that the LM rate exceeds the GMI even after the optimization of Q_X. In Section 6.3, we will present an achievable rate that further improves on the LM rate in this example, implying that the LM rate is strictly smaller than the mismatch capacity, i.e., $C_{\text{LM}} < C_{\text{M}}$. Of course, this implies the same for the GMI, which is no higher than the LM rate.

2.4.4 Compound Channels

In this example (taken from [26]), we consider the case that the channel is only known to lie in some set $\mathcal{W}$, but the exact channel within the set is unknown. This is commonly referred to as the *compound channel* model, and was discussed in Section 1.3. We initially focus on the case that $\mathcal{W}$ is closed and convex, and then comment on the more general case.

Using i.i.d. random coding with an input distribution Q_X, along with a fixed decoding metric q, we can achieve any rate up to

$$R_{\mathcal{W}}(Q_X, q) \triangleq \min_{W \in \mathcal{W}} I_{\text{GMI}}(Q_X, W, q), \tag{2.30}$$

where $I_{\text{GMI}}(Q_X, W, q)$ denotes the GMI with an explicit dependence on the channel and metric. We will show that upon optimizing Q_X and q, this recovers the best possible rate (i.e., the compound channel capacity), given as follows in the case that $\mathcal{W}$ is closed and convex [26, 56]:

$$C^* = \max_{Q_X} \min_{W \in \mathcal{W}} I(Q_X, W) = \min_{W \in \mathcal{W}} \max_{Q_X} I(Q_X, W), \tag{2.31}$$

where $I(Q_X, W) \triangleq I_{Q_X \times W}(X; Y)$.

Let (Q_X^*, W^*) denote the pair achieving the saddle-point in (2.31). We choose $q(x, y) = W^*(y|x)$ and $Q_X(x) = Q_X^*(x)$, and proceed by lower bounding the achievable rate in (2.30). For any $W \in \mathcal{W}$, setting $s = 1$ in (2.10) yields

$$I_{\text{GMI}}(Q_X^*, W, W^*) \geq \sum_{x,y} Q_X^*(x) W(y|x) \log \frac{W^*(y|x)}{P_Y^*(y)}, \tag{2.32}$$

where P_Y^* is the Y-marginal of the joint distribution $Q_X^* \times W^*$. It now only remains to prove that

$$\sum_{x,y} Q_X^*(x)W(y|x) \log \frac{W^*(y|x)}{P_Y^*(y)} \geq I(Q_X^*, W^*), \tag{2.33}$$

since the right-hand side equals C^* in (2.31) by definition. To establish (2.33), we first note the following for any $\lambda \in [0,1]$:

$$I(Q_X^*, \lambda W + (1-\lambda)W^*) \geq I(Q_X^*, W^*), \tag{2.34}$$

which follows from the convexity of $\mathcal{W}$ and the fact that W^* minimizes $I(Q_X^*, W')$ among $W' \in \mathcal{W}$. By taking $\lambda \to 0$, it follows that the derivative of the left-hand side of (2.34) with respect to λ must be non-negative at $\lambda = 0$, and it is a simple differentiation exercise to show that this fact is equivalent to (2.33). It follows that $R_{\mathcal{W}}(Q_X^*, W^*) \geq C^*$, as desired.

If we remove the assumption that $\mathcal{W}$ is convex, then the best possible single-letter decoding metric $q(x,y)$ may still yield a highly suboptimal rate. For instance, from the example in Section 2.4.1, if $\mathcal{X} = \mathcal{Y} = \{0,1\}$ and $\mathcal{W}$ contains the deterministic channels with $Y = X$ and $Y = 1 - X$, then any maximum-metric decoder gives zero rate. However, the compound channel capacity is 1 bit/use, and can be achieved by a standard universal decoder [24, Ex. 6.20].

2.4.5 Parallel Channels

In this example (taken from [58]), we present another channel (in addition to that of Section 2.4.3) for which the LM rate is known to be strictly smaller than the mismatch capacity.

We let $\mathcal{X} = \mathcal{Y} = \{0,1\}^2$, and write the input and output as binary pairs, $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$. The output Y_1 is formed by passing X_1 through a binary symmetric channel (BSC), and similarly for Y_2 and X_2, with the two BSCs being independent of each other; see Figure 2.3. The corresponding crossover probabilities are denoted by $\delta_1, \delta_2 \in (0, \frac{1}{2})$, and the mismatched decoder incorrectly assumes that both crossover probabilities are equal. Hence, the channel is given by $W((y_1,y_2)|(x_2,x_2)) = W_1(y_1|x_2)W_2(y_2|x_2)$, and the decoding metric

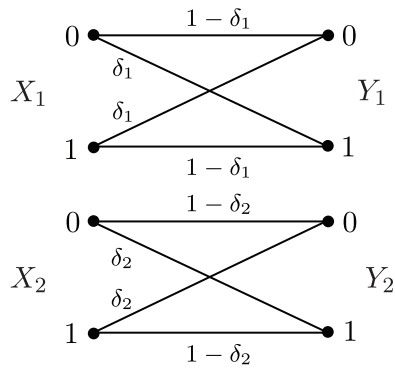


Figure 2.3: Parallel binary symmetric channels with crossover probabilities (δ_1, δ_2).

can be taken as

$$\log q((x_1, x_2), (y_1, y_2)) = \frac{1}{2}(\mathbb{1}\{x_1 = y_1\} + \mathbb{1}\{x_2 = y_2\}), \tag{2.35}$$

meaning that both $x_1 \neq y_1$ and $x_2 \neq y_2$ are penalized equally.

We let Q_X be the equiprobable distribution on $\{0,1\}^2$, and evaluate the LM rate using the primal expression. For any joint distribution $\widetilde{P}_{XY}$ satisfying the constraints in (2.15), we have

$$\begin{aligned} I_{\widetilde{P}}(X;Y) &= I_{\widetilde{P}}(X_1, X_2; Y_1, Y_2) && (2.36)\\ &= I_{\widetilde{P}}(X_1; Y_1, Y_2) + I_{\widetilde{P}}(X_2; Y_1, Y_2 | X_1) && (2.37)\\ &\geq I_{\widetilde{P}}(X_1; Y_1) + I_{\widetilde{P}}(X_2; Y_2), && (2.38) \end{aligned}$$

where (2.37) follows from the chain rule, and (2.38) follows by trivially lower bounding the first term, and bounding the second term as follows:

$$\begin{aligned} I_{\widetilde{P}}(X_2; Y_1, Y_2 | X_1) &= H_{\widetilde{P}}(X_2|X_1) - H_{\widetilde{P}}(X_2|Y_1, Y_2, X_1) && (2.39)\\ &\geq H_{\widetilde{P}}(X_2) - H_{\widetilde{P}}(X_2|Y_2) && (2.40)\\ &= I_{\widetilde{P}}(X_2; Y_2), && (2.41) \end{aligned}$$

where (2.40) uses the independence of (X_1, X_2) following from $\widetilde{P}_X = Q_X$ (with Q_X chosen to be equiprobable on $\{0,1\}^2$), and the fact that conditioning does not increase entropy.

The constraints $\widetilde{P}_X = Q_X$ and $\widetilde{P}_Y = P_Y$ in (2.15) ensure that X_1, X_2, Y_1, and Y_2, are each equiprobable on $\{0,1\}$ under $\widetilde{P}_{XY}$. This is

only possible if the conditional marginals $\widetilde{P}_{Y_1|X_1}$ and $\widetilde{P}_{Y_2|X_2}$ are BSCs. Letting $H_2(\alpha) = \alpha \log \frac{1}{\alpha} + (1-\alpha) \log \frac{1}{1-\alpha}$ denote the binary entropy function, we deduce from (2.38) that the following holds for some auxiliary crossover probabilities $\widetilde{\delta}_1, \widetilde{\delta}_2 \in (0,1)$:

$$I_{\widetilde{P}}(X;Y) \geq (1 - H_2(\widetilde{\delta}_1)) + (1 - H_2(\widetilde{\delta}_1)) \tag{2.42}$$

$$\geq 2\Big(1 - H_2\Big(\frac{\widetilde{\delta}_1 + \widetilde{\delta}_2}{2}\Big)\Big) \tag{2.43}$$

$$\geq 2\Big(1 - H_2\Big(\frac{\delta_1 + \delta_2}{2}\Big)\Big), \tag{2.44}$$

where (2.43) follows from the concavity of entropy, and (2.44) follows since the constraint $\mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]$ reduces to $\frac{\widetilde{\delta}_1+\widetilde{\delta}_2}{2} \leq \frac{\delta_1+\delta_2}{2}$ under the metric in (2.35).

We now observe that all the inequalities up to and including (2.44) hold with equality when $\widetilde{\delta}_1 = \widetilde{\delta}_2 = \frac{\delta_1+\delta_2}{2}$, meaning the corresponding joint distribution $\widetilde{P}_{XY}$ must achieve the minimum in (2.15). Hence, we have shown that

$$I_{\mathrm{LM}}(Q_X) = 2\Big(1 - H_2\Big(\frac{\delta_1 + \delta_2}{2}\Big)\Big). \tag{2.45}$$

For $\delta_1 \neq \delta_2$, this is strictly smaller than the matched capacity $C = (1 - H_2(\delta_1)) + (1 - H_2(\delta_2))$, which is achieved by equiprobable Q_X. See Figure 2.4 for an illustration, where we fix $\delta_1 = 0.11$ and vary δ_2.

We will revisit this example in Section 5.4.1, and see that in fact $C_{\mathrm{M}} = C$, and that the LM rate is strictly smaller than the mismatch capacity.

2.5 Properties of the GMI and LM Rate

In this subsection, we overview some of the most important known properties of the GMI and LM rate.

2.5.1 Basic Properties

We start by stating the important property of continuity of the GMI and LM rate as a function of the input distribution and channel. This

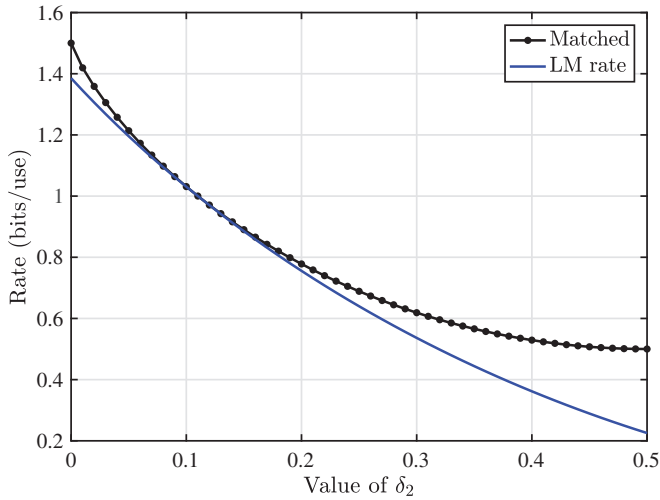


Figure 2.4: Parallel BSC example: Matched and mismatched achievable rates, with $\delta_1 = 0.11$ and various δ_2.

property will be used in the derivations of the primal expressions (2.9) and (2.15) based on the method of types.

In the following, recall that $I_{\text{GMI}}(Q_X, W)$ and $I_{\text{LM}}(Q_X, W)$ denote the GMI and LM rate with an explicit dependence on the channel, for a fixed metric q. The following was proved in [26].

Lemma 2.5 (Continuity)**.** In the discrete memoryless setting with a given decoding metric q, the achievable rates $I_{\text{GMI}}(Q_X, W)$ and $I_{\text{LM}}(Q_X, W)$ are continuous in the pair (Q_X, W) within the space of pairs satisfying

$$(Q_X(x) > 0 \cap q(x, y) = 0) \implies W(y|x) = 0. \tag{2.46}$$

The proof is somewhat technical, and is deferred to Appendix A.1. We note that the condition (2.46) is mild, since both achievable rates are zero when it fails [26]. Indeed, if (2.46) fails then we have $\mathbb{E}_{Q_X \times W}[\log q(x, y)] = -\infty$, which implies that the choice $\widetilde{P}_{XY} = Q_X \times P_Y$ in (2.15) is feasible, and gives $I_{\widetilde{P}}(X; Y) = 0$. At a more intuitive level, if $q(x, y) = 0$ and $Q_X(x)W(y|x) > 0$ then the metric associated with the transmitted codeword will be zero with high probability, precluding correct decoding.

Next, we present necessary and sufficient conditions for the GMI, LM rate, and mismatch capacity to be positive [26]. Notably, as long as we optimize the input distribution, both i.i.d. and constant-composition random coding achieve a positive rate whenever there exists *any* coding scheme achieving a positive rate.

Lemma 2.6 (Conditions for Positivity)**.** For any mismatched DMC (W, q) and input distribution Q_X, we have the following:

(i) $I_{\mathrm{GMI}}(Q_X) > 0 \iff I_{\mathrm{LM}}(Q_X) > 0$, with both being positive if and only if

$$\mathbb{E}_{Q_X \times W}[\log q(X,Y)] > \mathbb{E}_{Q_X \times P_Y}[\log q(X,Y)], \tag{2.47}$$

where P_Y is the Y-marginal of $P_{XY} = Q_X \times W$. Moreover, in the case that this condition holds, the primal expressions (2.9) and (2.15) remain unchanged when the inequality constraint is replaced by an equality constraint.

(ii) The mismatch capacity C_{M} satisfies

$$C_{\mathrm{M}} > 0 \iff C_{\mathrm{LM}} > 0 \iff C_{\mathrm{GMI}} > 0, \tag{2.48}$$

where C_{GMI} and C_{LM} are defined in (2.11) and (2.17), respectively.

Proof. In the primal expressions of both the GMI and LM rate (see (2.9) and (2.15)), the objective function is zero if and only if $\widetilde{P}_{XY} = Q_X \times P_Y$ is feasible. For the GMI, this follows since the KL divergence is zero if and only if the two distributions are equal, and for the LM rate, it follows from the marginal constraints and the fact that mutual information is zero if and only if X and Y are independent. The first claim follows since the choice $\widetilde{P}_{XY} = Q_X \times P_Y$ is feasible in the minimization problems if and only if (8.5) fails.

Next, recall that both (2.9) and (2.15) are convex minimization problems with linear constraints. If the inequality constraint were to be removed in either case, the product distribution would be feasible, yielding an objective of zero. This means that whenever the rate is positive, the inequality constraint is *active* (i.e., removing it changes the optimal objective value), and it is then a standard result in convex optimization theory that it must hold with equality [18, Sec. 5.5].

The proof of the second part is rather technical, so it is given in Appendix A.2. □

Lemma 2.6 is concerned with the distinction between a zero and non-zero rate, but does not give any indication of how close the achievable rates might be to the mismatch capacity. The following lemma addresses the other extreme, giving necessary and sufficient conditions for the achievable rate to equal the mutual information $I(Q_X, W) = I(X;Y)$ under $(X,Y) \sim Q_X \times W$, which is the rate that would be achieved under optimal decoding. Of course, if Q_X is a (matched) capacity-achieving input distribution, then achieving $I(Q_X, W)$ amounts to achieving the matched capacity, and hence also the mismatch capacity.

Lemma 2.7 (Conditions for Achieving the Matched Performance)**.** For any mismatched DMC (W, q) and input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, letting P_Y denote the Y-marginal of $P_{XY} = Q_X \times W$, we have:

(i) $I_{\text{GMI}}(Q_X, W) \leq I(Q_X, W)$, with equality if and only if

$$\inf_{s \geq 0, b(\cdot)} \max_{\substack{(x,y) \in \mathcal{X} \times \mathcal{Y}: \\ Q_X(x) > 0, P_Y(y) > 0}} |W(y|x) - q(x,y)^s e^{b(y)}| = 0. \tag{2.49}$$

(ii) $I_{\text{LM}}(Q_X, W) \leq I(Q_X, W)$, with equality if and only if

$$\inf_{s \geq 0, a(\cdot), b(\cdot)} \max_{\substack{(x,y) \in \mathcal{X} \times \mathcal{Y}: \\ Q_X(x) > 0, P_Y(y) > 0}} |W(y|x) - q(x,y)^s e^{a(x)} e^{b(y)}| = 0. \tag{2.50}$$

This result will be proved via an alternative formulation of the dual expressions after Lemma 2.8 below. We note that conditions (2.49) and (2.50) are slightly more general than the conditions

$$\log W(y|x) = b(y) + s \log q(x,y), \tag{2.51}$$

$$\log W(y|x) = a(x) + b(y) + s \log q(x,y), \tag{2.52}$$

which are easily seen to be sufficient for $I_{\text{GMI}}(Q_X) = I(Q_X, W)$ and $I_{\text{LM}}(Q_X) = I(Q_X, W)$ respectively: By Propositions 2.1–2.2, these conditions imply that mismatched decoding and maximum-likelihood decoding are equivalent when $s > 0$ (whereas if $s = 0$, it is straightforward to show that either of (2.51)–(2.52) imply $I(Q_X, W) = 0$).

To give an example of a case where the more general conditions (2.49)–(2.50) are required, consider the binary erasure channel (BEC): $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{0, e, 1\}$, and the channel and metric are described by the matrices

$$\boldsymbol{W} = \begin{bmatrix} 1-\epsilon & \epsilon & 0 \\ 0 & \epsilon & 1-\epsilon \end{bmatrix}, \quad \boldsymbol{q} = \begin{bmatrix} 1 & 1 & 0.5 \\ 0.5 & 1 & 1 \end{bmatrix} \tag{2.53}$$

for some $\epsilon \in (0, 1)$. Due to the entries with $W(y|x) = 0$, there do not exist fixed choices of s, $a(\cdot)$, and $b(\cdot)$ such that $W(y|x) = q(x, y)^s e^{a(x)} e^{b(y)}$ for all (x, y). However, if we raise each entry of $\boldsymbol{q}$ to the power of $s > 0$, then in the limit as $s \to \infty$, we recover the standard erasures-only metric $\widetilde{q}(x, y) = \mathbb{1}\{W(y|x) > 0\}$. Multiplying this metric by $e^{\widetilde{b}(y)}$ recovers the maximum-likelihood metric when we let $e^{\widetilde{b}(0)} = e^{\widetilde{b}(1)} = 1 - \epsilon$ and $e^{\widetilde{b}(e)} = \epsilon$. Hence, the condition (2.49) is satisfied (and consequently, so is the more general condition (2.50)).

It is also insightful to provide a direct proof that (2.51)–(2.52) are sufficient; we do this for the GMI, but a similar argument applies to the LM rate. First note that the upper bound $I_{\mathrm{GMI}}(Q_X) \le I(Q_X, W)$ follows immediately from the fact that the choice $\widetilde{P}_{XY} = P_{XY}$ (with $P_{XY} = Q_X \times W$) is feasible in (2.9). Moreover, for any feasible $\widetilde{P}_{XY}$, if (2.51) holds then we have

$$\begin{aligned}
&D(\widetilde{P}_{XY} \| Q_X \times P_Y) \\
&\quad = I_{\widetilde{P}}(X; Y) + D(\widetilde{P}_X \| Q_X) && (2.54)\\
&\quad \ge I_{\widetilde{P}}(X; Y) && (2.55)\\
&\quad = H_{\widetilde{P}}(Y) - H_{\widetilde{P}}(Y|X) && (2.56)\\
&\quad = H_P(Y) + \mathbb{E}_{\widetilde{P}}\left[\log \left(\frac{\widetilde{P}_{Y|X}(Y|X)}{W(Y|X)} W(Y|X) \right) \right] && (2.57)\\
&\quad = H_P(Y) + \mathbb{E}_{\widetilde{P}}[\log W(Y|X)] + D(\widetilde{P}_{Y|X} \| W | Q_X) && (2.58)\\
&\quad \ge H_P(Y) + \mathbb{E}_{\widetilde{P}}[\log W(Y|X)] && (2.59)\\
&\quad \ge H_P(Y) + \mathbb{E}_P[\log W(Y|X)] && (2.60)\\
&\quad = I_P(X; Y), && (2.61)
\end{aligned}$$

where (2.54) and (2.57) use the constraint $\widetilde{P}_Y = P_Y$, (2.55) and (2.59) follow from the non-negativity of KL divergence, and (2.60) follows since

the metric constraint $\mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]$ is equivalent to $\mathbb{E}_{\widetilde{P}}[\log W(Y|X)] \geq \mathbb{E}_P[\log W(Y|X)]$ when (2.51) holds with $s > 0$ (recall that $\widetilde{P}_Y = P_Y$). We conclude that the choice $\widetilde{P}_{XY} = P_{XY}$ must achieve the minimum in (2.9).

Next, we provide alternative forms of the dual expressions, which will also allow us to prove Lemma 2.7.

Lemma 2.8 (Alternative Formulations of the Dual Expressions)**.** For any mismatched DMC (W, q), we have the following:

(i) The GMI and LM rate can be written as

$$I_{\mathrm{GMI}}(Q_X, W) = I(Q_X, W) - \inf_{s \geq 0} D(W' \| V'_s | P_Y), \tag{2.62}$$

$$I_{\mathrm{LM}}(Q_X, W) = I(Q_X, W) - \inf_{s \geq 0, a(\cdot)} D(W' \| V'_{s,a} | P_Y), \tag{2.63}$$

where $W'(x|y) = P_{X|Y}(x|y) = \frac{W(y|x)Q_X(x)}{P_Y(y)}$ and $P_Y(y)$ is the marginal of $P_{XY} = Q_X \times W$, and where

$$V'_s(x|y) = \frac{Q_X(x)q(x,y)^s}{\sum_{\overline{x}} Q_X(\overline{x})q(x,y)^s}, \tag{2.64}$$

$$V'_{s,a}(x|y) = \frac{Q_X(x)q(x,y)^s e^{a(x)}}{\sum_{\overline{x}} Q_X(\overline{x})q(\overline{x},y)^s e^{a(\overline{x})}}. \tag{2.65}$$

(ii) The LM rate can be written as

$$I_{\mathrm{LM}}(Q_X, W) = \sup_{s \geq 0, b(\cdot)} \sum_{x,y} Q_X(x)W(y|x) \log \frac{q(x,y)^s e^{b(y)}}{\sum_{\overline{y}} P_Y(\overline{y})q(x,\overline{y})^s e^{b(\overline{y})}} \tag{2.66}$$

$$= I(Q_X, W) - \inf_{s \geq 0, b(\cdot)} D(W \| V_{s,b} | Q_X), \tag{2.67}$$

where

$$V_{s,b}(y|x) = \frac{P_Y(y)q(x,y)^s e^{b(y)}}{\sum_{\overline{y}} P_Y(\overline{y})q(x,\overline{y})^s e^{b(\overline{y})}}. \tag{2.68}$$

Proof. To prove the first part, we write the LM rate (2.16) as

$$I_{\mathrm{LM}}(Q_X) = \sup_{s \geq 0, a(\cdot)} \sum_{x,y} P_Y(y)W'(x|y) \log \left(\frac{V'_{s,a}(x|y)}{Q_X(x)} \cdot \frac{W'(x|y)}{W'(x|y)} \right), \tag{2.69}$$

where we have used the fact that $P_Y \times W' = Q_X \times W = P_{XY}$, along with the definition of $V'_{s,a}$. Splitting the logarithm as $\log\left(\frac{V'_{s,a}(x|y)}{Q_X(x)} \cdot \frac{W'(x|y)}{W'(x|y)}\right) = \log \frac{W'(x|y)}{Q_X(x)} - \log \frac{W'(x|y)}{V'_{s,a}(x|y)}$ and re-arranging yields (2.63), and the GMI is handled in an identical manner.

To prove the second part, we define $\phi_a = \mathbb{E}_Q[a(X)]$ and $\phi_b = \mathbb{E}_P[b(Y)]$, and proceed as follows:

$$
\begin{aligned}
I_{\mathrm{LM}}(Q_X) = & \sup_{s \geq 0, a(\cdot)} \sum_{x,y} Q_X(x) W(y|x) \log q(x,y)^s \\
& - \sum_y P_Y(y) \log \sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x},y)^s e^{a(\overline{x}) - \phi_a} \qquad (2.70) \\
= & \sup_{s \geq 0, a(\cdot), b(\cdot)} \sum_{x,y} Q_X(x) W(y|x) \log q(x,y)^s \\
& - \sum_y P_Y(y) \log \sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x},y)^s e^{a(\overline{x}) - \phi_a} e^{b(y) - \phi_b} \qquad (2.71) \\
= & \sup_{s \geq 0, a(\cdot), b(\cdot)} \sum_{x,y} Q_X(x) W(y|x) \log q(x,y)^s \\
& - \log \sum_{\overline{x},y} Q_X(\overline{x}) P_Y(y) q(\overline{x},y)^s e^{a(\overline{x}) - \phi_a} e^{b(y) - \phi_b} \qquad (2.72)
\end{aligned}
$$

where (2.70) follows by re-arranging (2.16), (2.71) follows since the factor $e^{b(y)-\phi_b}$ cancels to zero upon expanding the logarithm, and (2.72) follows by moving the average over Y inside the logarithm using Jensen's inequality, and noting that $b(y)$ can always be chosen to make Jensen's inequality hold with equality (i.e., such that $\sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x},y)^s e^{a(\overline{x})-\phi_a} e^{b(y)-\phi_b}$ does not vary with y).

Renaming $(\overline{x}, y)$ as $(x, \overline{y})$ in (2.72) and reversing the steps of (2.70)–(2.72) with the roles of x and y interchanged, we obtain (2.66). Finally, (2.67) follows from (2.66) by the same argument as the first part. □

The equivalence in (2.66) can be understood by noting that the primal expression (2.15) for the LM rate is symmetric with respect to X and Y, so the roles of the two can also be swapped in the dual expression. In contrast, the GMI does not exhibit such symmetry.

As discussed in Section 2.3.3, the formulation in (2.67) has the natural interpretation that the loss due to the mismatch is the smallest conditional KL divergence $D(W\|V|Q_X)$ among the set of conditional distributions $V(y|x) \in \mathcal{P}(\mathcal{Y}|\mathcal{X})$ that are equivalent to $q(x,y)$ in the sense of Proposition 2.2.[5] For the GMI, the interpretation is slightly less natural, with the loss taking the form $D(W'\|V'_s|P_Y)$ for the "reverse channel" $W'(x|y)$.

We conclude by using Lemma 2.8 to prove Lemma 2.7. We focus on the GMI, but analogous arguments apply for the LM rate. The idea is to note that $D(W'\|V'_s|P_Y) \geq 0$, with equality if and only if $V'_s(x|y) = W'(x|y)$ for all y such that $P_Y(y) > 0$ [22, Thm. 2.6.3]. Suppose first that the supremum in (2.49) is attained by a finite value of s. Writing $V'_s(x|y) = \frac{1}{\Omega(y)} Q_X(x) q(x,y)^s$ with $\Omega(y) = \sum_{\bar{x}} Q_X(\bar{x}) q(\bar{x},y)^s$, the necessary and sufficient condition $V'_s(x|y) = W'(x|y)$ becomes

$$W'(x|y) = \frac{1}{\Omega(y)} Q_X(x) q(x,y)^s, \tag{2.73}$$

and substituting $W'(x|y) = \frac{Q_X(x)W(y|x)}{P_Y(y)}$ yields the equivalent condition

$$W(y|x) = \frac{1}{\Omega(y)} P_Y(y) q(x,y)^s \tag{2.74}$$

for all (x,y) with $Q_X(x) > 0$ and $P_Y(y) > 0$. Defining $b(y)$ to satisfy $e^{b(y)} = \frac{P_Y(y)}{\Omega(y)}$, we find that this condition coincides with (2.49).

In the case that (2.49) is not attained by a finite value of s, the argument is similar but more tedious, so we omit the details. While the KL divergence is discontinuous in cases where the second argument has a zero transition but the first argument does not, such cases are not relevant here, because we are only considering conditions for vanishing KL divergence.

[5] Assuming without loss of generality that $P_Y(y) > 0$ for all y, we can rename $P_Y(y)e^{b(y)}$ as $e^{\widetilde{b}(y)}$ for some $\widetilde{b}(y)$. Moreover, to ensure that $V(y|x)$ is a valid conditional distribution, one must set $a(x) = -\log \sum_{\bar{y}} P_Y(\bar{y}) q(x,\bar{y}) e^{b(\bar{y})}$ when considering (2.5) with $\widetilde{q}(x,y) = V(y|x)$.

2.5.2 Ensemble Tightness

The vast majority of the mismatched decoding literature has focused on achievable rates, whereas upper bounds on C_{M} have mostly remained elusive (see Section 8 for an overview). We have also seen that there are mismatched channels for which the GMI and LM rate are strictly smaller than C_{M}, i.e., the achievable rates are not tight.

Based on these observations, it is natural to ask the following: *Is the weakness in the GMI and LM rate due to an inherent weakness in i.i.d. and constant-composition random coding, or is it due to loose bounds in the mathematical analysis?* This question is answered by the following *ensemble tightness* result [65], showing that i.i.d. random coding cannot achieve any rate higher than the GMI, and constant-composition random coding cannot achieve any rate better than the LM rate. Thus, we cannot hope to obtain improved rates by sharpening the mathematical analysis; we must devise alternative codebook constructions.

In the following, $\bar{p}_{\mathrm{e}} = \bar{p}_{\mathrm{e}}(n, M)$ denotes the random-coding error probability with block length n and codebook size M.

Lemma 2.9 (Ensemble Tightness)**.** For any mismatched DMC and a given input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, we have the following:

- Under i.i.d. random coding, $\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor) \to 1$ as $n \to \infty$ for any $R > I_{\mathrm{GMI}}(Q_X)$;
- Under constant-composition random coding, $\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor) \to 1$ as $n \to \infty$ for any $R > I_{\mathrm{LM}}(Q_X)$.

The proof is given in Section 2.6.

2.5.3 Equivalence of Primal and Dual Expressions

Thus far, we have taken for granted that the primal and dual expressions of the GMI (2.9)–(2.10) are equal, and similarly for the LM rate (2.15)–(2.16). The following lemma, adapted from [65], provides the justification for doing so.

Lemma 2.10 (Primal-Dual Equivalence)**.** For any mismatched DMC (W, q) and a given input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, the right-hand sides of (2.9) and (2.10) are equal, and similarly for (2.15) and (2.16).

Proof. We use a direct analysis to show that the primal expression for the LM rate is lower bounded by the dual expression. The matching upper bound is deduced using Lagrange duality techniques [18]; the details are omitted here, and can be found in [77, Appendix E]. The GMI is handled in a similar (and slightly simpler) manner, so its details are also omitted.

For any $\widetilde{P}_{XY}$ satisfying the constraints in (2.15), the mutual information can be lower bounded as follows for any $s \geq 0$ and $a(\cdot)$:

$$
\begin{aligned}
&I_{\widetilde{P}}(X;Y) \\
&= \sum_{x,y} \widetilde{P}_{XY}(x,y) \log \frac{\widetilde{P}_{XY}(x,y)}{Q_X(x)P_Y(y)} && (2.75)\\
&\geq \sum_{x,y} \widetilde{P}_{XY}(x,y) \log \frac{\widetilde{P}_{XY}(x,y)}{Q_X(x)P_Y(y)} + \sum_x a(x)(Q_X(x) - \widetilde{P}_X(x)) \\
&\quad + s \sum_{x,y} \log q(x,y)(P_{XY}(x,y) - \widetilde{P}_{XY}(x,y)) && (2.76)\\
&= \sum_{x,y} \widetilde{P}_{XY}(x,y) \log \frac{\widetilde{P}_{XY}(x,y)}{Q_X(x)P_Y(y)q(x,y)^s e^{a(x)}} + \sum_x Q_X(x)a(x) \\
&\quad + s \sum_{x,y} P_{XY}(x,y) \log q(x,y) && (2.77)\\
&\geq \sum_y \widetilde{P}_Y(y) \log \frac{\widetilde{P}_Y(y)}{\sum_{\overline{x}} Q_X(\overline{x})P_Y(y)q(\overline{x},y)^s e^{a(\overline{x})}} + \sum_x Q_X(x)a(x) \\
&\quad + s \sum_{x,y} P_{XY}(x,y) \log q(x,y) && (2.78)\\
&= \sum_{x,y} P_{XY}(x,y) \log \frac{q(x,y)^s e^{a(x)}}{\sum_{\overline{x}} Q_X(\overline{x})q(\overline{x},y)^s e^{a(\overline{x})}}, && (2.79)
\end{aligned}
$$

where (2.75) follows from the marginal constraints $\widetilde{P}_X = Q_X$ and $\widetilde{P}_Y = P_Y$, (2.76) follows from the X-marginal constraint and the metric constraint $\mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]$ with $P_{XY} = Q_X \times W$, (2.77) follows from simple re-arranging, (2.78) follows from Jensen's

inequality in the form of the log-sum inequality $\sum_{i=1}^{n} \alpha_i \log \frac{\alpha_i}{\beta_i} \geq (\sum_{i=1}^{n} \alpha_i) \log \frac{\sum_{i=1}^{n} \alpha_i}{\sum_{i=1}^{n} \beta_i}$ [22, Sec. 2.7], and (2.79) follows from the marginal constraints along with simple re-arranging. Taking the supremum over $s \geq 0$ and $a(\cdot)$ establishes the desired claim. □

2.5.4 Additional Properties: Differences to Mutual Information

Along with the above properties that the GMI and LM rate satisfy, it is useful to highlight certain properties that, despite holding in the matched setting (i.e., for the mutual information), may fail to hold in the mismatched setting. The following lemmas, adapted from [26, 39, 65], provide three such properties.

Lemma 2.11 (Non-Tightness). There exist mismatched DMCs for which $C_{\mathrm{LM}} < C_{\mathrm{M}}$ (and consequently $C_{\mathrm{GMI}} < C_{\mathrm{M}}$).

The claim regarding the GMI follows directly from that of the LM rate, since $C_{\mathrm{GMI}} \leq C_{\mathrm{LM}}$ (see (2.18)). There are several known examples showing that the LM rate is not tight:

- In Section 5.4, we will return to the parallel BSC example of Section 2.4.5, and see that the mismatch capacity equals the matched capacity, which is strictly higher than the LM rate.
- In Section 6.3.3, we will return the zero-undetected error example of Section 2.4.3, and see that the optimized LM rate can be improved upon.
- In Section 6.3.4, we provide an example of a binary-input DMC for which C_{LM} is strictly smaller than the mismatch capacity, thus providing a counter-example to a reported converse for the binary-input case [14].
- The zero-error capacity (*cf.* Section 1.2.8) is a special case of the mismatch capacity, and it is known that i.i.d. and constant-composition random coding can be strictly suboptimal in this context [26].

The first and last of these examples lead to analytical proofs of Lemma 2.11, whereas the other two are numerical.

Lemma 2.12 (Non-Concavity in Q_X)**.** There exist mismatched DMCs (W, q) such that the functions $I_{\text{GMI}}(Q_X)$ and $I_{\text{LM}}(Q_X)$ are non-concave in the input distribution Q_X.

We omit a formal proof, and instead discuss how this is established:

- For the LM rate, the non-concavity can be observed in the example in Figure 2.2. In the same example, it was shown *analytically* in [101, Ch. 4] that all three permutations of $Q_X = (\frac{1}{2}, \frac{1}{2}, 0)$ yield the same LM rate, and that this rate is strictly higher than that of $Q_X = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Due to Jensen's inequality, this would not be possible if the LM rate were concave in Q_X.
- An alternative proof for the LM rate can be found in [39] based on the mismatch capacity per unit cost.
- The non-concavity of the GMI can be inferred from a result on the zero-undetected error capacity given in [21]; we provide the details in Appendix A.2.

Finally, again recalling the notation $I_{\text{GMI}}(Q_X, W)$ and $I_{\text{LM}}(Q_X, W)$ with an explicit dependence on the channel, we have the following.

Lemma 2.13 (Lack of Data Processing Inequality)**.** There exist scenarios in which $I_{\text{GMI}}(Q_X, W) < I_{\text{GMI}}(Q_X, T \circ W)$ and $I_{\text{LM}}(Q_X, W) < I_{\text{LM}}(Q_X, T \circ W)$ for some transformation $T\colon \mathcal{Y} \to \mathcal{Y}$, where $\circ$ denotes the composition operation.

This result is in contrast to the mutual information $I(Q_X, W)$ with respect to $(X, Y) \sim Q_X \times W$, for which we know that $I(Q_X, W) \geq I(Q_X, T \circ W)$ for any transformation T by the data processing inequality [22, Sec. 2.8].

To establish Lemma 2.13, consider the following example: Let $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, let W be the noiseless binary channel $X = Y$, and let q be the maximum-likelihood metric of the binary channel with $Y = 1 - X$ (i.e., a deterministic flip of the input). From the example in

Section 2.4.1, $I_{\mathrm{GMI}}(Q_X, W) = I_{\mathrm{LM}}(Q_X, W) = 0$. However, if we choose T to represent a deterministic bit flip, we obtain from the same example that $I_{\mathrm{GMI}}(Q_X, T \circ W) = I_{\mathrm{LM}}(Q_X, T \circ W) = 1$ bit/use.

In fact, since we know from Lemma 2.6 that $C_{\mathrm{M}} > 0 \iff C_{\mathrm{LM}} > 0 \iff C_{\mathrm{GMI}} > 0$, this example demonstrates that the mismatch capacity C_{M} also does not satisfy the data processing inequality.

2.6 Proofs of Achievable Rates

In this subsection, we prove the achievability of the GMI and LM rate, as stated in Theorems 2.2 and 2.3, as well as their ensemble tightness, as stated in Lemma 2.9. We will provide primal and dual derivations for both the GMI and LM rate. While there is significant overlap in the proofs of the two rates, we present them separately so that their differences can be properly understood.

Throughout the subsection, we let $\bar{p}_{\mathrm{e}} = \bar{p}_{\mathrm{e}}(n, M)$ be the error probability averaged over the random codebook. By the symmetry of the random construction, the random-coding error probability is identical conditioned on any given value of m, so we may assume without loss of generality that $m = 1$. Recalling that ties are broken as errors, the starting point of all four achievability proofs will be the following:

$$\bar{p}_{\mathrm{e}} = \mathbb{E}\left[\mathbb{P}\left[\bigcup_{j=2}^{M}\left\{q^n(\boldsymbol{X}^{(j)}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}^{(1)}, \boldsymbol{Y})\right\} \,\middle|\, \boldsymbol{X}^{(1)}, \boldsymbol{Y}\right]\right] \tag{2.80}$$

$$\leq \mathbb{E}\Big[\min\{1, (M-1)\mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y}) | \boldsymbol{X}, \boldsymbol{Y}]\}\Big], \tag{2.81}$$

with $(\boldsymbol{X}, \boldsymbol{Y}, \overline{\boldsymbol{X}}) \sim P_{\boldsymbol{X}}(\boldsymbol{x}) W^n(\boldsymbol{y}|\boldsymbol{x}) P_{\boldsymbol{X}}(\overline{\boldsymbol{x}})$. In (2.80), we have written the probability as an expectation given $(\boldsymbol{X}^{(1)}, \boldsymbol{Y})$ for convenience, and in (2.81) we have applied the *truncated union bound*, i.e., the better out of the union bound and the trivial upper bound of one. Equation (2.81) is a simple extension of the *random-coding union bound* [69] to mismatched decoding.

In our analysis, we will handle the outer expectation over $(\boldsymbol{X}, \boldsymbol{Y})$ using the law of large numbers, which suffices for establishing the achievable rates. However, both the primal and dual analyses can be

strengthened to obtain *error exponents* establishing the exponential rate of decay of $\bar{p}_{\mathrm{e}}$; see Section 7.2 for details.

2.6.1 Preliminaries on the Method of Types

Here we present some useful properties of types that will be used in deriving the primal expressions. Since these are standard in the literature, we omit the proofs; further details can be found in [24, Ch. 2] and [37].

Recall from (2.12)–(2.13) that $Q_{X,n}$ is a type (i.e., an empirical distribution) approximating Q_X, and that $\mathcal{T}^n(Q_{X,n})$ is the set of all sequences having type $Q_{X,n}$. We will also need the notion of a *joint type class*, which is simply a type class in the case of a product alphabet, say $\mathcal{X} \times \mathcal{Y}$:

$$\mathcal{T}^n(\widetilde{P}_{XY}) = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \colon \widehat{P}_{\boldsymbol{xy}} = \widetilde{P}_{XY}\}, \tag{2.82}$$

where $\widehat{P}_{\boldsymbol{xy}}(x, y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{(x_i, y_i) = (x, y)\}$.

The properties that we will use are as follows:

- There are at most $(n+1)^{|\mathcal{X}|}$ types on $\mathcal{X}$, and at most $(n+1)^{|\mathcal{X}| \cdot |\mathcal{Y}|}$ joint types on $\mathcal{X} \times \mathcal{Y}$.
- The constant-composition codeword distribution can be upper bounded in terms of the corresponding i.i.d. distribution:

$$P_{\boldsymbol{X}}(\boldsymbol{x}) \le (n+1)^{|\mathcal{X}|} Q_{X,n}^n(\boldsymbol{x}), \tag{2.83}$$

  under the choice of $P_{\boldsymbol{X}}$ in Definition 2.2.
- For a fixed sequence $\boldsymbol{y} \in \mathcal{T}^n(P_Y)$, an independent codeword $\overline{\boldsymbol{X}} \sim Q_X^n$, and a joint type $\widetilde{P}_{XY}$ with $\widetilde{P}_Y = P_Y$, it holds for any $\delta > 0$ and sufficiently large n that

$$\begin{aligned} e^{-n(D(\widetilde{P}_{XY} \| Q_X \times P_Y) + \delta)} &\le \mathbb{P}[(\overline{\boldsymbol{X}}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{XY})] \\ &\le e^{-nD(\widetilde{P}_{XY} \| Q_X \times P_Y)}. \end{aligned} \tag{2.84}$$

- For a fixed sequence $\boldsymbol{y} \in \mathcal{T}^n(P_Y)$, a constant-composition codeword $\overline{\boldsymbol{X}} \sim P_{\boldsymbol{X}}$, and a joint type $\widetilde{P}_{XY}$ with $\widetilde{P}_X = Q_{X,n}$ and

$\widetilde{P}_Y = P_Y$, it holds for any $\delta > 0$ and sufficiently large n that

$$e^{-n(I_{\widetilde{P}}(X;Y)+\delta)} \le \mathbb{P}[(\overline{\boldsymbol{X}}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{XY})] \le e^{-n(I_{\widetilde{P}}(X;Y)-\delta)}. \tag{2.85}$$

- For $\boldsymbol{X}$ drawn from either the i.i.d. or constant-composition codeword distribution with input distribution Q_X, and $\boldsymbol{Y}$ conditionally drawn from $W^n(\cdot\,|\boldsymbol{x})$, the joint type P_{XY} of $(\boldsymbol{X}, \boldsymbol{Y})$ satisfies

$$\|P_{XY} - Q_X \times W\|_\infty \le \delta \tag{2.86}$$

  with probability approaching one, for arbitrarily small $\delta > 0$.

The last of these follows directly from the law of large numbers; the rest are non-trivial but still straightforward to derive [24, Ch. 2], [37].

2.6.2 Primal Derivation of the GMI

The key idea of the primal analysis is to observe that if $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}^n(P_{XY})$ and $(\overline{\boldsymbol{x}}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{XY})$, then the condition $q^n(\overline{\boldsymbol{x}}, \boldsymbol{y}) \ge q^n(\boldsymbol{x}, \boldsymbol{y})$ in (2.81) is equivalent to $\mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \ge \mathbb{E}_P[\log q(X,Y)]$. This follows immediately from the definition of a joint type following (2.82). Hence, for fixed $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}^n(P_{XY})$, we have

$$\begin{aligned}
&\mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{y}) \ge q^n(\boldsymbol{x}, \boldsymbol{y})] \\
&\quad= \sum_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}): \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \ge \mathbb{E}_P[\log q(X,Y)]}} \mathbb{P}[(\overline{\boldsymbol{X}}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{XY})] && (2.87) \\
&\quad\le \sum_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \ge \mathbb{E}_P[\log q(X,Y)]}} e^{-nD(\widetilde{P}_{XY}\|Q_X \times P_Y)} && (2.88) \\
&\quad\le (n+1)^{|\mathcal{X}|\cdot|\mathcal{Y}|} \max_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \ge \mathbb{E}_P[\log q(X,Y)]}} e^{-nD(\widetilde{P}_{XY}\|Q_X \times P_Y)} && (2.89) \\
&\quad\le (n+1)^{|\mathcal{X}|\cdot|\mathcal{Y}|} e^{-n(I_{\mathrm{GMI}}(P_{XY})+D(P_X\|Q_X))}, && (2.90)
\end{aligned}$$

where (2.88) follows from (2.84) and the fact that $\widetilde{P}_Y = P_Y$ by construction, (2.89) follows since there are at most $(n+1)^{|\mathcal{X}|\cdot|\mathcal{Y}|}$ joint types,

and (2.90) follows by upper bounding the maximum over joint types $\mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$ by that over all joint distributions $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$, and defining $I_{\text{GMI}}(P_{XY})$ to be the GMI in (2.10) with input distribution P_X and channel $P_{Y|X}$.

Next, recall from (2.86) that $\|P_{XY} - Q_X \times W\|_\infty \leq \delta$ with probability approaching one, and note that when this holds for sufficiently small δ, (2.90) can be weakened to

$$\mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})] \leq e^{-n(I_{\text{GMI}}(Q_X \times W) - \delta')} \tag{2.91}$$

for arbitrarily small δ' and sufficiently large n, by the continuity of the GMI (*cf.*, Lemma 2.5) and the fact that the KL divergence in (2.90) approaches zero as $\delta \to 0$. By upper bounding $\min\{1, \cdot\}$ in (2.81) by one whenever $\|P_{XY} - Q_X \times W\|_\infty > \delta$, and by (2.91) otherwise, we deduce that the error probability vanishes for any rate $R = \frac{1}{n} \log M$ arbitrarily close to the GMI, as desired.

2.6.3 Primal Derivation of the LM Rate

The core part of the argument is very similar to that of the GMI, with the main difference being that we can introduce the constraint $\widetilde{P}_X = P_X$ by the fact that all codewords have the same composition. Specifically, for $\overline{\boldsymbol{X}} \sim P_{\boldsymbol{X}}$ following the constant-composition codeword distribution in Definition 2.2, and for fixed $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}^n(P_{XY})$, we have

$$\begin{aligned}
&\mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})] \\
&\quad = \sum_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y}): \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} \mathbb{P}[(\overline{\boldsymbol{X}}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{XY})] && (2.92) \\
&\quad \leq \sum_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y}): \widetilde{P}_X = P_X, \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} e^{-nI_{\widetilde{P}}(X;Y)} && (2.93) \\
&\quad \leq (n+1)^{|\mathcal{X}| \cdot |\mathcal{Y}|} \max_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y}): \ \widetilde{P}_X = P_X, \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} e^{-nI_{\widetilde{P}}(X;Y)} && (2.94) \\
&\quad \leq (n+1)^{|\mathcal{X}| \cdot |\mathcal{Y}|} e^{-nI_{\text{LM}}(P_{XY})}, && (2.95)
\end{aligned}$$

where (2.93) follows from (2.85) and the fact that $\widetilde{P}_X = P_X$ and $\widetilde{P}_Y = P_Y$ by construction, (2.94) follows since there are at most $(n+1)^{|\mathcal{X}|\cdot|\mathcal{Y}|}$ joint types, and (2.90) follows by upper bounding the maximum over joint types $\mathcal{P}_n(\mathcal{X}\times\mathcal{Y})$ by that over all joint distributions $\mathcal{P}(\mathcal{X}\times\mathcal{Y})$, and defining $I_{\text{LM}}(P_{XY})$ to be the LM rate in (2.16) with input distribution P_X and channel $P_{Y|X}$.

The desired result now follows by substituting (2.95) into (2.81) and using the continuity of the LM rate (*cf.*, Lemma 2.5).

2.6.4 Dual Derivation of the GMI

The dual analysis bears some resemblance to that of Gallager [38, Ch. 5] for maximum-likelihood decoding. Raising both of the q^n terms to the power of $s > 0$ in (2.81) clearly leaves the probability unchanged, and upon applying Markov's inequality, we obtain[6]

$$\bar{p}_{\text{e}} \le \mathbb{E}\left[\min\left\{1, M\mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}},\boldsymbol{Y})}{q^n(\boldsymbol{X},\boldsymbol{Y})}\right)^s \,\middle|\, \boldsymbol{X},\boldsymbol{Y}\right]\right\}\right] \tag{2.96}$$

$$\le \mathbb{P}\left[\log M + \log\mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}},\boldsymbol{Y})}{q^n(\boldsymbol{X},\boldsymbol{Y})}\right)^s \,\middle|\, \boldsymbol{X},\boldsymbol{Y}\right] \ge \log\delta\right] + \delta, \tag{2.97}$$

where in (2.96) we also used $M - 1 \le M$, and (2.97) holds for any $\delta > 0$ by upper bounding $\min\{1, z\}$ by one when $z > \delta$, and by δ otherwise.

To simplify the notation, we write (2.97) as

$$\bar{p}_{\text{e}} \le \mathbb{P}[i_s^n(\boldsymbol{X},\boldsymbol{Y}) \le \log M - \log\delta] + \delta, \tag{2.98}$$

where

$$i_s^n(\boldsymbol{x},\boldsymbol{y}) = -\log\mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}},\boldsymbol{Y})}{q^n(\boldsymbol{X},\boldsymbol{Y})}\right)^s \,\middle|\, \boldsymbol{X},\boldsymbol{Y}\right]. \tag{2.99}$$

We proceed by showing that $i_s^n(\boldsymbol{X},\boldsymbol{Y})$ can be expressed as a sum of i.i.d. random variables with mean $I_{\text{GMI}}(Q_X)$. We have

$$i_s^n(\boldsymbol{x},\boldsymbol{y}) = -\log\sum_{\overline{\boldsymbol{x}}} Q_X^n(\overline{\boldsymbol{x}})\left(\frac{q^n(\overline{\boldsymbol{x}},\boldsymbol{y})}{q^n(\boldsymbol{x},\boldsymbol{y})}\right)^s \tag{2.100}$$

[6] We can also consider $s = 0$, as it makes the right-hand side of (2.96) equal one.

$$= -\log \sum_{\overline{\boldsymbol{x}}} \prod_{i=1}^{n} \left(Q_X(\overline{x}_i) \left(\frac{q(\overline{x}_i, y_i)}{q(x_i, y_i)} \right)^s \right) \tag{2.101}$$

$$= -\sum_{i=1}^{n} \log \sum_{\overline{x}} Q_X(\overline{x}) \left(\frac{q(\overline{x}, y_i)}{q(x_i, y_i)} \right)^s, \tag{2.102}$$

where (2.102) follows by writing $\sum_{\boldsymbol{x}} = \sum_{x_1} \dots \sum_{x_n}$ and then distributing the sums. The pair $(\boldsymbol{X}, \boldsymbol{Y})$ is i.i.d. on $Q_X \times W$, and averaging any summand of (2.102) over $(X_i, Y_i) \sim Q_X \times W$ yields the dual form of the GMI (*cf.*, (2.10)) upon optimizing s, as desired.

By the law of large numbers, we conclude from (2.98) that if the rate $R = \frac{1}{n} \log M$ is smaller than $I_{\mathrm{GMI}}(Q_X) - \delta$, then the upper bound (2.97) can be made arbitrarily close to zero for suitably-chosen δ.

2.6.5 Dual Derivation of the LM Rate

In the dual derivation of the GMI, the first two steps leading to (2.97) were not specific to any particular codeword distribution, so they remain valid here with the distribution $(\boldsymbol{X}, \boldsymbol{Y}, \overline{\boldsymbol{X}}) \sim P_{\boldsymbol{X}}(\boldsymbol{x}) W^n(\boldsymbol{y}|\boldsymbol{x}) P_{\boldsymbol{X}}(\overline{\boldsymbol{x}})$ adjusted accordingly. In addition, when $P_{\boldsymbol{X}}$ is the constant-composition codeword distribution, we can further weaken (2.97) as follows:

$$\bar{p}_{\mathrm{e}} \le \mathbb{P}\left[\log M + \log \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y})}{q^n(\boldsymbol{X}, \boldsymbol{Y})} \right)^s \frac{e^{a^n(\overline{\boldsymbol{x}})}}{e^{a^n(\boldsymbol{x})}} \,\middle|\, \boldsymbol{X}, \boldsymbol{Y} \right] \ge \log \delta \right] + \delta \tag{2.103}$$

$$\le \mathbb{P}\Bigg[\log M + |\mathcal{X}| \log(n+1) + \log \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}', \boldsymbol{Y})}{q^n(\boldsymbol{X}, \boldsymbol{Y})} \right)^s \frac{e^{a^n(\overline{\boldsymbol{x}})}}{e^{a^n(\boldsymbol{x})}} \,\middle|\, \boldsymbol{X}, \boldsymbol{Y} \right] \ge \log \delta \Bigg] + \delta, \tag{2.104}$$

where (2.103) holds with $a^n(\boldsymbol{x}) = \sum_{i=1}^n a(x_i)$ for any function $a(\cdot)$ since $\boldsymbol{X}$ and $\overline{\boldsymbol{X}}$ have the same composition, and (2.104) holds with $\overline{\boldsymbol{X}}' \sim Q_{X,n}^n$ by upper bounding $P_{\boldsymbol{X}}$ according to (2.83).

Now, similarly to (2.98), we can write (2.104) as

$$\bar{p}_{\mathrm{e}} \le \mathbb{P}[i_s^n(\boldsymbol{X}, \boldsymbol{Y}) \le \log M + |\mathcal{X}| \log n - \log \delta] + \delta, \tag{2.105}$$

where

$$i_{s,a}^n(\boldsymbol{x},\boldsymbol{y}) = -\log \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}',\boldsymbol{y})}{q^n(\boldsymbol{x},\boldsymbol{y})}\right)^s \frac{e^{a^n(\overline{\boldsymbol{X}}')}}{e^{a^n(\boldsymbol{x})}}\right]. \tag{2.106}$$

Since the expectation is taken with respect to an i.i.d. distribution for $\overline{\boldsymbol{X}}'$, we can expand it in the same way as (2.100)–(2.102) to obtain

$$i_{s,a}^n(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1}^n -\log \sum_{\overline{x}} Q_{X,n}(\overline{x})\left(\frac{q(\overline{x},y_i)}{q(x_i,y_i)}\right)^s \frac{e^{a(\overline{x})}}{e^{a(x_i)}}. \tag{2.107}$$

Observe that for any fixed $\boldsymbol{x} \in \mathcal{T}^n(Q_{X,n})$, if we replace $\boldsymbol{y}$ by $\boldsymbol{Y} \sim W^n(\cdot\,|\boldsymbol{x})$ in (2.107), then we get a sum of independent and *non-identically distributed* random variables, $nQ_{X,n}(x)$ of which are distributed as $-\log \sum_{\overline{x}} Q_{X,n}(\frac{q(\overline{x},Y)}{q(x,Y)})\frac{e^{a(\overline{x})}}{e^{a(x)}}$ with $Y \sim W(\cdot|x)$. The law of large numbers again applies in this case, and the corresponding normalized mean $\frac{1}{n}\mathbb{E}[i_{s,a}^n(\boldsymbol{X},\boldsymbol{Y})]$ is given by

$$\mu_n(s,a) \triangleq -\sum_{x,y} Q_{X,n}(x)W(y|x)\log \sum_{\overline{x}} Q_{X,n}(\overline{x})\left(\frac{q(\overline{x},y)}{q(x,y)}\right)^s \frac{e^{a(\overline{x})}}{e^{a(x)}}. \tag{2.108}$$

Using (2.105) and noting that $\frac{|\mathcal{X}|\log n - \log\delta}{n} \to 0$, we find that when $R \le \mu_n(s,a) - \delta$, the random-coding error probability vanishes. The achievability of $I_{\mathrm{LM}}(Q_X)$ now follows by recalling that $Q_{X,n}(x) \to Q_X(x)$ by definition, and taking the supremum over $s \ge 0$ and $a(\cdot)$.

2.6.6 Ensemble Tightness

The ensemble tightness of the GMI and LM rate (*cf.*, Lemma 2.9) are proved by following the primal achievability proofs, while replacing each upper bounding step by an analogous lower bound. We give the details for the LM rate (which turns out to be more difficult), and then only briefly comment on the GMI.

Since we have assumed that ties are broken as errors, the *exact* random-coding error probability in (2.80) can be written as

$$\bar{p}_{\mathrm{e}} = 1 - \mathbb{E}\left[\left(1 - \mathbb{P}[q^n(\overline{\boldsymbol{X}},\boldsymbol{Y}) \ge q^n(\boldsymbol{X},\boldsymbol{Y})\,|\,\boldsymbol{X},\boldsymbol{Y}]\right)^{M-1}\right] \tag{2.109}$$

with $(\boldsymbol{X}, \boldsymbol{Y}, \overline{\boldsymbol{X}}) \sim P_{\boldsymbol{X}}(\boldsymbol{x}) W^n(\boldsymbol{y}|\boldsymbol{x}) P_{\boldsymbol{X}}(\overline{\boldsymbol{x}})$, due to the independence of the codewords.

Since $(1 - \frac{1}{\alpha_n})^{\beta_n} \to 0$ whenever $\alpha_n \geq 1$ and $\lim_{n\to\infty} \frac{\alpha_n}{\beta_n} = 0$, we deduce from (2.109) that $\bar{p}_{\mathrm{e}} \to 1$ whenever the following holds for all $(\boldsymbol{x}, \boldsymbol{y})$ within some "typical set" (i.e., a set with probability tending to one):

$$\mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})] \geq e^{-n(R-\delta)}, \tag{2.110}$$

where $R = \frac{1}{n} \log M$, and $\delta > 0$ is arbitrarily small.

To establish (2.110), we analyze the left-hand side in terms of joint types, similarly to the primal achievability proof. Letting $P_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$ denote the joint type of $(\boldsymbol{x}, \boldsymbol{y})$, we have

$$\begin{aligned}
&\mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})] && (2.111)\\
&= \sum_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}):\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} \mathbb{P}[(\overline{\boldsymbol{X}}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{XY})] && (2.112)\\
&\geq \sum_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}): \widetilde{P}_X = P_X, \widetilde{P}_Y = P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} e^{-n(I_{\widetilde{P}}(X;Y)+\delta)} && (2.113)\\
&\geq \max_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}): \widetilde{P}_X = P_X, \widetilde{P}_Y = P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} e^{-n(I_{\widetilde{P}}(X;Y)+\delta)}, && (2.114)
\end{aligned}$$

where (2.113) follows from (2.85), and (2.114) follows by lower bounding the summation by the maximum.

We would now like to show that we can lower bound (2.114) in terms of a similar maximization over *all* joint distributions, not only joint types. Intuitively, the two should be essentially equivalent due to the fact that all joint distributions are increasingly close to the nearest joint type as $n \to \infty$. However, due to the constraints in (2.114), the proof is surprisingly tricky in general. We provide the formal statement as follows, and defer the proof to Appendix A.1.

Lemma 2.14 (Passing from Types to General Distributions)**.** For any decoding metric $q(x, y)$ and joint type $P_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$, it holds that

$$\begin{aligned} &\min_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}): \widetilde{P}_X = P_X, \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y) \\ &\quad \leq \min_{\substack{\widetilde{P}_{XY} \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}): \widetilde{P}_X = P_X, \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y) + \delta \end{aligned} \tag{2.115}$$

for arbitrarily small $\delta > 0$ and sufficiently large n.

By Lemma 2.14 and (2.114), we have

$$\mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})] \geq e^{-n(I_{\mathrm{LM}}(P_{XY})+2\delta)}, \tag{2.116}$$

where $I_{\mathrm{LM}}(P_{XY})$ denotes the LM rate (2.15) with input distribution P_X and channel $P_{Y|X}$. As a result, recalling the condition (2.110) and the fact that $\|P_{XY} - Q_X \times W\|_\infty \leq \delta$ with probability approaching one, we conclude that $\bar{p}_{\mathrm{e}} \to 1$ whenever

$$R \geq \max_{P_{XY}:\ \|P_{XY} - Q_X \times W\|_\infty \leq \delta} I_{\mathrm{LM}}(P_{XY}) + 3\delta. \tag{2.117}$$

Since the LM rate is continuous in P_{XY} by Lemma 2.5, and δ can be arbitrarily small, this establishes the ensemble tightness of $I_{\mathrm{LM}}(Q_X)$.

For the GMI, similar steps are applied, including the lower bound in (2.84). However, the analog of Lemma 2.14 turns out to be simpler to prove; see Appendix A.1 for details.

2.7 Multi-Letter Improvements

Since the GMI and the LM rate do not achieve the mismatch capacity in general (*cf.*, Lemma 2.11), it is natural to ask how we might achieve improved rates. The first proposed method for improvement, due to Csiszár and Narayan [26], was to apply the GMI and LM rate to the *product channel*, described by

$$W^2((y_1, y_2)|(x_1, x_2)) = W(y_1|x_1)W(y_2|x_2) \tag{2.118}$$

$$q^2((x_1, x_2), (y_1, y_2)) = q(x_1, y_1)q(x_2, y_2). \tag{2.119}$$

Indeed, it is easy to see that if the rate $R^{(2)}$ is achievable for (W^2, q^2), then the rate $R = \frac{1}{2}R^{(2)}$ is achievable for (W, q). This argument corresponds to coding over *pairs* of symbols, and perhaps surprisingly, this can lead to strict improvements in the mismatched setting.

More generally, we can code over k-tuples of symbols for any fixed integer k, and divide the resulting achievable rate for (W^k, q^k) by k to get the achievable rate for (W, q). By doing so, we obtain the following:

$$C_{\mathrm{GMI}}^{(k)} = \frac{1}{k} \max_{Q_{X^k}} I_{\mathrm{GMI}}(Q_{X^k}, W^k, q^k) \tag{2.120}$$

$$C_{\mathrm{LM}}^{(k)} = \frac{1}{k} \max_{Q_{X^k}} I_{\mathrm{LM}}(Q_{X^k}, W^k, q^k), \tag{2.121}$$

where on the right-hand sides we use the GMI and LM rate with an explicit dependence on the channel and metric. Formally, we have the following.

Lemma 2.15 (Multi-Letter Extensions)**.** For any mismatched DMC (W, q) and any positive integer k, the rates $C_{\mathrm{GMI}}^{(k)}$ and $C_{\mathrm{LM}}^{(k)}$ are achievable, and consequently, so are the rates $C_{\mathrm{GMI}}^{(\infty)} \triangleq \sup_{k\in\mathbb{Z}} C_{\mathrm{GMI}}^{(k)}$ and $C_{\mathrm{LM}}^{(\infty)} \triangleq \sup_{k\in\mathbb{Z}} C_{\mathrm{LM}}^{(k)}$. Moreover, there exist mismatched DMCs for which $C_{\mathrm{GMI}}^{(2)} > C_{\mathrm{GMI}}$ and $C_{\mathrm{LM}}^{(2)} > C_{\mathrm{LM}}$.

There are at least two cases where the strict improvement stated in the second part has been established:

- We will see that the zero-undetected error capacity example of Section 2.4.3 serves as a numerical example when we return to it in Section 6.3.3.
- An analytical proof of the second part is given in [26] based on the reduction to zero-error capacity (*cf.*, Section 1.2.8).

Csiszár and Narayan [26] conjectured that $C_{\mathrm{LM}}^{(k)}$ approaches the mismatch capacity as $k \to \infty$ (i.e., $C_{\mathrm{LM}}^{(\infty)} = C_{\mathrm{M}}$), but this conjecture remains open in general. We revisit this idea in detail in Section 8.

3

Continuous-Alphabet Memoryless Channels

3.1 Introduction

While discrete memoryless channels (DMCs) are convenient to analyze and capture a variety of interesting communication settings, there is substantial motivation to understand channels defined on continuous alphabets, such as the real or complex numbers.

In the absence of mismatch, coding theorems for continuous-alphabet channels are often obtained by performing an increasingly fine quantization of the inputs and outputs. However, such an approach is less suitable in the mismatched setting for at least two reasons. Firstly, quantizing the output amounts to changing the decoder, and it is non-trivial to study the resulting impact on the original mismatched decoder of interest. Secondly, it is often of interest to characterize the performance of random coding under a specific continuous input distribution (e.g., Gaussian).

An inspection of the dual i.i.d. analysis in Section 2.6, yielding the GMI, reveals that it directly extends to continuous-alphabet memoryless channels (see also Section 3.3 below). However, we saw that the constant-composition ensemble, defined on finite alphabets, yields the improved LM rate. One of the main goals of this section is to present a unified

random coding technique, known as *cost-constrained random coding*, that generalizes the LM rate to the continuous setting, as well as achieving it for DMCs (see Section 3.4).

In Section 3.8, we will apply the GMI and LM rate to three continuous-alphabet mismatched channels of interest: The additive white Gaussian noise (AWGN) channel with a mismatched signal level, non-Gaussian additive channels where an AWGN-type coding scheme is employed anyway, and a channel fading scenario with imperfect knowledge of the fading coefficients.

This section is predominantly based on the works of Ganti *et al.* [39] and Scarlett *et al.* [79], and includes examples based on the works of Merhav *et al.* [65], Lapidoth [53], and Lapidoth and Shamai [57].

3.2 Problem Setup

We continue to consider the setup described in Section 1.1, but we drop the assumption that $\mathcal{X}$ and $\mathcal{Y}$ are finite. For concreteness, we focus primarily on real alphabets, i.e., $\mathcal{X} = \mathcal{Y} = \mathbb{R}$; however, the analysis will extend directly to other alphabets such as $\mathbb{C}$ or $\mathbb{R}^d$, as well as finite or countably infinite discrete alphabets.

We let $W(y|x)$ be a conditional probability density function, and consider input distributions $Q_X(x)$ in the form of probability densities. As a result, the decoding metric $q(x, y)$ is also defined on $\mathbb{R} \times \mathbb{R}$. We initially focus on general choices of (W, q), but we will later specialize to the important class of additive noise channels, taking the form $Y = X + Z$ with Z representing random noise.

For most continuous-alphabet channels, the capacity is infinite unless suitable restrictions are placed on the input. To address this, we assume throughout the section that the transmitted codeword $\boldsymbol{x} = (x_1, \ldots, x_n)$ must satisfy an *input constraint* of the form

$$\frac{1}{n} \sum_{i=1}^{n} c(x_i) \leq \Gamma \tag{3.1}$$

for some cost function $c(x)$ and threshold Γ. Perhaps the most widely-adopted special case is a power constraint, in which $c(x) = x^2$, and Γ is the permitted per-symbol power averaged over the block. We denote

the minimal cost of a single symbols as

$$\Gamma_{\min} = \inf_{x \in \mathcal{X}} c(x), \tag{3.2}$$

and we assume throughout the section that $\Gamma > \Gamma_{\min}$.

Definition 3.1 (Input-Constrained Mismatch Capacity)**.** The input-constrained mismatch capacity $C_{\mathrm{M}}(\Gamma)$ of the pair (W, q) with input constraint (c, Γ) and $\Gamma > \Gamma_{\min}$ is defined to be the supremum of all achievable rates under the decoding rule (1.1) for codebooks $\mathcal{C}$ satisfying $\frac{1}{n}\sum_{i=1}^{n} c(x_i) \leq \Gamma$ for all $\boldsymbol{x} \in \mathcal{C}$.

Before proceeding, we present the following useful lemma [39].

Lemma 3.1 (Concavity in Γ)**.** Under the preceding setup, the function $C_{\mathrm{M}}(\Gamma)$ is a concave non-decreasing function of $\Gamma \in (\Gamma_{\min}, \infty)$.

Our use of this lemma will be in noting that concavity implies continuity, though the stronger concavity property is also of interest in its own right. The non-decreasing property is immediate from Definition 3.1, and concavity follows from a standard concatenation argument [39]: Choose the first λn symbols from a codebook achieving $C_{\mathrm{M}}(\Gamma_1)$, and the last $1 - \lambda n$ symbols from a codebook achieving $C_{\mathrm{M}}(\Gamma_2)$. The total cost is $\lambda\Gamma_1 + (1-\lambda)\Gamma_2$, and the rate $\lambda C_{\mathrm{M}}(\Gamma_1) + (1-\lambda)C_{\mathrm{M}}(\Gamma_2)$ is achieved. Recalling that we are considering decoding metrics of the form $q^n(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} q(x_i, y_i)$, we can factorize q^n as the product of the metrics for the first λn and last $(1-\lambda)n$ symbols. Hence, the product nature of the codebook implies that the overall mismatched decoder is equivalent to separately decoding the two blocks, and the overall error probability is upper bounded by the sum of the two blocks' error probabilities.

3.3 i.i.d. Random Coding and the GMI

Similarly to the discrete setting, we fix a density function Q_X and consider i.i.d. random coding of the form

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{n} Q_X(x_i). \tag{3.3}$$

In this case, the dual analysis of Section 2.6.4 applies without change upon replacing the summations by integrals, and we recover a continuous-alphabet counterpart of the GMI [48], stated in the following.

Theorem 3.2 (GMI for Continuous-Alphabet Channels)**.** For any mismatched memoryless channel (W, q), under i.i.d. random coding with input distribution Q_X, the random coding error probability vanishes for any fixed rate $R < I_{\text{GMI}}(Q_X)$, where

$$I_{\text{GMI}}(Q_X) = \sup_{s \geq 0} \mathbb{E}\left[\log \frac{q(X,Y)^s}{\mathbb{E}[q(\overline{X},Y)^s|Y]} \right], \tag{3.4}$$

with $(X, Y, \overline{X}) \sim Q_X(x)W(y|x)Q_X(\overline{x})$.

This result can also be inferred from a more general analysis given in Section 3.6 below.

Theorem 3.2 does not *immediately* provide a lower bound on the input-constrained mismatch capacity in the sense of Definition 3.1, as some codewords may violate the input constraint. However, we can form a feasible codebook using a standard expurgation argument similar to that of Remark 1.2: As long as $\mathbb{E}_Q[c(X)] < \Gamma$, any given codeword will indeed be feasible with probability approaching one (e.g., by the law of large numbers). Hence, for arbitrarily small $\epsilon > 0$, the average number of infeasible codewords is at most ϵM for sufficiently large n, and Markov's inequality implies that the probability of having $\frac{M}{2}$ or more infeasible codewords is at most 2ϵ. In the high-probability case that this event does not occur, the sub-codebook with only the feasible half (or more) of the codewords achieves vanishing error probability with rate approaching $I_{\text{GMI}}(Q_X)$. Hence, defining the optimized GMI as

$$C_{\text{GMI}}(\Gamma) \triangleq \sup_{Q_X : \mathbb{E}_Q[c(X)] \leq \Gamma} I_{\text{GMI}}(Q_X), \tag{3.5}$$

we have for any $\Gamma > \Gamma_{\min}$ that

$$C_{\text{GMI}}(\Gamma) \leq C_{\text{M}}(\Gamma). \tag{3.6}$$

The case $\mathbb{E}_Q[c(X)] = \Gamma$ is justified by the concavity (and hence continuity) of $C_{\text{M}}(\Gamma)$ in Γ, stated in Lemma 3.1.

3.4 Cost-Constrained Random Coding and the LM Rate

Based on the findings of Section 2, one would hope to be able to achieve a continuous variation of the LM rate (*cf.*, Theorem 2.3) in order to improve on the GMI. In this subsection, we present a codeword distribution constructed precisely for this purpose [39]. Specifically, we fix an input distribution Q_X and real-valued function $a(\cdot)$, and consider

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\Omega_n} \prod_{i=1}^{n} Q_X(x_i) \mathbb{1}\{\boldsymbol{x} \in \mathcal{D}_n\}, \tag{3.7}$$

where

$$\mathcal{D}_n \triangleq \left\{ \boldsymbol{x} \colon \left| \frac{1}{n} \sum_{i=1}^{n} c(x_i) - \phi_c \right| \leq \delta, \ \left| \frac{1}{n} \sum_{i=1}^{n} a(x_i) - \phi_a \right| \leq \delta \right\}, \tag{3.8}$$

$$\phi_c \triangleq \mathbb{E}_Q[c(X)], \quad \phi_a \triangleq \mathbb{E}_Q[a(X)], \tag{3.9}$$

and where $\delta > 0$ is a parameter, and Ω_n is a normalizing constant. Hence, $P_{\boldsymbol{X}}$ follows an i.i.d. distribution conditioned on both $c(\cdot)$ and $a(\cdot)$ having an empirical mean close to the true mean.

The intuition behind this choice of $P_{\boldsymbol{X}}$ is as follows:

- If $\mathbb{E}_Q[c(X)] < \Gamma$, where Γ is the threshold in (3.1), then the first constraint in (3.9) with sufficiently small δ ensures that all codewords satisfy the input constraint (3.1).
- The dual analysis in Section 2.6 for constant-composition random coding used the fact that $\sum_{i=1}^{n} a(x_i)$ is the same for all codewords and any $a(\cdot)$, in order to deduce the improved LM rate rather than the GMI. For $P_{\boldsymbol{X}}$ in (3.7), this property is *approximately* true for *one particular function* $a(\cdot)$. The deviation by δ does not impact the achievable rate in the limit $\delta \to 0$, and the fact that this property holds for only one function is also sufficient given that we are free to design that function to our liking.

We refer to the ensemble corresponding to the choice of $P_{\boldsymbol{X}}$ in (3.7) as *cost-constrained random coding*, but it is important to understand the distinction between $c(\cdot)$ and $a(\cdot)$: The *system cost* $c(\cdot)$ corresponds to a constraint imposed by the problem formulation, whereas the *auxiliary*

cost $a(\cdot)$ is intentionally introduced in order to improve the random coding performance.

While the choice of constraints in (3.9) will suffice for our purposes, other variations exist in the literature, such as one-sided system costs [38, Sec. 7.3] and choices of δ that decrease with n [79]. The choices of δ could also differ for the two constraints, but this would not affect the results that we present corresponding to $\delta \to 0$.

We formally state the achievability of a generalization of the LM rate in the following theorem [39], which will be proved in Section 3.6.

Theorem 3.3 (LM Rate for Continuous-Alphabet Channels)**.** For any mismatched memoryless channel (W, q) with input constraint (c, Γ), and any input distribution Q_X such that $\mathbb{E}_Q[c(X)] < \Gamma$, the following rate is achievable via cost-constrained random coding with suitably-chosen $a(\cdot)$ and $\delta > 0$:

$$I_{\text{LM}}(Q_X) = \sup_{s \geq 0, a(\cdot)} \mathbb{E}\left[\log \frac{q(X,Y)^s e^{a(X)}}{\mathbb{E}[q(\overline{X},Y)^s e^{a(\overline{X})}|Y]}\right], \tag{3.10}$$

where $(X, Y, \overline{X}) \sim Q_X(x)W(y|x)Q_X(\overline{x})$, and the supremum is over all $a(\cdot)$ such that $\mathbb{E}_Q[a(X)]$ is finite.

Analogously to (3.6), we deduce the following from Theorem 3.3: Defining the optimized LM rate

$$C_{\text{LM}}(\Gamma) \triangleq \sup_{Q_X \colon \mathbb{E}_Q[c(X)] \leq \Gamma} I_{\text{LM}}(Q_X) \tag{3.11}$$

we have for any $\Gamma > \Gamma_{\min}$ that

$$C_{\text{LM}}(\Gamma) \leq C_{\text{M}}(\Gamma). \tag{3.12}$$

Here we do not need to expurgate any codewords violating the power constraint, since we have ensured that there is zero probability of such an event happening.

Theorem 3.3 comes with an important caveat not present in the constant-composition random coding result of Theorem 2.3. Under constant-composition random coding, we can introduce an *arbitrary* function $a(\cdot)$ in the *analysis*, and then optimize the final result to deduce

the LM rate. In contrast, in the present setting, we need to choose a *specific* function $a(\cdot)$ in the random-coding *construction* itself. The optimal function $a(\cdot)$ in (3.10) clearly depends on both the channel W and the decoding metric q. As a result, the cost-constrained approach to achieving the LM rate requires knowing both of these. In fact, even when they are known, explicitly constructing an optimal choice of $a(\cdot)$ may be difficult, since the supremum in (3.10) corresponds to an infinite-dimensional optimization problem. In the following subsection, we present a more general random coding ensemble and analysis addressing these considerations.

A related fact is that while the decoding metrics $q(x, y)$ and $q(x, y)^s e^{a(x)} e^{b(y)}$ (with $s > 0$) always yield the same mismatch capacity when $\mathcal{X}$ is finite (*cf.*, Corollary 2.1), it is unclear whether this is true for infinite alphabets. The ability to include $a(x)$ in this claim for finite $\mathcal{X}$ is based on the fact that any good code also has a good constant-composition sub-code, and this fact has no obvious extension to the continuous-alphabet case. However, from (3.10), one can at least deduce that the metrics $q(x, y)$ and $q(x, y)^s e^{a(x)} e^{b(y)}$ yield the same LM rate.

3.5 Cost-Constrained Random Coding with Multiple Auxiliary Costs

We first consider the question of how cost-constrained random coding performs with a suboptimal choice of the auxiliary cost $a(\cdot)$ (again considering the limit $\delta \to 0$). A natural guess is that we can achieve the rate (3.10) with the supremum over $a(\cdot)$ replaced by the fixed choice. Such a rate is indeed achievable, but it turns out that a better rate can be derived by introducing an additional auxiliary parameter, allowing us to replace $a(\cdot)$ by $ra(\cdot)$ for any $r \in \mathbb{R}$ (see Theorem 3.4 below).

We know that in order to achieve the LM rate, having a single optimized auxiliary cost is enough. However, when such optimization is not possible, it is also of interest to understand the behavior in the presence of *multiple* auxiliary costs [78, 79]. Generalizing (3.7), we fix $Q_X \in \mathcal{P}(\mathcal{X})$ and let L denote the number of auxiliary costs, denote the l-th one by $a_l(\cdot)$ and its mean by $\phi_l = \mathbb{E}_Q[a_l(X)]$, and consider the

codeword distribution

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\Omega_n} \prod_{i=1}^{n} Q_X(x_i) \mathbb{1}\{\boldsymbol{x} \in \mathcal{D}_n\}, \tag{3.13}$$

where

$$\mathcal{D}_n \triangleq \Bigg\{ \boldsymbol{x} \colon \left| \frac{1}{n} \sum_{i=1}^{n} c(x_i) - \phi_c \right| \leq \delta, \\ \left| \frac{1}{n} \sum_{i=1}^{n} a_l(x_i) - \phi_l \right| \leq \delta, \forall l = 1, \ldots, L \Bigg\}. \tag{3.14}$$

The following theorem states an achievable rate for random coding under this more general ensemble.

Theorem 3.4 (Achievable Rate with Fixed Auxiliary Costs)**.** For any mismatched memoryless channel (W, q) with input constraint (c, Γ), and any input distribution Q_X such that $\mathbb{E}_Q[c(X)] < \Gamma$, the following rate is achievable via cost-constrained random coding with fixed auxiliary costs $a_1(\cdot), \ldots, a_L(\cdot)$ having finite means $\phi_l = \mathbb{E}_Q[a_l(X)]$:

$$I'_{\text{LM}}(Q_X, \{a_l\}_{l=1}^{L}) = \sup_{s \geq 0, \{r_l\}_{l=1}^{L}} \mathbb{E}\left[\log \frac{q(X,Y)^s e^{\sum_{l=1}^{L} r_l a_l(X)}}{\mathbb{E}[q(\overline{X}, Y)^s e^{\sum_{l=1}^{L} r_l a_l(\overline{X})} | Y]} \right], \tag{3.15}$$

where $(X, Y, \overline{X}) \sim Q_X(x) W(y|x) Q_X(\overline{x})$.

The intuition behind the presence of the parameters $\{r_l\}_{l=1}^{L}$ is that since $a_l(\cdot)$ has an empirical mean within δ of the true mean, $r_l a_l(\cdot)$ has an empirical mean within $r_l \delta$ of the true mean. Hence, by letting δ be sufficiently small, we can perform the analysis as if the l-th auxiliary cost function were scaled by r_l.

In fact, we could also include $rc(\cdot)$ (with $r \in \mathbb{R}$) inside both of the exponential terms in (3.15) to potentially improve the rate further. We omit such terms to avoid cumbersome notation, and since the resulting rate could equivalently be recovered by including an additional auxiliary cost with $a_l(\cdot) = c(\cdot)$.

We note the following relation between the achievable rates:

$$I_{\text{LM}}(Q_X) \geq I'_{\text{LM}}(Q_X, \{a_l\}_{l=1}^{L}) \geq I_{\text{GMI}}(Q_X), \tag{3.16}$$

where the right inequality follows by setting all values of r_l to zero in (3.15), and the left inequality follows since given any choices of auxiliary costs and parameters in (3.15), we recover the objective in (3.10) upon setting $a(x) = \sum_{l=1}^{L} r_l a_l(x)$. A generalization of the right inequality in (3.16) is that adding further auxiliary costs can never reduce the achievable rate.

The preceding results and discussions apply to both discrete and continuous channels. In the discrete case, it is also worth noting that if one chooses $L = |\mathcal{X}|$ and $a_l(x) = \mathbb{1}\{x = l\}$ (assuming without loss of generality that $\mathcal{X} = \{1, \ldots, |\mathcal{X}|\}$), then the distribution in (3.13) produces codewords whose types are within δ of Q_X in the entry-wise sense. Hence, this choice recovers a similar ensemble to the constant-composition ensemble. However, it is not necessary to increase the number of auxiliary costs as a function of the alphabet size to attain the LM rate, as Theorem 3.3 shows that a single auxiliary cost is sufficient.

3.6 Proofs of Achievable Rates

Here we present the proof of the fixed-cost LM rate in Theorem 3.4, which immediately implies the GMI (Theorem 3.2) upon setting $L = 0$, and the LM rate (Theorem 3.3) upon setting $L = 1$ and optimizing $a_1(\cdot)$. The analysis is similar to the dual constant-composition analysis of Section 2.6, but we provide the details for completeness.

We start with the following key property of the cost-constrained ensemble, which follows directly from the law of large numbers.

Lemma 3.5 (Normalizing Constant for Cost-Constrained Coding)**.** For the cost-constrained codeword distribution in (3.13) with L fixed auxiliary costs, if ϕ_c and $\{\phi_l\}_{l=1}^{L}$ are finite, then $\Omega_n \to 1$ as $n \to \infty$ for any fixed $\delta > 0$.

We now consider the bound (2.97), which holds for a general codeword distribution, and is repeated here for convenience:

$$\bar{p}_{\mathrm{e}} \leq \mathbb{P}\left[\log M + \log \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y})}{q^n(\boldsymbol{X}, \boldsymbol{Y})}\right)^s \,\middle|\, \boldsymbol{X}, \boldsymbol{Y}\right] \geq \log \delta\right] + \delta, \tag{3.17}$$

where $(\boldsymbol{X}, \boldsymbol{Y}, \overline{\boldsymbol{X}}) \sim P_{\boldsymbol{X}}(\boldsymbol{x}) W^n(\boldsymbol{y}|\boldsymbol{x}) P_{\boldsymbol{X}}(\overline{\boldsymbol{x}})$, and $\delta > 0$ is arbitrary. To form a convenient upper bound, we define $a_l^n(\boldsymbol{x}) \triangleq \sum_{i=1}^n a_l(x_i)$, and observe that the constraint set in (3.14) implies that

$$1 \le \frac{e^{\sum_{l=1}^L (r_l a_l^n(\overline{\boldsymbol{x}}) + |r_l|\delta n)}}{e^{\sum_{l=1}^L (r_l a_l^n(\boldsymbol{x}) - |r_l|\delta n)}} = \frac{e^{\sum_{l=1}^L r_l a_l^n(\overline{\boldsymbol{x}})}}{e^{\sum_{l=1}^L r_l a_l^n(\boldsymbol{x})}} \cdot e^{2\delta n \sum_{l=1}^L r_l} \tag{3.18}$$

for any parameters $\{r_l\}_{l=1}^L$, and any codewords $\boldsymbol{x}$ and $\overline{\boldsymbol{x}}$ such that $P_{\boldsymbol{X}}(\boldsymbol{x}) > 0$. Hence, defining $r_{\text{sum}} = \sum_{l=1}^L |r_l|$ for brevity, we obtain

$$\log \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}, \boldsymbol{y})}{q^n(\boldsymbol{x}, \boldsymbol{y})}\right)^s\right] \le \log \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}, \boldsymbol{y})}{q^n(\boldsymbol{x}, \boldsymbol{y})}\right)^s \frac{e^{\sum_{l=1}^L r_l a_l^n(\overline{\boldsymbol{X}})}}{e^{\sum_{l=1}^L r_l a_l^n(\boldsymbol{x})}} e^{2\delta r_{\text{sum}} n}\right]. \tag{3.19}$$

Substituting into (3.17) and using the fact that $P_{\boldsymbol{X}}(\boldsymbol{x}) \le \frac{1}{\Omega_n} Q_X^n(\boldsymbol{x})$ (see (3.13)), we obtain

$$\begin{aligned} \bar{p}_{\text{e}} \le \frac{1}{\Omega_n} \mathbb{P}\Bigg[&- \log \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}', \boldsymbol{Y})}{q^n(\boldsymbol{X}', \boldsymbol{Y})}\right)^s \frac{e^{\sum_{l=1}^L r_l a_l^n(\overline{\boldsymbol{X}}')}}{e^{\sum_{l=1}^L r_l a_l^n(\boldsymbol{X}')}} \,\middle|\, \boldsymbol{X}', \boldsymbol{Y}\right] \\ &\le \log \frac{M}{\delta \Omega_n} + 2\delta r_{\text{sum}} n \Bigg] + \delta, \end{aligned} \tag{3.20}$$

where $(\boldsymbol{X}', \boldsymbol{Y}, \overline{\boldsymbol{X}}') \sim Q_X^n(\boldsymbol{x}) W^n(\boldsymbol{y}|\boldsymbol{x}) Q_X^n(\overline{\boldsymbol{x}})$ is now a triplet of random vectors with the corresponding $(X_i', Y_i', \overline{X}_i)$ being i.i.d. with respect to $i = 1, \ldots, n$. For notational convenience, we rewrite (3.20) as

$$\bar{p}_{\text{e}} \le \frac{1}{\Omega_n} \mathbb{P}\Big[i_{s,\{r_l\}}(\boldsymbol{X}', \boldsymbol{Y}) \le \log M - \log(\delta \Omega_n) + 2\delta r_{\text{sum}} n \Big] + \delta, \tag{3.21}$$

where

$$i_{s,\{r_l\}}(\boldsymbol{x}, \boldsymbol{y}) \triangleq - \log \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}', \boldsymbol{y})}{q^n(\boldsymbol{x}, \boldsymbol{y})}\right)^s \frac{e^{\sum_{l=1}^L r_l a_l^n(\overline{\boldsymbol{X}}')}}{e^{\sum_{l=1}^L r_l a_l^n(\boldsymbol{x})}}\right]. \tag{3.22}$$

Since the average in (3.22) is taken with respect to $\overline{\boldsymbol{X}}' \sim Q_X^n$, we can expand it in the same way as (2.100)–(2.102) to obtain

$$i_{s,\{r_l\}}(\boldsymbol{x}, \boldsymbol{y}) = - \sum_{i=1}^n \log \sum_{\overline{x}} Q_X(\overline{x}) \left(\frac{q(\overline{x}, y_i)}{q(x_i, y_i)}\right)^s \frac{e^{\sum_{l=1}^L r_l a_l(\overline{x})}}{e^{\sum_{l=1}^L r_l a_l(x_i)}}. \tag{3.23}$$

Hence, when we substitute the pair $(\boldsymbol{X}', \boldsymbol{Y})$, we have a sum of i.i.d. random variables, each having mean

$$\mu(s, \{r_l\}_{l=1}^L)$$
$$= -\sum_{x,y} Q_X(x)W(y|x) \log \sum_{\overline{x}} Q_X(\overline{x}) \left(\frac{q(\overline{x}, y)}{q(x, y)}\right)^s \frac{e^{\sum_{l=1}^L r_l a_l(\overline{x})}}{e^{\sum_{l=1}^L r_l a_l(x)}}. \quad (3.24)$$

By applying the law of large numbers in (3.20) and noting that $\Omega_n \to 1$ (see Proposition 3.5) and δ is arbitrarily small, we conclude that $\mu(s, \{r_l\}_{l=1}^L)$ is an achievable rate. Optimizing over $s \geq 0$ and $\{r_l\}_{l=1}^L$ establishes Theorem 3.4.

3.7 Ensemble Tightness Results

We know from Section 2 that in the finite-alphabet setting, the GMI and LM rate are ensemble-tight for the i.i.d. and constant-composition ensembles. In addition, when the input constraint is absent or trivial (e.g., $c(x) = 0$ for all x, and $\Gamma = 0$), or when one of the auxiliary costs $a_l(\cdot)$ is equal to $c(\cdot)$ (which can be assumed without loss of generality), it can also be shown using the techniques of Section 2 that $I'_{\mathrm{LM}}(Q_X, \{a_l\}_{l=1}^L)$ in Theorem 3.4 is ensemble-tight for the cost-constrained ensemble when specialized to any mismatched DMC [79].

In this subsection, we present an ensemble tightness result for the GMI with continuous alphabets, which holds under mild technical assumptions, and follows from the large-deviations analysis of a distinct but related rate-distortion setup [28]. Related proofs of ensemble tightness for certain additive noise channels can also be found in [53, 57, 105], and an example from [53] with a simpler direct argument will be illustrated in the proof of Theorem 3.9 below.

We will also discuss a *plausibility argument* for the ensemble tightness of the LM rate under cost-constrained random coding in the continuous-alphabet setting, but we leave open the problem of establishing this rigorously.

The large deviations result that we utilize [28, Thm. 1] is stated as follows without proof. We note that the roles of X and Y are reversed here compared to [28], and the negative log-metric $-\log q(x, y)$ plays

the role of the distortion function in [28]. This result will be re-stated in Section 4.7 (see Lemma 4.6 therein) in the context of rate-distortion theory, using notation that more closely resembles that of [28].

Lemma 3.6 (Large Deviations Result for i.i.d. Random Coding)**.** Fix an input distribution Q_X, an output distribution P_Y, and a decoding metric q satisfying $q(x,y) \in (0,1]$ for all (x,y), and define[1]

$$\gamma_{\min} = \mathbb{E}_{P_Y}[\operatorname{ess\,inf}_{Q_X} \{-\log q(X,Y)\}], \tag{3.25}$$

$$\gamma_{\text{prod}} = \mathbb{E}_{Q_X \times P_Y}[-\log q(X,Y)]. \tag{3.26}$$

Then, if $\gamma_{\text{prod}} < \infty$, we have for any $\gamma \in (\gamma_{\min}, \gamma_{\text{prod}})$ that the following holds with probability one with respect to $\boldsymbol{Y}$:

$$\begin{aligned} &\lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}[-\log q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \geq n\gamma | \boldsymbol{Y}] \\ &\quad = \sup_{s \geq 0} \big\{ -s\gamma - \mathbb{E}[\log \mathbb{E}[q(\overline{X}, Y)^s | Y]] \big\}, \end{aligned} \tag{3.27}$$

where $(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \sim Q_X^n(\overline{\boldsymbol{x}}) P_Y^n(\boldsymbol{y})$ and $(\overline{X}, Y) \sim Q_X(\overline{x}) P_Y(y)$.

Using this result, we can establish the ensemble tightness of the GMI for continuous-alphabet memoryless channels, stated in the following theorem. We adopt some technical assumptions that follow from those regarding q, $\gamma_{\min}$, and γ_{prod} in Lemma 3.6, and will be discussed shortly.

Theorem 3.7 (Ensemble Tightness of the GMI)**.** Consider a mismatched memoryless channel (W, q) and an input distribution Q_X, and let P_Y be the marginal distribution of $P_{XY} = Q_X \times W$. Then, under the assumptions

$$q(x,y) \in (0, q_{\max}], \ \forall (x,y), \quad \text{for some } q_{\max} < \infty, \tag{3.28}$$

$$\mathbb{E}_{Q_X \times P_Y}[\log q(X,Y)] > -\infty, \tag{3.29}$$

$$\mathbb{E}_{Q_X \times W}[\log q(X,Y)] < \mathbb{E}_{P_Y}[\operatorname{ess\,sup}_{Q_X} \{\log q(X,Y)\}], \tag{3.30}$$

$$\mathbb{E}_{Q_X \times W}[\log q(X,Y)] > \mathbb{E}_{Q_X \times P_Y}[\log q(X,Y)], \tag{3.31}$$

we have for any $R > I_{\text{GMI}}(Q_X)$ that $\bar{p}_{\text{e}}(n, \lfloor e^{nR} \rfloor) \to 1$ as $n \to \infty$ under i.i.d. random coding with input distribution Q_X.

[1] The *essential infimum* of a function $g(X)$ with respect to $X \sim Q_X$ is defined to be supremum of $t \in \mathbb{R}$ for which $\mathbb{P}_Q[g(X) > t] = 1$. The essential supremum is defined analogously.

Before presenting the proof, we briefly discuss the assumptions. We observe from (3.28) that $q(x,y) = 0$ is disallowed, but we are not aware of any continuous-alphabet examples that consider decoding metrics taking value zero. Similarly, the upper bound $q(x,y) \leq q_{\max}$ states that the metric is bounded, and we are not aware of any counter-examples considered previously. In particular, for standard distance-based decoding metrics such as $q(x,y) = e^{-(y-x)^2}$ (see Section 3.8), (3.28) is satisfied with $q_{\max} = 1$.

The assumption (3.29) rules out scenarios in which incorrect codewords yield a log-decoding metric with a very heavy tail in the negative direction; such situations are not commonly encountered. Similarly, (3.30) rules out pathological scenarios where the transmitted codeword always has the highest metric possible. This can occur naturally for certain mismatched DMCs (e.g., see Section 1.2.7), but we are not aware of any such examples in continuous-alphabet settings.

Finally, we note that the condition (3.31) was already considered in Lemma 2.6, where it was shown that, at least for DMCs, the mismatch capacity is zero when this fails to hold. While we are not aware of an analogous claim for continuous-alphabet channels, it is shown in [39, Prop. 1] that the LM rate (and therefore, the GMI) is indeed zero when (3.31) fails.

Proof of Theorem 3.7. We first argue that we can consider the case that $q_{\max} = 1$ in (3.28) without loss of generality. Indeed, if $q_{\max} > 1$, then an equivalent decoding metric is attained by dividing all values by $q_{\max}$, and the resulting normalized metric is upper bounded by one.

We start with the exact expression for the error probability in (2.109):

$$\bar{p}_{\mathrm{e}} = 1 - \mathbb{E}\left[\left(1 - \mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y}) \,|\, \boldsymbol{X}, \boldsymbol{Y}]\right)^{M-1}\right] \tag{3.32}$$

$$= 1 - \mathbb{E}\left[\left(1 - \mathbb{P}[-\log q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \leq -\log q^n(\boldsymbol{X}, \boldsymbol{Y}) \,|\, \boldsymbol{X}, \boldsymbol{Y}]\right)^{M-1}\right], \tag{3.33}$$

where $(\boldsymbol{X}, \boldsymbol{Y}, \overline{\boldsymbol{X}}) \sim Q_X^n(\boldsymbol{x}) W^n(\boldsymbol{y}|\boldsymbol{x}) Q_X^n(\overline{\boldsymbol{x}})$, and $M = \lfloor e^{nR} \rfloor$. Observe that $-\log q^n(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=1}^n -\log q(X_i, Y_i)$ is an i.i.d. summation with

mean $n\mathbb{E}[-\log q(X,Y)]$, where $(X,Y) \sim Q_X \times W$. Hence, by the law of large numbers, we have for any $\delta > 0$ and sufficiently large n that the following holds with $\gamma = \mathbb{E}[-\log q(X,Y)] - \delta$:

$$\bar{p}_{\mathrm{e}} \geq 1 - \delta - \mathbb{E}\left[\left(1 - \mathbb{P}[-\log q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \leq n\gamma | \boldsymbol{Y}]\right)^{M-1}\right]. \tag{3.34}$$

This inner probability matches that in (3.27), and we proceed by verifying the technical conditions of Lemma 3.6.

Multiplying both sides in (3.29)–(3.31) by -1 gives $\gamma_{\min} < \mathbb{E}[-\log q(X,Y)] < \gamma_{\mathrm{prod}} < \infty$, where $\gamma_{\min}$ and γ_{prod} are defined in (3.25)–(3.26). Hence, under the choice $\gamma = \mathbb{E}[-\log q(X,Y)] - \delta$, we have for sufficiently small δ that $\gamma \in (\gamma_{\min}, \gamma_{\mathrm{prod}})$, and Lemma 3.6 gives

$$\begin{aligned} &\lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}[-\log q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \geq n\gamma | \boldsymbol{Y}] \\ &\qquad = \sup_{s \geq 0}(-s\gamma - \mathbb{E}[\log \mathbb{E}[q(\overline{X}, Y)^s | Y]]) \end{aligned} \tag{3.35}$$

almost surely with respect to $\boldsymbol{Y}$.

Observe that when δ is replaced by zero in the choice of γ, the right-hand side of (3.35) simplifies as follows: Defining $(X, Y, \overline{X}) \sim Q_X(x)W(y|x)Q_X(\overline{x})$, we have

$$\begin{aligned} &\sup_{s \geq 0}(-s\mathbb{E}[-\log q(X,Y)] - \mathbb{E}[\log \mathbb{E}[q(\overline{X}, Y)^s | Y]]) \\ &\qquad = \sup_{s \geq 0} \mathbb{E}\left[\log \frac{q(X,Y)^s}{\mathbb{E}[q(\overline{X}, Y)^s | Y]}\right] &(3.36) \\ &\qquad = I_{\mathrm{GMI}}(Q_X). &(3.37) \end{aligned}$$

Thus, assuming momentarily that we can replace δ by zero in (3.34), we deduce that $\bar{p}_{\mathrm{e}} \to 1$ for $R > I_{\mathrm{GMI}}(Q_X)$, using $M = \lfloor e^{nR} \rfloor$ and the fact that $(1 - \frac{1}{\alpha_n})^{\beta_n} \to 0$ whenever $\lim_{n\to\infty} \frac{\alpha_n}{\beta_n} = 0$.

To justify taking $\delta \to 0$, it suffices to show that the right-hand side of (3.27) is continuous with respect to $\gamma \in (\gamma_{\min}, \gamma_{\mathrm{prod}})$. This follows from the fact that the pointwise supremum of a collection of linear functions is convex [18, Sec. 3.2.3], and a convex function is continuous when restricted to the region in which it is finite. The assumptions of Lemma 3.6 ensure that (3.27) is finite for $\gamma \in (\gamma_{\min}, \gamma_{\mathrm{prod}})$ (see the discussion following [28, Thm. 1]), which implies the desired claim. □

Discussion. The key tool in proving Lemma 3.6 is the Gärtner-Ellis theorem in large deviations theory; the interested reader is referred to [28] for the details. Intuitively, the parameter s in the dual expression for the GMI can be viewed as a Chernoff parameter resulting from applying the Chernoff bound to $\sum_{i=1}^{n} \log q(\overline{X}_i, y_i)$.

Similarly, in the fixed-cost LM rate of Theorem 3.4 each parameter r_l can be viewed as combining two Chernoff parameters: one for $\sum_{i=1}^{n} a(\overline{X}_i) \geq n(\phi_a - \delta)$ and one for $\sum_{i=1}^{n} a(\overline{X}_i) \leq n(\phi_a + \delta)$. By taking the difference of two arbitrary positive parameters, we end up with an arbitrary real-valued parameter. In light of this interpretation, we expect that counterparts of Lemma 3.6 and Theorem 3.7 should hold for the LM rate and cost-constrained random coding, but currently this remains an open problem.

3.8 Examples

In this subsection, we evaluate the preceding achievable rates for three continuous-alphabet channels. For the sake of analytical tractability, all examples are based on a Gaussian input distribution and a form of nearest-neighbor decoding; however, the three are still fundamentally distinct due to the different forms of the mismatch.

3.8.1 Mismatched Signal Level

We consider a variant of the additive white Gaussian noise (AWGN) channel in which the input is scaled by an unknown constant $\alpha > 0$:

$$Y = \alpha X + Z, \tag{3.38}$$

where $Z \sim \mathrm{N}(0, \sigma^2)$ for some noise power $\sigma^2 > 0$, and X and Z are independent. We consider a power constraint corresponding to (3.1) with $c(x) = x^2$ and some $\Gamma > 0$. Accordingly, we adopt a Gaussian input distribution Q_X corresponding to $X \sim \mathrm{N}(0, \Gamma)$, which is capacity-achieving in the matched case.

The optimal decoding rule minimizes $\|\boldsymbol{y} - \alpha \boldsymbol{x}\|_2^2$, and requires knowledge of α. We consider a scenario in the decoder is unaware of the

scaling effect, and therefore acts as if $\alpha = 1$:[2]

$$\hat{m} = \operatorname*{arg\,min}_{j=1,\ldots,M} \|\boldsymbol{y} - \boldsymbol{x}^{(j)}\|_2^2. \tag{3.39}$$

This is the *nearest-neighbor* decoding rule, and corresponds to the decoding metric $q(x, y) = e^{-(y-x)^2}$.

In this subsection, we will derive the following expressions for the GMI and LM rate [65].

Theorem 3.8 (Mismatched Signal Level)**.** Under the preceding mismatched signal level setting with input power $\Gamma > 0$, noise power $\sigma^2 > 0$, and signal level $\alpha > 0$, we have the following:

(i) The GMI is given by

$$I_{\mathrm{GMI}}(Q_X) = \sup_{s \geq 0} \frac{1}{2} \log(1 + 2s\Gamma) + \frac{s(\alpha^2\Gamma + \sigma^2)}{1 + 2\Gamma s} - s((\alpha - 1)^2\Gamma + \sigma^2). \tag{3.40}$$

(ii) The LM rate is given by

$$I_{\mathrm{LM}}(Q_X) = \frac{1}{2} \log\left(1 + \frac{\alpha^2\Gamma}{\sigma^2}\right), \tag{3.41}$$

and equals the matched (and hence mismatched) capacity subject to a power constraint Γ.

The intuition behind the fact that we achieve the matched capacity in the second part is that the mismatched and maximum-likelihood decoding rules are equivalent for codebooks in which all codewords have the same power $\|\boldsymbol{x}\|_2^2$. By using cost-constrained random coding with an auxiliary cost $a(x) = x^2$, we can ensure that the powers are *nearly* identical, and that we still achieve the optimal rate.

We compare the GMI and LM rate numerically in Figure 3.1. In addition, we plot the achievable rate of Theorem 3.4 with a single auxiliary cost (i.e., $L = 1$) under the suboptimal choices $a_1(x) = |x|$ and

[2] More generally, one could consider a decoder that has an estimate $\hat{\alpha}$ of α and decodes accordingly. Taking $\hat{\alpha} = 1$ in fact entails no loss of generality, since one can reduce to this scenario by re-scaling (and accordingly modifying the noise level σ^2 and input power Γ).

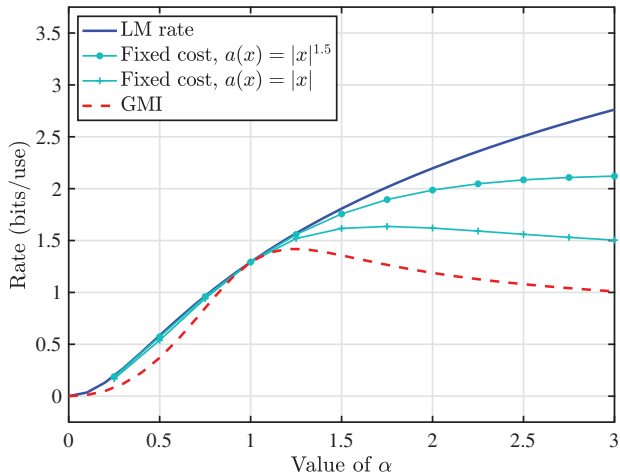


Figure 3.1: Unknown signal level example: GMI, LM rate, and its variation with a suboptimal auxiliary cost, each plotted as a function of α.

$a_1(x) = |x|^{1.5}$. These rates were evaluated using numerical integration, along with gradient ascent for optimizing (s, r_1).

We observe that the LM rate exceeds the GMI for all $\alpha > 0$, except for the matched case $\alpha = 1$. Moreover, the GMI is seen to saturate as $\alpha \to \infty$, while the LM rate grows unbounded. Moreover, even with a suboptimal auxiliary cost, we get a strict improvement over the GMI, and nearly match the LM rate for $\alpha < 1$. The choice $a_1(x) = |x|^{1.5}$ provides a better rate than $a_1(x) = |x|$, as should be expected due to the fact that the former more closely resembles the optimal choice $a_1(x) = x^2$.

Proof of Theorem 3.8

We first consider the LM rate, and then the GMI.

Evaluation of the LM Rate. The right-hand side of (3.41) is the matched capacity of the AWGN channel with input scaled by α. To see that it is achieved by the LM rate, we write the mismatched metric q

and the optimal (maximum-likelihood) metric q^* as

$$\log q(x,y) = -y^2 - x^2 + 2xy \tag{3.42}$$

$$\log q^*(x,y) = -y^2 - \alpha^2 x^2 + 2\alpha xy, \tag{3.43}$$

which implies that $\log q^*(x,y) = \alpha \log q(x,y) + (\alpha - \alpha^2)x^2 + (\alpha - 1)y^2$. By Lemma 2.7 (whose proof extends to the continuous setting), we conclude that $I_{\mathrm{LM}}(Q_X)$ equals the matched decoding rate (i.e., the mutual information $I(X;Y)$), which in turn is known to be given by (3.41).

Evaluation of the GMI. We write the GMI in (3.4) as

$$I_{\mathrm{GMI}}(Q_X) = \sup_{s \geq 0} \; s\mathbb{E}[\log q(X,Y)] - \mathbb{E}[\log \mathbb{E}[q(\overline{X},Y)^s|Y]] \tag{3.44}$$

$$= \sup_{s \geq 0} \; -s\mathbb{E}[(Y-X)^2] - \mathbb{E}[\log \mathbb{E}[e^{-s(Y-\overline{X})^2}|Y]], \tag{3.45}$$

where we have substituted $q(x,y) = e^{-(y-x)^2}$. We then obtain (3.40) by evaluating the relevant Gaussian integrals according to $Y = aX + Z$ with $X \sim \mathrm{N}(0,\Gamma)$ and $Z \sim \mathrm{N}(0,\sigma^2)$. The technical details of these integral evaluations are omitted.

3.8.2 Mismatched Noise Distribution – Nearest-Neighbor Decoding for Additive Non-Gaussian Channels

Continuing with the theme of additive noise channels, we turn to a particularly important setting in which the noise distribution is mismatched. The input-output relationship is given by

$$Y = X + Z, \tag{3.46}$$

where X and Z are independent. We again consider an input constraint according to (3.1) with $c(x) = x^2$ (i.e., a power constraint of $\Gamma > 0$), and we assume that Z follows an *arbitrary distribution* having a finite mean and variance:

$$\mathbb{E}[Z] = \mu, \quad \mathrm{Var}[Z] = \sigma^2 \tag{3.47}$$

for some $\mu \in \mathbb{R}$ and $\sigma^2 > 0$.

AWGN Case. As we already saw in the previous example, when the noise is Gaussian, the matched capacity is given by

$$C_{\text{AWGN}}(\Gamma, \sigma^2) = \frac{1}{2} \log \left(1 + \frac{\Gamma}{\sigma^2}\right). \tag{3.48}$$

Moreover, when $\mu = 0$, the optimal decoding rule is nearest-neighbor decoding according to (3.39), which we repeat here:

$$\hat{m} = \arg\min_{j=1,\ldots,M} \|\boldsymbol{y} - \boldsymbol{x}^{(j)}\|_2^2. \tag{3.49}$$

More generally, replacing $\boldsymbol{y}$ by $\boldsymbol{y} - \mu\mathbf{1}$ in (3.49) gives the optimal rule, where $\mathbf{1}$ is the vector of ones. The capacity-achieving input distribution under the power constraint $\mathbb{E}[X^2] \leq \Gamma$ is given by $X \sim \mathrm{N}(0, \Gamma)$.

Non-Gaussian Noise. Gaussian noise is convenient to analyze mathematically, is often well-motivated in applications, and has been studied extensively from both theoretical and practical viewpoints. On the other hand, it is certainly not the only type of additive noise of interest.

In general, under non-Gaussian noise, the *matched* channel capacity is given by $C = \sup_{Q_X : \mathbb{E}_Q[X^2] \leq \Gamma} I(X;Y)$. However, achieving this may be difficult for several reasons: The distribution P_Z may not be known, and even if it is known, the required codebook structure and decoder may be prohibitively complex.

Motivated by these considerations, we consider the following question [53]: *If we adopt a Gaussian coding scheme, but apply it to a channel that actually consists of non-Gaussian additive noise, how well will it perform?* More precisely, we consider random coding with input distribution $Q_X \sim \mathrm{N}(0, \Gamma)$, along with the nearest-neighbor decoding rule (3.49) corresponding to $q(x, y) = e^{-(y-x)^2}$. These choices are optimal for the *zero-mean* AWGN channel; we will also comment on the alternative choice $q(x, y) = e^{-(y-\mu-x)^2}$ corresponding to optimal decoding for $\mathrm{N}(\mu, \sigma^2)$ noise.

If we apply such a Gaussian coding scheme to a non-Gaussian channel, we may expect a performance degradation due to the fact that we have designed a code for the wrong channel. On the other hand, we may expect a higher rate due to the fact that for a given

variance $\mathrm{Var}[Z] = \sigma^2$, Gaussian noise is the most harmful in the matched scenario (i.e., yields the smallest capacity) [22, Ex. 9.21]. Intriguingly, the following result reveals that these effects cancel each other, and all noise distributions lead to the same rate as Gaussian noise. The first part of this theorem can be found in [53], and the second part is a straightforward variant that is new to the best of our knowledge.

Theorem 3.9 (Gaussian Coding Schemes for Non-Gaussian Channels)**.** Under the preceding mismatched nearest-neighbor decoding setting with input distribution $Q_X \sim \mathrm{N}(0, \Gamma)$ and an arbitrary additive noise distribution P_Z having mean μ and variance σ^2, we have the following:

(i) The GMI is given by

$$I_{\mathrm{GMI}}(Q_X) = \frac{1}{2} \log \left(1 + \frac{\Gamma}{\mu^2 + \sigma^2} \right). \tag{3.50}$$

(ii) The fixed-cost LM rate with a single auxiliary cost $a_1(x) = x$ is given by

$$I'_{\mathrm{LM}}(Q_X, a_1) = \frac{1}{2} \log \left(1 + \frac{\Gamma}{\sigma^2} \right). \tag{3.51}$$

In the proof of Theorem 3.9 (given below), we will additionally present an elementary proof of ensemble tightness for the GMI that is specific to this setting [53]. In contrast, we make no claims regarding the ensemble tightness of the fixed-cost LM rate.

Theorem 3.9 can be viewed both positively and negatively. On the positive side, we achieve C_{AWGN} even though we designed the code for the wrong channel, meaning that Gaussian codes form a *robust* communication scheme. On the negative side, since Gaussian noise is the most harmful noise, the matched capacity is strictly higher than C_{AWGN}, so we have failed to achieve the best possible rate. Stated differently, by designing for the worst case noise, we are guaranteed to achieve the worst-case capacity and no better.

From the second part of Theorem 3.9, we can conclude that the improvement of the fixed-cost LM rate amounts *at least* to replacing $\mathbb{E}[Z^2] = \mu^2 + \sigma^2$ by $\mathrm{Var}[Z] = \sigma^2$. The gap between the two can be arbitrarily large; we demonstrate this in Figure 3.2, where we plot the rates as a function of μ with $\Gamma = \sigma^2 = 1$.

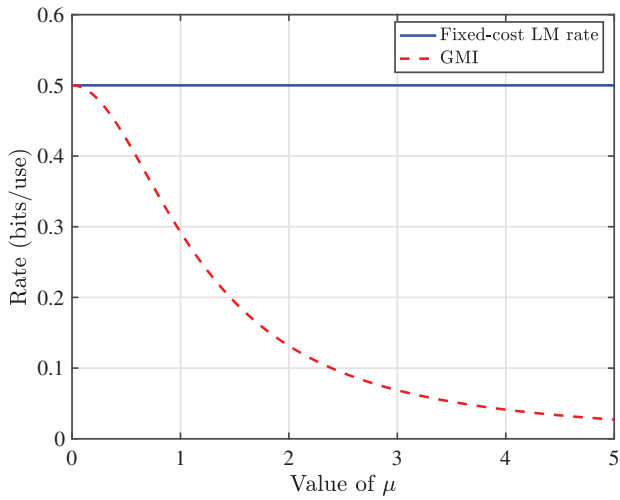


Figure 3.2: Mismatched noise distribution example: GMI and fixed-cost LM rate as a function of μ with $\Gamma = \sigma^2 = 1$.

To see why the choice $a(x) = x$ in the second part removes the dependence on the mean μ, we consider the decoding rule $\widetilde{q}(x, y) = e^{(y-\mu-x)^2}$, which provides a natural counterpart to the above choice when μ is known. Expanding the square yields

$$-\log \widetilde{q}(x, y) = (y - x)^2 - 2\mu y + 2\mu x. \tag{3.52}$$

The term $2\mu y$ has no impact on the decoder, since it depends only on the output. On the other hand, the term $2\mu x$ does have an impact in general. However, if we use a codebook such that $\sum_{i=1}^{n} x_i$ is the same for each codeword $\boldsymbol{x}$, then there is no impact. By using cost-constrained random coding with $a(x) = x$, we ensure that $\sum_{i=1}^{n} x_i$ is *nearly* the same for every codeword, which is enough to ensure that the LM rates corresponding to $q(x, y) = e^{-(y-x)^2}$ and $\widetilde{q}(x, y)$ above are the same.

Under the decoding metric $\widetilde{q}(x, y)$, the GMI also achieves the improved rate $\frac{1}{2}\log(1+\frac{\Gamma}{\sigma^2})$. However, such a choice amounts to subtracting the mean μ at the output, and hence, it requires the mean to be known in the first place. In contrast, using the fixed-cost LM rate, we attain the improvement without knowledge of μ.

Proof of Theorem 3.9

Evaluation of the GMI. We rewrite (3.4) as

$$I_{\mathrm{GMI}}(Q_X) = \sup_{s\geq 0} s\mathbb{E}[\log q(X,Y)] - \mathbb{E}[\log \mathbb{E}[q(\overline{X},Y)^s|Y]] \tag{3.53}$$

$$= \sup_{s\geq 0} \frac{1}{2}\log(1+2s\Gamma) + \frac{s(\Gamma+\mu^2+\sigma^2)}{1+2\Gamma s} - s(\mu^2+\sigma^2), \tag{3.54}$$

where (3.50) follows by substituting $q(x,y) = e^{-(y-x)^2}$ and evaluating the expectations explicitly; this requires some technical but elementary integration. Differentiating the right-hand side of (3.54), it is a straightforward exercise to verify that the choice $s = \frac{1}{2(\mu^2+\sigma^2)}$ makes the derivative vanish. Hence, since the objective function defining the GMI is always concave in s, we conclude that this choice of s must be globally optimal. With some algebraic manipulation, this choice leads to (3.50).

An Elementary Ensemble-Tightness Proof for the GMI. While the ensemble-tightness of the GMI is already established in Theorem 3.7, it is instructive to provide a direct proof that is not only significantly simpler, but also comes with an interesting alternative proof of the achievability part. The analysis proceeds in three steps, outlined as follows:

- By the circular symmetry of $P_{\boldsymbol{X}} = Q_X^n$, and similar symmetry in the nearest-neighbor decoding rule, it can be verified that the conditional error probability $\mathbb{P}[\mathrm{error}|\boldsymbol{z}]$ depends on the noise realization $\boldsymbol{z}$ only through its power $\|\boldsymbol{z}\|^2$ [53].
- Moreover, the conditional random-coding error probability $\mathbb{P}[\mathrm{error}|\boldsymbol{z}]$ is non-decreasing as a function of $\|\boldsymbol{z}\|^2$. To see this, we suppose that $\|\boldsymbol{y}-\overline{\boldsymbol{x}}\|_2 \leq \|\boldsymbol{y}-\boldsymbol{x}\|_2$ for some incorrect $\overline{\boldsymbol{x}}$ and $\boldsymbol{y} = \boldsymbol{x}+\boldsymbol{z}$, and proceed by showing that the same is true with

$\boldsymbol{y}' = \boldsymbol{x} + (1+\delta)\boldsymbol{z}$ for some $\delta > 0$:

$$\begin{aligned} \|\boldsymbol{y}' - \overline{\boldsymbol{x}}\|_2 &= \|\boldsymbol{y} + \delta\boldsymbol{z} - \overline{\boldsymbol{x}}\|_2 && (3.55)\\ &\le \|\boldsymbol{y} - \overline{\boldsymbol{x}}\|_2 + \delta\|\boldsymbol{z}\|_2 && (3.56)\\ &\le \|\boldsymbol{z}\|_2 + \delta\|\boldsymbol{z}\|_2 && (3.57)\\ &= \|\boldsymbol{y}' - \boldsymbol{x}\|_2, && (3.58) \end{aligned}$$

where (3.56) follows from the triangle inequality, and (3.57) follows since $\boldsymbol{x}$ is closer to $\boldsymbol{y}$ than $\overline{\boldsymbol{x}}$ is.

- As a result, the asymptotic behavior of the error probability is the same for any noise distribution such that $\frac{1}{n}\|\boldsymbol{Z}\|^2 \to \mathbb{E}[Z^2]$ almost surely. In particular, by the well-known fact that (3.50) is the (tight) rate achieved under $\mathrm{N}(0, \mathbb{E}[Z^2])$ additive Gaussian noise, we deduce that it is also the (ensemble-tight) rate achieved for any memoryless channel with the same second moment $\mathbb{E}[Z^2] = \mu^2 + \sigma^2$.

The interested reader is referred to [53] for the details.

Evaluation of the Fixed-Cost LM Rate. We rewrite (3.15) as

$$\begin{aligned} I'_{\mathrm{LM}}(Q_X, a) = \sup_{s\ge 0, r} \; & s\mathbb{E}[\log q(X,Y)] + r\mathbb{E}[a(X)] \\ & - \mathbb{E}[\log \mathbb{E}[q(\overline{X}, Y)^s e^{ra(\overline{X})} | Y]] && (3.59)\\ = \sup_{s\ge 0, r} \; & \frac{1}{2}\log(1+2s\Gamma) - \frac{\Gamma \cdot r(r+4s\mu)}{2(1+2\Gamma s)} \\ & + \frac{s(\Gamma + \mu^2 + \sigma^2)}{1+2\Gamma s} - s(\mu^2+\sigma^2), && (3.60) \end{aligned}$$

where (3.60) follows by substituting the choices of q and $a_1(\cdot) = a(\cdot)$ and evaluating the expectations. Only the second term in (3.60) depends on r, and it is a simple differentiation exercise to show that the optimal choice is $r = -2s\mu$. Substituting this choice and suitably canceling terms, we obtain

$$I'_{\mathrm{LM}}(Q_X, a) = \sup_{s\ge 0} \frac{1}{2}\log(1+2s\Gamma) + \frac{s(\Gamma+\sigma^2)}{1+2\Gamma s} - s\sigma^2. \tag{3.61}$$

This matches (3.54) with σ^2 in place of $\mu + \sigma^2$. Hence, since we already know that (3.54) leads to (3.50), we deduce that (3.61) leads to (3.51).

Variations and Generalizations

The first part of Theorem 3.9 holds not only for i.i.d. Gaussian codes, but also *shell codes*, in which $\boldsymbol{X}$ is drawn uniformly from the sphere of radius $\sqrt{n\Gamma}$ [53]. This distribution can be thought of as corresponding to cost-constrained random coding with $a(x) = x^2$ and $\delta = 0$, and indeed, it can be verified analytically that the fixed-cost rate of Theorem 3.4 does not improve on the GMI in this example when $L = 1$ and $a_1(x) = x^2$.

The first part of Theorem 3.9 has also been extended to additive multiple-access channels [53], fading channels [53, 57, 105] (see also Section 3.8.3), and parallel additive noise channels [53], and refined characterizations of the fixed-error asymptotics have been given [84].

In addition, regarding the second part of Theorem 3.9, a similar improvement replacing the second moment $\mathbb{E}[Z^2]$ by the variance $\mathrm{Var}[Z]$ is known to hold in the low-SNR regime under binary phase shift keying (BPSK) and constant-composition random coding [39].

3.8.3 Fading Channels

In our final example, we turn to a class of channels exhibiting *fading*, which is fundamental to wireless communication scenarios (*cf.*, Section 1.2.2). Fading channels are usually modeled as being *complex-valued*, and we adopt such an approach here.[3]

The setup that we consider adopts fairly strong memorylessness assumptions to keep the analysis simple, while still conveying some of the main ideas behind mismatch in fading channels. A more comprehensive treatment (including this special case) can be found in [57, 105].

We consider a memoryless time-varying model of the form

$$Y_i = H_i X_i + Z_i, \tag{3.62}$$

where $X_i \in \mathbb{C}$ is the input, $Z_i \in \mathbb{C}$ is additive noise, and $H_i \in \mathbb{C}$ is a *fading coefficient*. We assume that $\{Z_i\}_{i=1}^n$ is i.i.d. on $\mathrm{CN}(0, \sigma^2)$ (i.e., the variance is $\frac{\sigma^2}{2}$ each for the real and imaginary parts), and that $\{H_i\}$ is an i.i.d. sequence with density function P_H. In addition, we assume

[3] The achievability results (i.e., GMI and LM rate) we have given for real-valued channels are directly applicable to the complex-valued setting, and the proofs apply without change.

that $\boldsymbol{X} = (X_1, \ldots, X_n)$, $\boldsymbol{H} = (H_1, \ldots, H_n)$, and $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ are all mutually independent.

Perfect Channel Knowledge. If each random realization $H_i = h_i$ were known perfectly at the decoder, then due to the Gaussianity of the noise, the optimal decoding rule would be the following *weighted* version of the nearest-neighbor rule:

$$\hat{m} = \arg\min_{j=1,\ldots,M} \sum_{i=1}^{n} |y_i - h_i x_i^{(j)}|^2. \tag{3.63}$$

In addition, under a power constraint $\mathbb{E}[|X|^2] \leq \Gamma$ (i.e., $c(x) = |x|^2$), the capacity is achieved by a complex Gaussian input distribution $\mathrm{CN}(0, \Gamma)$, and is given by [57, 102]

$$C_{\text{AWGN}}(\Gamma, \sigma^2, P_H) = \mathbb{E}\left[\log\left(1 + \frac{|H|^2\Gamma}{\sigma^2}\right)\right], \tag{3.64}$$

where $H \sim P_H$, and Γ is the input power. Note that unlike the previous examples, there is no factor of $\frac{1}{2}$ here, as we are in the complex-valued setting rather than the real setting.

In fact, the achievability of (3.64) extends to non-Gaussian noise distributions analogously to the first part of Theorem 3.9 above [53], but we focus on Gaussian noise here for simplicity.

Imperfect Channel Knowledge. Our focus in this example is on *uncertainty in the fading coefficients* H_i. To address this, we adopt a simple uncertainty model in which

$$H_i = \widehat{H}_i + \widetilde{H}_i, \quad \mathbb{E}[\widetilde{H}_i | \widehat{H}_i] = 0, \tag{3.65}$$

where $\widehat{H}_i$ is a possibly-random estimate of H known at the decoder, and $\widetilde{H}_i$ represents an unknown *conditionally zero-mean* error term. We make the simplifying assumption that the pairs $\{(\widehat{H}_i, \widetilde{H}_i)\}_{i=1}^n$ are i.i.d. with respect to $i = 1, \ldots, n$, and independent of the channel input and noise.

In the case that the joint density function of $(\widehat{H}_i, \widetilde{H}_i)$ is unknown (or even when it is known but difficult to design a corresponding optimal coding scheme), it is natural to apply weighted nearest-neighbor coding

similarly to (3.63), but with each weight given only by the realization $\widehat{H}_i = \widehat{h}_i$ of the corresponding estimate:

$$\hat{m} = \underset{j=1,\dots,M}{\arg\min} \sum_{i=1}^{n} |y_i - \widehat{h}_i x_i^{(j)}|^2. \tag{3.66}$$

This is a mismatched decoding rule, in the sense that it would be optimal under a model of the form $Y = \widehat{H}X + Z$ with zero-mean Gaussian noise, in contrast with the true model. The corresponding decoding metric is given by $q(x, (y, \widehat{h})) = e^{-(y-\widehat{h}x)^2}$. Note that here and subsequently, since $\widehat{H}$ is known at the decoder, it is treated as part of the output, so that $(Y, \widehat{H})$ plays the role of the usual Y.

The following result gives an exact expression for the GMI [57]; to our knowledge, the LM rate is yet to be studied in fading scenarios.

Theorem 3.10 (Fading Channels)**.** Consider the preceding complex-valued channel fading setup with a known estimate $\widehat{H}$ at the output, and a conditionally zero-mean error term $\widetilde{H}$. Under i.i.d. random coding with $X \sim \mathrm{N}(0, \Gamma)$, along with weighted nearest-neighbor decoding according to (3.66), the GMI is given by

$$I_{\mathrm{GMI}}(Q_X) = \mathbb{E}\left[\log\left(1 + \frac{|\widehat{H}|^2\Gamma}{\mathbb{E}[|\widetilde{H}|^2|\widehat{H}]\Gamma + \sigma^2}\right)\right]. \tag{3.67}$$

We observe that the GMI takes a similar form to the matched capacity (3.64), except that the signal power in the numerator is multiplied by $|\widehat{H}|^2$ instead of $|H|^2$, and the denominator contains an extra term $\mathbb{E}[|\widetilde{H}|^2|\widehat{H}]\Gamma$. The intuition here is that we are treating the unknown error term $\widetilde{H}$ as noise, and since that term is multiplied by X, the overall additional "noise power" given $\widehat{H}$ is $\mathbb{E}[|\widetilde{H}|^2|\widehat{H}]\Gamma$.

Theorem 3.10 reveals that mismatch in the fading coefficient can be much more harmful than mismatch in the noise distribution (*cf.*, Section 3.8.2). For instance, supposing for simplicity that $\widehat{H}$ takes some *deterministic* non-zero value, we find that $I_{\mathrm{GMI}}(Q_X)$ remains *bounded in the high-power regime* $\Gamma \to \infty$, converging to $\log(1 + \frac{|\widehat{H}|^2}{\mathbb{E}[\widetilde{H}^2]})$. In contrast, the matched capacity (3.64) grows unbounded. An illustration is given in Figure 3.3, where we plot the GMI and matched capacity as a function of Γ when $\Gamma = \sigma^2 = 1$, $|\widehat{H}|^2 = 1$ (deterministically), and $\mathbb{E}[|\widetilde{H}|^2] = 1$.

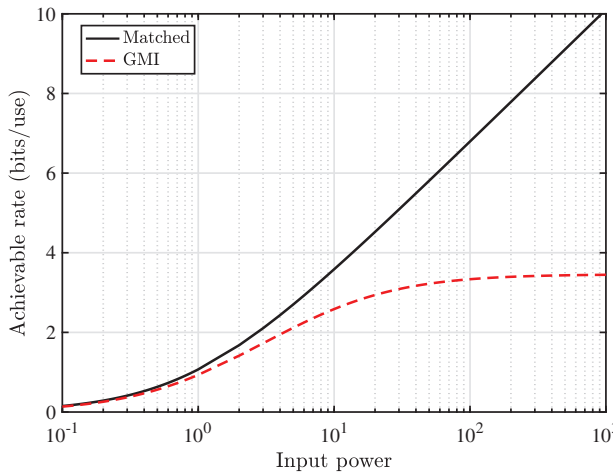


Figure 3.3: Fading channel example: Matched capacity and GMI as a function of Γ with $\sigma^2 = 1$, $|\widehat{H}|^2 = 1$ (deterministically) and $\mathbb{E}[|\widetilde{H}|^2] = 1$.

Proof of Theorem 3.10

Recall that the pair $(Y, \widehat{H})$ plays the role of the channel output. We first write the GMI as

$$I_{\mathrm{GMI}}(Q_X) = \sup_{s \geq 0} \mathbb{E}[I_{\mathrm{GMI}}(Q_X, s, \widehat{H})], \tag{3.68}$$

where

$$I_{\mathrm{GMI}}(Q_X, s, \widehat{h}) = s\mathbb{E}[\log q(X, (Y, \widehat{h}))] - \mathbb{E}[\log \mathbb{E}[q(\overline{X}, (Y, \widehat{h}))^s)]] \tag{3.69}$$

$$= -s\mathbb{E}[|Y - \widehat{h}X|^2] - \mathbb{E}[\log \mathbb{E}[e^{-|Y - \widehat{h}\overline{X}|^2} | Y]] \tag{3.70}$$

with implicit conditioning on $\widehat{H} = \widehat{h}$ throughout. Notice that this expression is similar to that of (3.45) in the mismatched signal level example. While we are working with complex-valued (rather than real-valued) random variables here, we can still explicitly evaluate the expectations in a similar manner, leading to

$$I_{\mathrm{GMI}}(Q_X, s, \widehat{h}) = \log(1 + |\widehat{h}|^2\Gamma s) + \frac{s(|\widehat{h}|^2\Gamma + \Gamma_{XY}(\widehat{h}))}{1 + |\widehat{h}|^2\Gamma s} - s\Gamma_{XY}(\widehat{h}), \tag{3.71}$$

where we write $\Gamma_{XY}(\widehat{h}) = \mathbb{E}[|\widetilde{H}|^2 | \widehat{H} = \widehat{h}]\Gamma + \sigma^2$ for brevity.

Next, it is a simple exercise to differentiate (3.71) and verify that the choice $s = \frac{1}{\Gamma_{XY}}$ yields a derivative of zero, regardless of the value of $\widehat{h}$. Since the objective function is concave in s, we conclude that such a choice is globally optimal. Substituting this optimal choice, denoted by s^*, we obtain the following after some algebraic manipulation:

$$I_{\text{GMI}}(Q_X, s^*, \widehat{h}) = \log\left(1 + \frac{|\widehat{h}|^2\Gamma}{\Gamma_{XY}(\widehat{h})}\right). \tag{3.72}$$

The proof is concluded by averaging over $\widehat{H}$ and substituting $\Gamma_{XY}(\cdot)$.

4

Mismatch in Rate Distortion Theory

4.1 Introduction

Thus far, we have focused on channel coding with a mismatched decoding rule. In this section, we turn to the distinct problem of source coding subject to a fidelity constraint, i.e., rate-distortion theory [22, Ch. 10]. As with channel coding, this represents one of the most fundamental coding problems in information theory.

The setup of rate-distortion theory (detailed in Section 4.2 below) consists of a memoryless source, an encoder mapping source sequences to indices, and a decoder mapping these indices to a reconstructed sequence. The quality of the reconstruction is measured by a distortion function. Assuming that the decoder is deterministic, we can view the set of all reconstructed sequences as forming a codebook. In addition, given such a codebook, the optimal encoding rule is to map the source sequence to the index corresponding to the codeword with the lowest distortion.

If the encoder does not know the true distortion measure under consideration, or if the optimal encoding rule is ruled out due to implementation constraints, then it is natural to consider a variant of this problem with *mismatched encoding*, in which the encoder minimizes a

given (possibly suboptimal) *encoding metric* between the source and its reconstruction. Some motivating examples are given in Section 4.2.1.

As well as being of interest in its own right, the mismatched rate-distortion problem has interesting analogies with the mismatched channel coding problem. Among other things, we will see that the multi-letter extension (*cf.*, Section 2.7) of a certain random-coding achievability result has a matching converse. In contrast, for channel coding it remains open as to whether the multi-letter extension of the LM rate coincides with the mismatch capacity in general.

In Sections 4.7 and 4.8, we consider a different type of mismatch, in which a single distortion measure is adopted throughout the system, but a suboptimal random coding distribution is used. Section 4.7 considers discrete memoryless sources, while Section 4.8 focuses on Gaussian coding for non-Gaussian sources. In the latter setting, in analogy with the mismatched channel coding example of Section 3.8.2 (Gaussian coding with non-Gaussian noise), we find that all real-valued sources with a given second moment yield the same rate-distortion trade-off as the Gaussian source.

This section is predominantly based on the work of Lapidoth [54], with various aspects also relating to Sakrison [73, 74], Yang and Kieffer [106], and Dembo and Kontoyiannis [28]. We note that compared to [54], we consider a slightly simplified mismatched encoding problem formulation (see below).

4.2 Problem Setup

We consider a discrete memoryless source (DMS), in which each symbol is drawn from a distribution $\Pi_X(x)$ on a finite alphabet $\mathcal{X}$. The resulting n-letter distribution is denoted by $\Pi_X^n(\boldsymbol{x}) \triangleq \prod_{i=1}^n \Pi_X(x_i)$. The source sequence $\boldsymbol{X} \sim \Pi_X^n$ is received as the input to an *encoder* that maps $\mathcal{X}^n$ to $\{1, \ldots, M\}$ for some integer M. The *rate* of the code is denoted by $R = \frac{1}{n} \log M$.

We consider encoding rules that are dictated by a codebook $\mathcal{C} = \{\widehat{\boldsymbol{x}}^{(1)}, \ldots, \widehat{\boldsymbol{x}}^{(M)}\}$, where each codeword lies in $\widehat{\mathcal{X}}^n$ for some finite reconstruction alphabet $\widehat{\mathcal{X}}$. In particular, in accordance with the

discussion in Section 4.1, we assume that the encoder chooses[1]

$$m = \arg\min_{j=1,\ldots,M} d_0^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}^{(j)}), \quad d_0^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = \sum_{i=1}^{n} d_0(x_i, \widehat{x}_i), \tag{4.1}$$

where $d_0(x, \widehat{x})$ is a given *encoding metric* on $\mathcal{X} \times \widehat{\mathcal{X}}$. Except where stated otherwise, the results that we present hold under arbitrary tie-breaking methods in (4.1).

We focus on the trivial decoder that maps m to $\widehat{\boldsymbol{x}}^{(m)}$. This is in slight contrast to Lapidoth [54], who allowed the decoder to apply post-processing to the codeword to obtain the final estimate. The former setup will lead to slightly simpler results, while still conveying the same key ideas and techniques.

When the source realization is $\boldsymbol{x}$ and its reconstruction is $\widehat{\boldsymbol{x}}$, the distortion incurred is given by $d_1^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = \sum_{i=1}^n d_1(x_i, \widehat{x}_i)$ for some *true distortion measure* d_1 on $\mathcal{X} \times \widehat{\mathcal{X}}$. Our goal is to achieve a normalized distortion $\frac{1}{n} d_1^n(\boldsymbol{x}, \widehat{\boldsymbol{x}})$ that, with high probability, does not exceed a given threshold D. Throughout the section, we assume that both d_0 and d_1 are non-negative and finite-valued, and we define $d_{1,\max} \triangleq \max_{x,\widehat{x}} d_1(x, \widehat{x})$.

Formally, we have the following.

Definition 4.1 (Mismatched Rate-Distortion and Distortion-Rate Functions)**.** Under the preceding mismatched rate-distortion setup specified by (Π_X, d_0, d_1), a rate-distortion pair (R, D_1) is said to be *achievable* if, for all $\delta > 0$, there exists a sequence of codebooks (indexed by n) with $M \leq e^{n(R+\delta)}$ codewords such that

$$\mathbb{P}\left[\frac{1}{n} d_1^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}) \geq D_1 + \delta\right] \leq \delta \tag{4.2}$$

for sufficiently large n (depending on δ), where $\boldsymbol{X} \sim \Pi_X^n$, and $\widehat{\boldsymbol{X}}$ is the resulting estimate. Given D_1, the *mismatched rate-distortion function* $R^*(D_1)$ is defined to be the smallest R such that (R, D_1) is achievable, and given R, the *mismatched distortion-rate function* $D_1^*(R)$ is defined to be smallest D_1 such that (R, D_1) is achievable.

[1] Here we minimize an additive metric, whereas in channel coding we maximized a multiplicative metric. In both settings, we can readily switch between these notions, e.g., maximizing $\prod_{i=1}^n q(x_i, y_i)$ is equivalent to minimizing $\sum_{i=1}^n (-\log q(x_i, y_i))$.

Before proceeding, we state the following useful result, which can be viewed as an analog of the result in Lemma 3.1 concerning the input-constrained mismatch capacity.

Lemma 4.1 (Convexity of Mismatched Distortion-Rate Function)**.** Under the preceding setup, the function $D_1^*(R)$ is convex and non-increasing in R.

The monotonicity is immediate from Definition 4.1, and the convexity follows by a standard concatenation argument: Use a codebook achieving $D_1^*(R_1)$ for the first λn symbols, and a codebook achieving $D_1^*(R_2)$ for the last $(1-\lambda)n$ symbols, yielding a total distortion of $\lambda D_1^*(R_1) + (1-\lambda)D_1^*(R_2)$ at rate $\lambda R_1 + (1-\lambda)R_2$.

Remark 4.1 (Expected vs. High-Probability Distortion Criteria)**.** The results that we present under the excess distortion criterion (4.2) also hold under an average distortion criterion of the form $\mathbb{E}[\frac{1}{n}d_1^n(\boldsymbol{X}, \widehat{\boldsymbol{X}})] \leq D_1 + \delta$. The achievability result is stated in [19, Thm. 3], and can also be deduced from the high-probability result by noting that (4.2) implies $\mathbb{E}[\frac{1}{n}d_1^n(\boldsymbol{X}, \widehat{\boldsymbol{X}})] \leq D_1 + \delta + \delta d_{1,\max}$, and δ can be scaled down by $1 + d_{1,\max}$ since it is arbitrarily small in Definition 4.1. In Section 4.6, we will present a multi-letter converse result for the excess distortion criterion, but its proof will immediately imply the same for the expected distortion criterion.

4.2.1 Discussion and Motivating Examples

Here we discuss two motivating examples where the mismatched rate-distortion problem may be of interest.

Unknown Distortion Measure. The mismatched rate-distortion problem naturally arises in the case that it is not possible to know the true distortion measure $d_1(x, \widehat{x})$. For instance, when it comes to compressing an image or audio signal,[2] it is difficult to mathematically quantify what makes a reconstruction "good". Nevertheless, one might still try

[2]These are mentioned for the sake of intuition, but it should be noted that the memoryless property and the assumption of additive distortion measures should not be expected to hold in such cases.

to perform compression according to some heuristic choice, such as $d_0(x, \widehat{x}) = (x - \widehat{x})^2$.

Similarly, after designing a coding scheme for some distortion measure d_0, the system designer might change their mind and decide that a different measure d_1 is more suitable. It is then of interest to understand how the performance of the original coding scheme degrades compared to a new scheme based on d_1. If the loss is minimal, one may prefer to simply keep the mismatched design.

In both of these settings, the notion of mismatched rate-distortion region in Definition 4.1 is questionable, as the optimal codebook depends on both d_0 and d_1. Nevertheless, similarly to the channel coding setting, we will state the achievability results for a *fixed random coding distribution*, hence allowing us to assess the performance of random codebooks that are designed without knowledge of d_1.

Finite-Precision Arithmetic. Suppose for concreteness that the true distortion measure is $d_1(x, \widehat{x}) = (x - \widehat{x})^2$. If the encoder is constrained to use finite-precision arithmetic, it may consist of first applying a scalar quantizer to $\boldsymbol{x}$, and then choosing a similarly-quantized codeword $\widehat{\boldsymbol{x}}$ from the codebook based on a finite-precision distortion measure. For instance, letting $\Phi(x)$ be a scalar quantizer, and letting $\widehat{\mathcal{X}}$ be a its range (i.e., a finite subset of $\mathbb{R}$), one can consider the following distortion measure at the encoder, with $x \in \mathbb{R}$ and $\widehat{x} \in \widehat{\mathcal{X}}$:

$$d_0(x, \widehat{x}) = (\Phi(x) - \widehat{x})^2. \tag{4.3}$$

In this case, assuming the system model is known perfectly, the notion of mismatched rate-distortion region in Definition 4.1 is justified: One knows that the system is constrained by the finite-precision arithmetic and scalar quantizer, and hence, one can design a codebook specifically targeted at combating this constraint.

4.3 Achievability Result

In this subsection, we provide an achievable distortion level for a given rate, i.e., an upper bound on $D_1^*(R)$. The proof uses a form of constant-composition random coding, and in this sense, the result can be viewed

as a counterpart to the LM rate for channel coding (*cf.*, Section 2.3.2). A similar analysis could also be performed for i.i.d. random coding, but this is omitted for brevity.

We fix a reconstruction distribution $Q_{\widehat{X}} \in \mathcal{P}(\widehat{\mathcal{X}})$, let $Q_{\widehat{X},n} \in \mathcal{P}_n(\widehat{\mathcal{X}})$ be a type with the same support of $Q_{\widehat{X}}$ such that $\|Q_{\widehat{X},n} - Q_{\widehat{X}}\|_\infty \leq \frac{1}{n}$, and draw each codeword $\widehat{\boldsymbol{X}}^{(j)}$ independently from

$$P_{\widehat{\boldsymbol{X}}}(\widehat{\boldsymbol{x}}) = \frac{1}{|\mathcal{T}^n(Q_{\widehat{X},n})|} \mathbb{1}\{\widehat{\boldsymbol{x}} \in \mathcal{T}^n(Q_{\widehat{X},n})\}, \tag{4.4}$$

where $\mathcal{T}^n(Q_{\widehat{X},n})$ is the type class corresponding to $Q_{\widehat{X},n}$. We have the following [54].

Theorem 4.2 (Achievable Mismatched Distortion-Rate Function)**.** Under the mismatched rate-distortion setup with a discrete memoryless source Π_X and distortion measures (d_0, d_1), the following distortion is achievable at rate R via constant-composition random coding with an auxiliary distribution $Q_{\widehat{X}} \in \mathcal{P}(\widehat{\mathcal{X}})$:

$$\overline{D}_1(Q_{\widehat{X}}, R) = \max_{\widetilde{P}_{X\widehat{X}} \in \widetilde{\mathcal{P}}} \mathbb{E}_{\widetilde{P}}[d_1(X, \widehat{X})], \tag{4.5}$$

where

$$\widetilde{\mathcal{P}} = \left\{ \widetilde{P}_{X\widehat{X}} \colon \widetilde{P}_{X\widehat{X}} \in \underset{\substack{\widetilde{P}_{X\widehat{X}} \colon \widetilde{P}_X = \Pi_X, \widetilde{P}_{\widehat{X}} = Q_{\widehat{X}}, \\ I_{\widetilde{P}}(X;\widehat{X}) \leq R}}{\arg\min} \mathbb{E}_{\widetilde{P}}[d_0(X, \widehat{X})] \right\}. \tag{4.6}$$

Consequently, we have $D_1^*(R) \leq \min_{Q_{\widehat{X}}} \overline{D}_1(Q_{\widehat{X}}, R)$.

The proof is given in Section 4.5. The result can be understood intuitively by interpreting $\widetilde{P}_{X\widehat{X}}$ as the joint empirical distribution of $(\boldsymbol{x}, \widehat{\boldsymbol{x}})$. The distortion incurred is $d_1^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = n\mathbb{E}_{\widetilde{P}}[d_1(X, \widehat{X})]$, and by the nature of the encoding function, the joint type is one that minimizes $\mathbb{E}_P[d_0(X, \widehat{X})]$. The constraint $\widetilde{P}_X = \Pi_X$ follows by a basic typicality argument, the constraint $\widetilde{P}_{\widehat{X}} = Q_{\widehat{X}}$ arises from the constant-composition nature of the codebook, and the constraint $I_{\widetilde{P}}(X; \widehat{X}) \leq R$ arises because the probability of *any* random codeword inducing $I_{\widetilde{P}}(X; \widehat{X}) > R$ is negligible. Finally, the maximum in (4.5) arises by treating ties in a

"worst-case" manner, i.e., assuming that when a tie occurs, the selected sequence is the one that maximizes d_1.

In the case that $d_0 = d_1$, all elements of $\widetilde{\mathcal{P}}$ provide the same distortion, and the bound of Theorem 4.2 simplifies to

$$\overline{D}_{1,\mathrm{Matched}}(Q_{\widehat{X}}, R) = \min_{\substack{\widetilde{P}_{X\widehat{X}}\colon\ \widetilde{P}_X = \Pi_X, \widetilde{P}_{\widehat{X}} = Q_{\widehat{X}}, \\ I_{\widetilde{P}}(X;\widehat{X}) \le R}} \mathbb{E}_{\widetilde{P}}[d_1(X, \widehat{X})]. \tag{4.7}$$

This corresponds to the usual distortion-rate function [22, Ch. 10], but with an additional constraint on the $\widehat{X}$-marginal. Moreover, this constraint can be removed by simply taking the minimum over all $Q_{\widehat{X}}$, since the latter is a free parameter.

We briefly that Theorem 4.2 is only stated and proved for discrete memoryless sources, and handling continuous sources is a potentially interesting open problem.

4.4 Examples

In this subsection, we provide two examples taken from [54].

4.4.1 Parallel Binary Sources

We begin with an analytical example showing that the achievable distortion of Theorem 4.2 is not tight, i.e., it can be strictly higher than the mismatched distortion-rate function. This example is closely related to the parallel channel example of Section 2.4.5, and leads to analogous observations. In fact, the analysis is also similar, and we therefore omit some repeated details.

The source and reconstruction alphabets are given by $\mathcal{X} = \widehat{\mathcal{X}} = \{0,1\}^2$, and we write $X = (X_1, X_2)$ and $\widehat{X} = (\widehat{X}_1, \widehat{X}_2)$ to denote the corresponding binary pairs. We let Π_X be uniform, so that $\Pi_X(x) = \frac{1}{4}$ for all x. For $x = (x_1, x_2)$ and $\widehat{x} = (\widehat{x}_1, \widehat{x}_2)$, the distortion measures are given as follows for some $\lambda \in [0,1]$:

$$d_0(x, \widehat{x}) = \lambda \mathbb{1}\{x_1 \neq \widehat{x}_1\} + (1-\lambda)\mathbb{1}\{x_2 \neq \widehat{x}_2\}, \tag{4.8}$$

$$d_1(x, \widehat{x}) = \frac{1}{2}\mathbb{1}\{x_1 \neq \widehat{x}_1\} + \frac{1}{2}\mathbb{1}\{x_2 \neq \widehat{x}_2\}. \tag{4.9}$$

Hence, the mismatched encoder treats the reconstructions of x_1 and x_2 with different weights, even though the two are treated equally under the true distortion measure.

We set $Q_{\widehat{X}}$ in Theorem 4.2 to be the uniform distribution. To understand the set $\widetilde{\mathcal{P}}$ in (4.6), we consider the mutual information constraint $I_{\widetilde{P}}(X;\widehat{X}) \leq R$. From the analysis in Section 2.4.5, we know that

$$I_{\widetilde{P}}(X;\widehat{X}) \geq (1 - H_2(\widetilde{\delta}_1)) + (1 - H_2(\widetilde{\delta}_2)), \tag{4.10}$$

where $\widetilde{\delta}_1$ and $\widetilde{\delta}_2$ are transition probabilities associated with the marginals $\widetilde{P}_{X_1\widehat{X}_1}$ and $\widetilde{P}_{X_2\widehat{X}_2}$. From the choice of d_0 in (4.8), the corresponding distortion incurred is $\lambda\widetilde{\delta}_1 + (1-\lambda)\widetilde{\delta}_2$, and we have

$$\begin{aligned} &\min_{\widetilde{P}_{X\widehat{X}}\colon\ \widetilde{P}_X = \Pi_X, \widetilde{P}_{\widehat{X}} = Q_{\widehat{X}}, I_{\widetilde{P}}(X;\widehat{X}) \leq R} \mathbb{E}_P[d_0(X,\widehat{X})] \\ &\quad \geq \min_{\widetilde{\delta}_1,\widetilde{\delta}_2\colon\ (1-H_2(\widetilde{\delta}_1))+(1-H_2(\widetilde{\delta}_2)) \leq R} \lambda\widetilde{\delta}_1 + (1-\lambda)\widetilde{\delta}_2. \end{aligned} \tag{4.11}$$

In fact, this lower bound holds with equality, since it is attained when $\widetilde{P}_{\widehat{X}_1|X_1}$ and $\widetilde{P}_{\widehat{X}_1|X_2}$ are independent BSCs with crossover probabilities $(\widetilde{\delta}_1^*, \widetilde{\delta}_2^*)$ achieving the minimum on the right-hand side. Moreover, it can be shown that the minimizer is unique on both sides of (4.11). Under the optimal parameters $(\widetilde{\delta}_1^*, \widetilde{\delta}_2^*)$ attained in (4.11), the resulting distortion incurred in (4.5) is

$$\overline{D}_1(Q_{\widehat{X}}, R) = \frac{1}{2}(\widetilde{\delta}_1^* + \widetilde{\delta}_2^*) \tag{4.12}$$

due to the choice of d_1 in (4.9).

Next, since the matched rate-distortion function of a binary symmetric source with Hamming distortion is $R^*_{\text{symm}}(D) = 1 - H_2(D)$ bits/symbol [22, Sec. 10.3.1], the matched rate-distortion function of the parallel source under consideration is $R^*_{\text{Matched}}(D_1) = 2(1-H_2(D_1))$. Hence, for a given rate R, the distortion D_1 incurred is the solution to $R = 2(1 - H_2(D_1))$. We claim that, in fact, the *mismatched* rate-distortion function $D_1^*(R)$ is identical to this matched one. This can be verified by generating a separate codebook for each source, and letting the overall codebook be the product of the two codebooks generated. The easiest way to see that this achieves the matched performance is

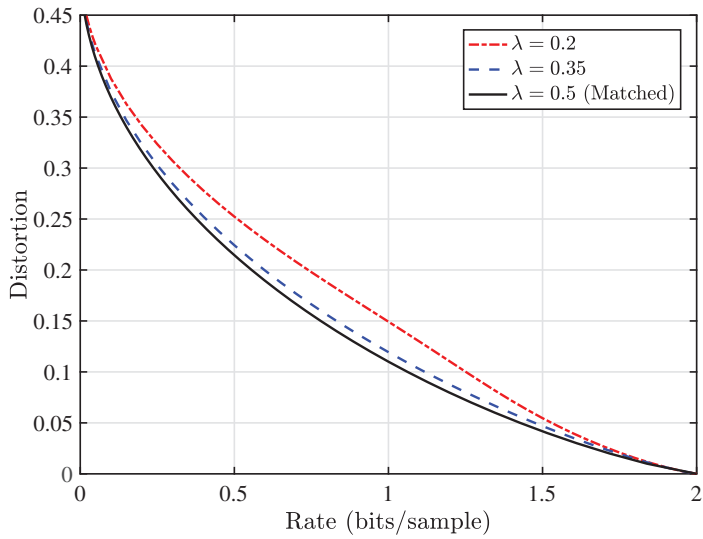


Figure 4.1: Parallel source example: Comparison of matched ($\lambda = 0.5$) and mismatched ($\lambda = 0.2, 0.35$) rate-distortion curves resulting from Theorem 4.2, in the case that the source and the auxiliary distribution are both uniform, i.e., $\Pi_X = Q_{\widehat{X}} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.

to note that d_0 and d_1 are equivalent under such a product codebook structure.

Summarizing the above, the mismatched distortion rate function yields the optimal value D_1^* satisfying $R = 2(1 - H_2(D_1^*))$, whereas the bound of Theorem 4.2 yields the distortion $\frac{1}{2}(\widetilde{\delta}_1^* + \widetilde{\delta}_2^*)$ under the pair $(\widetilde{\delta}_1, \widetilde{\delta}_2)$ minimizing (4.11). Since the latter constrains $(1 - H_2(\widetilde{\delta}_1)) + (1 - H_2(\widetilde{\delta}_2)) \leq R$, we can use the strict concavity of $H_2(\cdot)$ to deduce that $\frac{1}{2}(\widetilde{\delta}_1^* + \widetilde{\delta}_2^*) \geq D_1^*$, with strict inequality when $\delta_1^* \neq \delta_2^*$. Finally, one can verify that $\delta_1^* \neq \delta_2^*$ whenever $\lambda \neq \frac{1}{2}$ in (4.8); see Figure 4.1 for an illustrative numerical example. Hence, we conclude that Theorem 4.2 fails to achieve the mismatched distortion-rate function in such cases. While we only showed this for uniform $Q_{\widehat{X}}$, any other choice would be suboptimal even under matched encoding.

This example suggests that Theorem 4.2 could be improved via multi-user coding techniques, analogously to the studies of multi-user coding techniques for single-user channel coding surveyed in Section 6.

To our knowledge, this idea is yet to be investigated beyond the present example.

4.4.2 Ternary Symmetric Source

We consider an equiprobable ternary source: $\mathcal{X} = \{0, 1, 2\}$, and $\Pi_X(x) = \frac{1}{3}$ for each x. The distortion measures are given by

$$d_0(x, \widehat{x}) = \mathbb{1}\{x \neq \widehat{x}\}, \tag{4.13}$$

$$d_1(x, \widehat{x}) = \mathbb{1}\{x \neq \widehat{x} \cap x \neq 2\}. \tag{4.14}$$

Hence, the mismatched encoder uses the Hamming distance, whereas under the true distortion measure, no penalty is incurred for $x = 2$.

The analysis given in [54] is outlined as follows:

- Under the uniform distribution $Q_{\widehat{X}} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, the unique minimizer $\widetilde{P}_{\widehat{X}|X}$ in (4.6) is a ternary symmetric channel. Moreover, the transition probability δ^* between differing symbols is such that $I_{\widetilde{P}}(X; \widehat{X}) = R$. The resulting distortion is $\mathbb{E}_{\widetilde{P}}[d_1(X, \widehat{X})] = \frac{2}{3} \cdot 2\delta^*$.
- Under the distribution $Q_{\widehat{X}} = (\frac{1}{2}, \frac{1}{2}, 0)$, the unique minimizer $\widetilde{P}_{\widehat{X}|X}$ in (4.6) maps $x = 2$ to $\{0, 1\}$ with probability $\frac{1}{2}$ each, and maps $\{0, 1\}$ to $\{0, 1\}$ according to a BSC. The crossover probability δ^* is such that $\frac{2}{3}(1 - H_2(\delta^*)) = R$, and the resulting distortion is $\mathbb{E}_{\widetilde{P}}[d_1(X, \widehat{X})] = \frac{2}{3} \cdot \delta^*$.

The first of these choices of $Q_{\widehat{X}}$ would be optimal if the true distortion were also d_0, whereas the second is optimal under d_1.

The achievable rate-distortion curves are plotted in Figure 4.2. For the case $Q_{\widehat{X}} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, we also plot the *matched* function given in (4.7), whereas for $Q_{\widehat{X}} = (\frac{1}{2}, \frac{1}{2}, 0)$ we found the matched and mismatched curves to coincide.

4.5 Achievability Proof

In this subsection, we prove Theorem 4.2. We make use of the method of types, including some of the properties stated in Section 2.6.1.

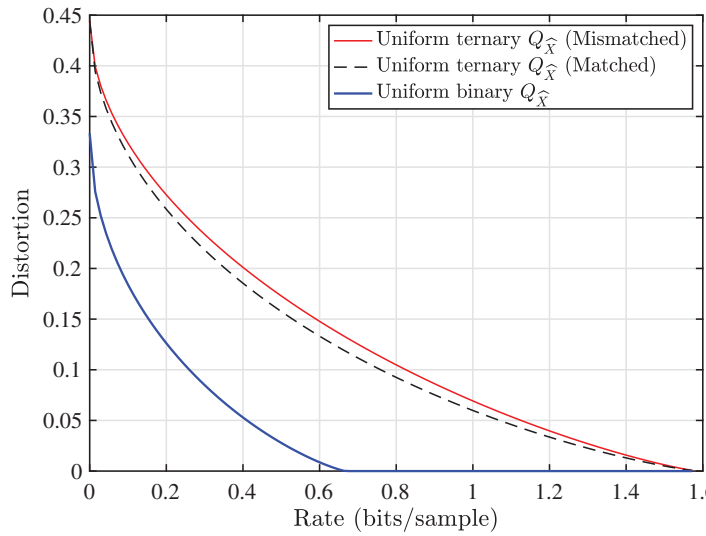


Figure 4.2: Ternary symmetric source example: Upper bounds on the mismatched distortion-rate function under the auxiliary distributions $Q_{\widehat{X}} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $Q_{\widehat{X}} = (\frac{1}{2}, \frac{1}{2}, 0)$.

Recall that we generate the codewords $\{\widehat{\boldsymbol{X}}^{(m)}\}_{m=1}^{M}$ independently according to the constant-composition codeword distribution $P_{\widehat{\boldsymbol{X}}}$ in (4.4). In accordance with Definition 4.1, we would like to show that the random-coding error probability

$$\bar{p}_{\mathrm{e}} \triangleq \mathbb{P}\left[\frac{1}{n} d_1^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}) \geq D_1 + \delta\right] \tag{4.15}$$

vanishes as $n \to \infty$, for arbitrarily small $\delta > 0$. Overloading the notation, we first consider the conditional error probability

$$\bar{p}_{\mathrm{e}}(\boldsymbol{x}) \triangleq \mathbb{P}\left[\frac{1}{n} d_1^n(\boldsymbol{x}, \widehat{\boldsymbol{X}}) \geq D_1 + \delta \,\middle|\, \boldsymbol{X} = \boldsymbol{x}\right]. \tag{4.16}$$

In the following, for a given type $P_X \in \mathcal{P}_n(\mathcal{X})$ associated with $\boldsymbol{x}$, we consider the possible joint types $\widetilde{P}_{X\widehat{X}} \in \mathcal{P}_n(\mathcal{X} \times \widehat{\mathcal{X}})$ induced between $\boldsymbol{x}$ and the codewords $\{\widehat{\boldsymbol{X}}^{(j)}\}_{j=1}^{M}$.

Lemma 4.3 (Occurrences of Joint Types)**.** Consider a random codebook $\mathcal{C}_n = \{\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(M)}\}$ with codewords of length n drawn independently from $P_{\widehat{\boldsymbol{X}}}$ in (4.4). For any $\delta' > 0$, conditioned on any $\boldsymbol{X} = \boldsymbol{x}$

with $\boldsymbol{x} \in \mathcal{T}^n(P_X)$, the following statements simultaneously hold with probability approaching one as $n \to \infty$:

1. For all $\widehat{\boldsymbol{x}} \in \mathcal{C}_n$, if $(\boldsymbol{x}, \widehat{\boldsymbol{x}}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})$ for some $\widetilde{P}_{X\widehat{X}} \in \mathcal{P}_n(\mathcal{X} \times \widehat{\mathcal{X}})$, then it must hold that $\widetilde{P}_X = P_X$ and $\widetilde{P}_{\widehat{X}} = Q_{\widehat{X},n}$.

2. Defining the set

$$\mathcal{S}_n^{(-)}(P_X) = \{\widetilde{P}_{X\widehat{X}} \in \mathcal{P}_n(\mathcal{X} \times \widehat{\mathcal{X}}): \widetilde{P}_X = P_X, \widetilde{P}_{\widehat{X}} = Q_{\widehat{X},n}, I_{\widetilde{P}}(X;\widehat{X}) \geq R + \delta'\}, \tag{4.17}$$

there does not exist a codeword $\widehat{\boldsymbol{x}} \in \mathcal{C}_n$ such that $(\boldsymbol{x}, \widehat{\boldsymbol{x}}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})$ with $\widetilde{P}_{X\widehat{X}} \in \mathcal{S}_n^{(-)}(P_X)$.

3. Conversely, for any joint type in the set

$$\mathcal{S}_n^{(+)}(P_X) = \{\widetilde{P}_{X\widehat{X}} \in \mathcal{P}_n(\mathcal{X} \times \widehat{\mathcal{X}}): \widetilde{P}_X = P_X, \widetilde{P}_{\widehat{X}} = Q_{\widehat{X},n}, I_{\widetilde{P}}(X;\widehat{X}) \leq R - \delta'\}, \tag{4.18}$$

there exists at least one codeword $\widehat{\boldsymbol{x}} \in \mathcal{C}_n$ for which $(\boldsymbol{x}, \widehat{\boldsymbol{x}}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})$.

Proof. The first claim follows trivially from the assumption that $\boldsymbol{x} \in \mathcal{T}^n(P_X)$ and the fact that we are using constant-composition random coding according to the type $Q_{\widehat{X},n}$.

The second and third claims are based on the following expression for the probability that there exists at least one codeword inducing a given joint type $\widetilde{P}_{X\widehat{X}}$:

$$\mathbb{P}\Big[\bigcup_{j=1}^{M}\{(\boldsymbol{x}, \widehat{\boldsymbol{X}}^{(j)}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})\}\Big] = 1 - (1 - \mathbb{P}[(\boldsymbol{x}, \widehat{\boldsymbol{X}}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})])^M, \tag{4.19}$$

where $\widehat{\boldsymbol{X}} \sim P_{\widehat{\boldsymbol{X}}}$ is independent of $\boldsymbol{x}$. We fix $\epsilon > 0$ and note the following:

- If $\mathbb{P}[(\boldsymbol{x}, \widehat{\boldsymbol{X}}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})] \geq e^{-n(R-\epsilon)}$, then the right-hand side of (4.19) is lower bounded by $1 - (1 - e^{-n(R-\epsilon)})^{e^{nR}} = 1 - ((1 - e^{-n(R-\epsilon)})^{e^{n(R-\epsilon)}})^{e^{n\epsilon}}$, which tends to one faster than exponentially since $(1 - e^{-n(R-\epsilon)})^{e^{n(R-\epsilon)}} \to \frac{1}{e}$.

- If $\mathbb{P}[(\boldsymbol{x}, \widehat{\boldsymbol{X}}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})] \leq e^{-n(R+\epsilon)}$, then the right-hand side of (4.19) is upper bounded by $1 - (1 - e^{-n(R+\epsilon)})^{e^{nR}} = 1 - ((1 - e^{-n(R+\epsilon)})^{e^{n(R+\epsilon)}})^{e^{-n\epsilon}}$, which tends to zero exponentially fast since $(1 - e^{-n(R+\epsilon)})^{e^{n(R+\epsilon)}} \to \frac{1}{e}$, and since for fixed $\alpha > 0$ it holds that $(\frac{1}{\alpha})^z = 1 - z \log \alpha + O(z^2)$ as $z \to 0$.

The second and third claims of the lemma now follow by combining the above calculations with the fact that $\mathbb{P}[(\boldsymbol{x}, \widehat{\boldsymbol{X}}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})]$ behaves as $e^{-nI_{\widetilde{P}}(X;\widehat{X})}$ times a sub-exponential pre-factor (*cf.*, (2.85)), taking the union bound over all possible joint types (of which there are only polynomially many), and noting that ϵ can be arbitrarily small. □

The remainder of the proof of Theorem 4.2 contains rather technical continuity arguments; we only present an outline of the arguments here, and refer the interested reader to [54] for further details.

Letting $\mathcal{S}_n(P_X)$ be the set of *all* joint types of $(\boldsymbol{x}, \widehat{\boldsymbol{x}})$ induced by some $\widehat{\boldsymbol{x}} \in \mathcal{C}_n$, Lemma 4.3 states that with high probability, $\mathcal{S}_n^{(+)}(P_X) \subseteq \mathcal{S}_n(P_X)$ and $\mathcal{S}_n^{(-)}(P_X) \cap \mathcal{S}_n(P_X) = \emptyset$. Moreover, by the first part of the lemma, $\mathcal{S}^{(-)}$ and $\mathcal{S}^{(+)}$ collectively account for all possible joint types except those for which $I_{\widetilde{P}}(X; \widehat{X}) \in (R - \delta', R + \delta')$.

By construction, the encoder maps $\boldsymbol{x}$ to an index $m \in \{1, \ldots, M\}$ such that the corresponding joint type $\widetilde{P}_{X\widehat{X}}$ has the highest possible value of $\mathbb{E}_{\widetilde{P}}[d_0(X, \widehat{X})]$. In addition, we have $Q_{\widehat{X},n} \to Q_{\widehat{X}}$ by definition, and $P_X \to \Pi_X$ by the law of large numbers. Hence, and taking $\delta' \to 0$ in Lemma 4.3, we can deduce that the joint type $\widetilde{P}_{X\widehat{X}}$ of $(\boldsymbol{X}, \widehat{\boldsymbol{X}})$ must be arbitrarily close to a minimizer of $\mathbb{E}_{\widetilde{P}}[d_0(X, \widehat{X})]$ subject to $\widetilde{P}_X = \Pi_X$, $\widetilde{P}_{\widehat{X}} = Q_{\widehat{X}}$, and $I_{\widetilde{P}}(X; \widehat{X}) \leq R$. Notice that these constraints coincide with those in (4.6). In particular, the normalized distortion incurred is $\frac{1}{n} d_1^n(\boldsymbol{x}, \widehat{\boldsymbol{X}}) = \mathbb{E}_{\widetilde{P}}[d_1(X, \widehat{X})]$, and is therefore upper bounded by $\max_{\widetilde{P}_{X\widehat{X}} \in \widetilde{\mathcal{P}}} \mathbb{E}_{\widetilde{P}}[d_1(X, \widehat{X})] + \delta$ for arbitrarily small $\delta > 0$.

The proof is completed by recalling the definition of an achievable (R, D) pair in Definition 4.1, and noting that we have only relied on the high-probability events in Lemma 4.3 as well as $P_X \to \Pi_X$.

4.6 Multi-Letter Improvement and Converse

We saw in Section 4.4 that the upper bound $\overline{D}_1(Q_{\widehat{X}}, R)$ is not tight in general, i.e., it can be strictly higher than the mismatched distortion-rate function $D_1^*(R)$, even after optimizing $Q_{\widehat{X}}$. As with the LM rate in channel coding (*cf.*, Section 2.7), one can attain improved rates by coding over blocks of k symbols. Applying Theorem 4.2 to the source Π_X^k, optimizing the input distribution $Q_{\widehat{X}^k}$ on the product alphabet $\widehat{\mathcal{X}}^k$, and normalizing the rate and the distortion incurred by k, we obtain an achievability result for the original source. Specifically, letting $\overline{D}_1(Q_{\widehat{X}}, R, \Pi_X)$ denote (4.5) with an explicit dependence on the source, the achievable distortion is given by

$$\overline{D}_1^{(k)}(R) \triangleq \min_{Q_{\widehat{X}^k}} \frac{1}{k} \overline{D}_1(Q_{\widehat{X}^k}, kR, \Pi_X^k). \tag{4.20}$$

Since the choice of k is arbitrary, it follows that the distortion

$$\overline{D}_1^{(\infty)}(R) \triangleq \inf_{k \in \mathbb{Z}} \overline{D}_1^{(k)}(R) \tag{4.21}$$

is also achievable at rate R.

While the tightness of the product extension of the LM rate for channel coding (i.e., the Csiszár–Narayan conjecture) remains an open problem, the following theorem reveals that the analogous statement is indeed true for the mismatched rate-distortion problem, at least up to tie-breaking issues [54]. Specifically, it is true when we assume a *pessimistic* view of tie-breaking: If multiple codewords achieve the minimum in (4.1), then the selected codeword is the one with the highest $d_1^n(\boldsymbol{x}, \widehat{\boldsymbol{x}})$. If there is still a tie among these, then it makes no difference which one is chosen.

Theorem 4.4 (Multi-Letter Converse)**.** Under the mismatched rate-distortion problem specified by (Π, d_0, d_1), under pessimistic tie-breaking, the mismatched distortion-rate function satisfies

$$D_1^*(R) = \overline{D}_1^{(\infty)}(R). \tag{4.22}$$

Proof. We have already established that the distortion $\overline{D}_1^{(\infty)}(R) = \inf_{k \in \mathbb{Z}} \overline{D}_1^{(k)}(R)$ is achievable, i.e., it is an upper bound on $D_1^*(R)$. Since

the distortion-rate function is continuous in R by Lemma 4.1, it suffices to show that for any $\delta > 0$, we can find an integer k such that $\overline{D}_1^{(k)}(R+\delta) \le D_1^*(R) + c\delta$, where $c > 0$ is a constant not depending on δ. From (4.20), such a statement is equivalent to showing that we can find an auxiliary distribution $Q_{\widehat{X}^k}$ such that

$$\frac{1}{k}\overline{D}_1(Q_{\widehat{X}^k}, k(R+\delta), \Pi_X^k) \le D_1^*(R) + c\delta. \tag{4.23}$$

By the definition of the mismatched distortion-rate function $D_1^*(R)$ (Definition 4.1), for any $\delta > 0$, there exists a sufficiently large block length k and a codebook $\mathcal{C}_k = \{\widehat{\boldsymbol{x}}^{(1)}, \ldots, \widehat{\boldsymbol{x}}^{(M)}\}$ with $M \le e^{k(R+\delta)}$, such that $\frac{1}{k} d_1^k(\boldsymbol{X}, \widehat{\boldsymbol{X}}) \le D_1^*(R) + \delta$ with probability at least $1-\delta$,[3] where $\boldsymbol{X} \sim \Pi_X^k$ and $\widehat{\boldsymbol{X}}$ is the resulting estimate. Recalling the definition $d_{1,\max} = \max_{x,\widehat{x}} d_1(x,\widehat{x})$ and the non-negativity of d_1, we deduce that

$$\frac{1}{k}\mathbb{E}[d_1^k(\boldsymbol{X}, \widehat{\boldsymbol{X}})] \le D_1^*(R) + \delta(1 + d_{1,\max}), \tag{4.24}$$

which follows by upper bounding the distortion by $d_{1,\max}$ in the case that $\frac{1}{k} d_1^k(\boldsymbol{X}, \widehat{\boldsymbol{X}}) > D_1^*(R) + \delta$.

In the following, we denote the mapping from source sequences to compressed sequences under the codebook $\mathcal{C}_k$ as $\phi^*(\cdot)$, i.e., $\widehat{\boldsymbol{X}} = \phi^*(\boldsymbol{X})$. In addition, we let $Q_{\widehat{X}^k}$ be the resulting marginal distribution of $\widehat{\boldsymbol{X}} \in \widehat{\mathcal{X}}^k$, corresponding to randomly generating $\boldsymbol{X} \sim \Pi_X^k$ and then setting $\widehat{\boldsymbol{X}} = \phi^*(X^k)$.

We now consider the achievable distortion level of Theorem 4.2 applied to the multi-letter source Π_X^k:

$$\overline{D}_1(Q_{\widehat{X}^k}, k(R+\delta), \Pi_X^k) = \max_{\widetilde{P}_{X^k\widehat{X}^k} \in \widetilde{\mathcal{P}}_k} \mathbb{E}_{\widetilde{P}}[d_1^k(X^k, \widehat{X}^k)], \tag{4.25}$$

where, in accordance with (4.6), we have

$$\widetilde{\mathcal{P}}_k = \Big\{ \widetilde{P}_{X^k\widehat{X}^k} \colon \widetilde{P}_{X^k\widehat{X}^k} \in \underset{\substack{\widetilde{P}_{X^k\widehat{X}^k} \colon \widetilde{P}_{X^k} = \Pi_X^k, \widetilde{P}_{\widehat{X}^k} = Q_{\widehat{X}^k}, \\ I_{\widetilde{P}}(X^k;\widehat{X}^k) \le k(R+\delta)}}{\arg\min} \mathbb{E}_{\widetilde{P}}[d_0(X, \widehat{X})] \Big\}. \tag{4.26}$$

[3] Here we use k instead of n for the block length to highlight the fact that we consider $\overline{D}_1^{(k)}(R)$ in (4.20) and choose $Q_{\widehat{X}^k}$ depending on $\mathcal{C}_k$.

Note that the $\widehat{X}$-marginal constraint ensures that $P_{\widehat{X}^k}$ takes mass only on the elements of $\mathcal{C}_k$, of which there are at most $e^{k(R+\delta)}$. It follows that $I_{\widetilde{P}}(X^k;\widehat{X}^k) \leq H_{\widetilde{P}}(\widehat{X}^k) \leq k(R+\delta)$, and hence the rate constraint is redundant (i.e., it is automatically satisfied).

Since $\phi^*(x^k)$ is a mismatched encoding rule minimizing d_0, the joint distribution resulting from $\widetilde{P}_{\widehat{X}^k|X^k}(\widehat{x}^k|x^k) = \mathbb{1}\{\widehat{x}^k = \phi^*(x^k)\}$ must be in the set $\widetilde{\mathcal{P}}_k$. Moreover, among *all* conditional distributions $\widetilde{P}_{\widehat{X}^k|X^k}$ producing the required marginal $Q_{\widehat{X}^k}$ supported on $\mathcal{C}_k$, it is $\phi^*(x^k)$ that achieves the minimum in (4.25). Indeed, this follows directly from our assumption of pessimistic tie-breaking – if two codewords achieve the same d_0, then the one with the highest d_1 is selected under ϕ^*.

As a result, using the bound on the average distortion under $\mathcal{C}_k$ in (4.24), we deduce from (4.25) that

$$\frac{1}{k}\overline{D}_1(\Pi_X^k, Q_{X^k}, k(R+\delta)) \leq D_1^*(R) + \delta(1 + d_{1,\max}). \tag{4.27}$$

We have therefore established (4.23) (with $c = 1 + d_{1,\max}$), as required. □

4.7 Mismatched Random Codebooks with Optimal Encoding

The preceding subsections considered the mismatched encoding problem in which the encoder chooses the minimum-distortion codeword with respect to d_0, but the actual distortion measure is d_1. In this subsection, we turn to a different form of mismatch in which a common distortion measure d is adopted throughout the system, but a suboptimal random coding distribution is used to generate the codebook. We are interested in the rate-distortion trade-off for the given random coding ensemble. This problem was studied in the case of finite alphabets in [106], and more general alphabets in [107]. We refer the reader to [28] for a more detailed survey, and discuss some other notions of mismatch and universality in Section 4.7.4.

We focus primarily on the finite-alphabet setting in this subsection, while occasionally providing references to more general results. In addition, Section 4.8 will consider a specific case of significant interest with

continuous alphabets, namely, Gaussian codebooks for non-Gaussian sources.

4.7.1 Achievable Rate-Distortion Functions

We consider the setup of Section 4.2 with a common distortion measure $d_0 = d_1 = d$, assumed to be non-negative and finite-valued. In this case, the encoder in (4.1) becomes the optimal minimum-distortion rule, which is repeated as follows for convenience:

$$m = \operatorname*{arg\,min}_{j=1,\ldots,M} d^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}^{(j)}), \quad d^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = \sum_{i=1}^{n} d(x_i, \widehat{x}_i). \tag{4.28}$$

Here the distortion incurred is identical for all tie-breaking strategies.

If the codebook is optimized in accordance with Definition 4.1, then the rate-distortion trade-off is characterized by the classical solution of Shannon [89], stated as follows:

$$R^*_{\text{Matched}}(D) = \min_{\substack{\widetilde{P}_{X\widehat{X}}:\ \widetilde{P}_X = \Pi_X \\ \mathbb{E}_{\widetilde{P}}[d(X,\widehat{X})] \le D}} I_{\widetilde{P}}(X;\widehat{X}). \tag{4.29}$$

Here and subsequently, we find it more convenient to consider the rate as a function of the distortion, but the distortion as a function of the rate follows by a simple inversion. For instance, (4.29) yields the following distortion-rate function:

$$D^*_{\text{Matched}}(R) = \min_{\substack{\widetilde{P}_{X\widehat{X}}:\ \widetilde{P}_X = \Pi_X \\ I_{\widetilde{P}}(X;\widehat{X}) \le R}} \mathbb{E}_{\widetilde{P}}[d(X,\widehat{X})]. \tag{4.30}$$

In analogy with the channel coding results surveyed in Section 2, we consider two random-coding ensembles, both of which consist of generating a codebook $\mathcal{C} = \{\widehat{\boldsymbol{X}}^{(1)}, \ldots, \widehat{\boldsymbol{X}}^{(M)}\}$ containing codewords drawn independently from some distribution $P_{\widehat{\boldsymbol{X}}}$. Under the i.i.d. ensemble, we set $P_{\widehat{\boldsymbol{X}}}(\widehat{\boldsymbol{x}}) = Q^n_{\widehat{X}}$, and under the constant-composition ensemble, $P_{\widehat{\boldsymbol{X}}}$ is as given in (4.4). Similarly to Definition 4.1, we are interested in the random-coding error probability

$$\bar{p}_{\text{e}}(n, M, D, Q_{\widehat{X}}, \delta) \triangleq \mathbb{P}\left[\frac{1}{n} d^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}) \ge D + \delta\right], \tag{4.31}$$

where the average is with respect to both $\boldsymbol{X} \sim \Pi_X^n$ and the random codebook of size M. We say that a rate-distortion pair (R, D) is achievable if $\bar{p}_{\mathrm{e}}(n, \lfloor e^{n(R+\delta)} \rfloor, D, Q_{\widehat{X}}, \delta) \to 0$ for any fixed $\delta > 0$.

The achievable rate-distortion functions for these ensembles are introduced as follows, and will be accompanied by a formal theorem below. For the i.i.d. ensemble, the primal form is given by

$$\overline{R}_{\mathrm{iid}}(D, Q_{\widehat{X}}) = \min_{\substack{\widetilde{P}_{X\widehat{X}}: \widetilde{P}_X = \Pi_X, \\ \mathbb{E}_{\widetilde{P}}[d(X,\widehat{X})] \le D}} D(\widetilde{P}_{X\widehat{X}} \| \Pi_X \times Q_{\widehat{X}}), \tag{4.32}$$

and the dual form is given by

$$\overline{R}_{\mathrm{iid}}(D, Q_{\widehat{X}}) = \sup_{s \ge 0} -\sum_x \Pi_X(x) \log \sum_{\widehat{x}} Q_{\widehat{X}}(\widehat{x}) e^{s(D - d(x,\widehat{x}))}. \tag{4.33}$$

Similarly, for the constant-composition ensemble, the primal form is given by

$$\overline{R}_{\mathrm{cc}}(D, Q_{\widehat{X}}) = \min_{\substack{\widetilde{P}_{X\widehat{X}}: \ \widetilde{P}_X = \Pi_X, \widetilde{P}_{\widehat{X}} = Q_{\widehat{X}}, \\ \mathbb{E}_{\widetilde{P}}[d(X,\widehat{X})] \le D}} I_{\widetilde{P}}(X; \widehat{X}), \tag{4.34}$$

and the dual form is given by

$$\overline{R}_{\mathrm{cc}}(D, Q_{\widehat{X}}) = \sup_{s \ge 0, b(\cdot)} -\sum_x \Pi_X(x) \log \sum_{\widehat{x}} Q_{\widehat{X}}(\widehat{x}) e^{s(D - d(x,\widehat{x}))} e^{b(\widehat{x}) - \phi_b} \tag{4.35}$$

with $\phi_b = \mathbb{E}_Q[b(\widehat{X})]$. A proof of the equivalence of (4.32)–(4.33) can be found in [28], and that of (4.34)–(4.35) can be proved in the same way as the LM rate for channel coding.

Interestingly, there is a close connection between these expressions and those of the GMI and LM rate: From (4.32) and (4.34), we obtain the primal channel coding rates upon identifying $(X, \widehat{X}) \leftrightarrow (Y, X)$, $d(x, \widehat{x}) \leftrightarrow -\log q(x, y)$, and $D \leftrightarrow \mathbb{E}_{Q_X \times W}[-\log q(X, Y)]$. In addition, we have the following analog of the fact that the LM rate is at least as high as the GMI:

$$\overline{R}_{\mathrm{iid}}(D, Q_{\widehat{X}}) \le \overline{R}_{\mathrm{cc}}(D, Q_{\widehat{X}}). \tag{4.36}$$

It is important, however, to remember that here we would like the rate to be *as small as possible* for a given distortion level, in stark contrast

with the channel coding setup in which high rates are desired. Thus, the ordering (4.36) in fact implies that i.i.d. random coding is *preferable* to constant-composition random coding for a given choice of $Q_{\widehat{X}}$. While strict inequality is possible in (4.36), equality holds upon minimizing both sides over $Q_{\widehat{X}}$ (see the discussion following (4.7)).

The intuition as to why i.i.d. random coding may be preferable is as follows. Attaining $d^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}) < n(D+\delta)$ is equivalent to the existence of a codeword $\widehat{\boldsymbol{X}}^{(j)}$ such that $(\boldsymbol{X}, \widehat{\boldsymbol{X}}^{(j)})$ induces a joint type $\widetilde{P}_{X\widehat{X}}$ satisfying $\mathbb{E}_{\widetilde{P}}[d(X,\widehat{X})] < D+\delta$. By a typicality argument, we may assume that the X-marginal of $\widetilde{P}_{X\widehat{X}}$ is arbitrarily close to Π_X. For the $\widehat{X}$-marginal, we consider two separate cases:

- Suppose that the $\widehat{X}$-marginal of $\widetilde{P}_{X\widehat{X}}$ is $Q_{\widehat{X}}$ (or more generally, the type $Q_{\widehat{X},n}$ approximating $Q_{\widehat{X}}$). Then, by the properties of types in Section 2.6.1, the i.i.d. and constant-composition codeword distributions yield matching behavior (on an exponential scale) in the probability of the event $(\boldsymbol{X}, \widehat{\boldsymbol{X}}^{(j)}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})$. Note that $I_{\widetilde{P}}(X;\widehat{X}) = D(\widetilde{P}_{X\widehat{X}} \| \Pi_X \times Q_{\widehat{X}})$ in this case.

- On the other hand, if the $\widehat{X}$-marginal of $\widetilde{P}_{X\widehat{X}}$ differs from $Q_{\widehat{X}}$, then the event $(\boldsymbol{X}, \widehat{\boldsymbol{X}}^{(j)}) \in \mathcal{T}^n(\widetilde{P}_{X\widehat{X}})$ never occurs under the constant-composition ensemble. However, it may still occur under the i.i.d. ensemble, with an associated per-codeword probability of roughly $e^{-nD(\widetilde{P}_{X\widehat{X}} \| \Pi_X \times Q_{\widehat{X}})}$.

Thus, high-probability success under the constant-composition ensemble essentially implies the same under the i.i.d. ensemble, but the opposite is not true in general. In this sense, the i.i.d. ensemble produces a *more diverse* codebook that improves the chances of a low-distortion codeword existing. Analogous observations were made for the error exponents of Gaussian codebooks in [111].

In light of this discussion, it is natural to ask why constant-composition coding was used in the proof of Theorem 4.2. Firstly, we note that the distribution $Q_{\widehat{X}}$ therein is optimized, rather than being fixed and possibly suboptimal. In addition, under the setup therein with two distortion measures d_0 and d_1, an additional subtlety arises: A more

diverse codebook may amount to more ways in which the encoder can be led astray due to the mismatch. In fact, in the parallel source example of Section 4.4.1 is an example in which additional codebook structure provably eliminates the performance loss due to mismatch.

4.7.2 Statement of Results

In the following, we provide a formal statement of achievability and ensemble tightness for the above rate-distortion functions $\overline{R}_{\mathrm{iid}}$ and $\overline{R}_{\mathrm{cc}}$. To avoid trivial cases, we consider technical assumptions stated in terms of the following definitions:

$$D_{\min} = \sum_{x} \Pi_X(x) \min_{\widehat{x}:\ Q_{\widehat{X}}(\widehat{x})>0} d(x, \widehat{x}), \tag{4.37}$$

$$D_{\mathrm{prod}} = \sum_{x,\widehat{x}} \Pi_X(x) Q_{\widehat{X}}(\widehat{x}) d(x, \widehat{x}). \tag{4.38}$$

We assume that $D \geq D_{\min}$, which is justified by the fact that $D_{\min}$ is the average distortion attained when each source symbol x_i is paired with the reconstructed symbol $\widehat{x}_i$ having the lowest possible distortion. In addition, we can assume without loss of generality that $D < D_{\mathrm{prod}}$, since a distortion of at most $D_{\mathrm{prod}} + \delta$ is trivially attained with high probability using a "codebook" with just a single codeword drawn from $Q_{\widehat{X}}^n$, due to the law of large numbers.

Theorem 4.5 (Achievable Rate-Distortion Functions with Mismatched Random Coding)**.** For any discrete memoryless source Π_X, distortion function d, distortion level $D \in [D_{\min}, D_{\mathrm{prod}})$, auxiliary distribution $Q_{\widehat{X}} \in \mathcal{P}(\widehat{\mathcal{X}})$, and parameter $\delta > 0$, the random-coding error probability (4.31) satisfies the following:

- Under the i.i.d. ensemble, we have $\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor, D, Q_{\widehat{X}}, \delta) \to 0$ for any $R > \overline{R}_{\mathrm{iid}}(D, Q_{\widehat{X}})$ and $\delta > 0$, and $\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor, D, Q_{\widehat{X}}, \delta) \to 1$ for any $R < \overline{R}_{\mathrm{iid}}(D, Q_{\widehat{X}})$ and sufficiently small $\delta > 0$.
- Under the constant-composition ensemble, we have $\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor, D, Q_{\widehat{X}}, \delta) \to 0$ for any $R > \overline{R}_{\mathrm{cc}}(D, Q_{\widehat{X}})$ and $\delta > 0$, and $\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor, D, Q_{\widehat{X}}, \delta) \to 1$ for any $R < \overline{R}_{\mathrm{cc}}(D, Q_{\widehat{X}})$ and sufficiently small $\delta > 0$.

Remark 4.2. As we will exemplify in Section 4.7.3, the constraint set in (4.34) may be empty, in which case we adopt the convention $\overline{R}_{\mathrm{cc}}(D, Q_{\widehat{X}}) = \infty$, meaning that the condition $R > \overline{R}_{\mathrm{cc}}(D, Q_{\widehat{X}})$ cannot be satisfied for any rate R, and no achievability claim is made for the given pair $(D, Q_{\widehat{X}})$. Specifically, this occurs whenever

$$D < \min_{\widetilde{P}_{X\widehat{X}}:\ \widetilde{P}_X = \Pi_X, \widetilde{P}_{\widehat{X}} = Q_{\widehat{X}}} \mathbb{E}_{\widetilde{P}}[d(X, \widehat{X})], \tag{4.39}$$

and the right-hand side may be strictly larger than $D_{\min}$ in (4.37). The reason for this phenomenon is that under a suboptimal choice of $Q_{\widehat{X}}$, a given typical sequence $\boldsymbol{X}$ may incur significant distortion with *every* sequence that has composition $Q_{\widehat{X}}$. In contrast, under the i.i.d. ensemble, although most codewords have composition close to $Q_{\widehat{X}}$, many codewords also have significantly different compositions.

The results of Theorem 4.5 can be proved in several ways. For the constant-composition ensemble, the achievability part is in fact a special case of Theorem 4.2, obtained by setting $d_0 = d_1$. This gives the distortion-rate function in (4.7),[4] which can be inverted to obtain the primal rate-distortion expression in (4.34). The ensemble tightness claim readily follows since the proof in Section 4.5 is based on exact error probability expressions and the method of types.

The i.i.d. ensemble can be analyzed similarly using types. Alternatively, an elementary achievability proof that avoids the use of types is possible via the techniques of Gallager [38, Lemma 9.3.1] and Sakrison [73]. To further highlight the connections between mismatched channel coding and rate-distortion theory, we instead outline an approach via a large-deviations result of [28], which we also used in Section 3.7 to prove the ensemble tightness of the GMI for channel coding.

We first form an exact expression for the random-coding error probability under the i.i.d. ensemble as follows (see also [52, Thm. 9]): Defining

[4] The subscript "Matched" in (4.7) refers to having $d_0 = d_1$, but the choice of $Q_{\widehat{X}}$ may still be mismatched in the sense of being fixed and suboptimal.

$(\boldsymbol{X}, \widehat{\boldsymbol{X}}') \sim \Pi_X^n(\boldsymbol{x}) Q_{\widehat{X}}^n(\widehat{\boldsymbol{x}})$, we have

$$\bar{p}_{\mathrm{e}} = \mathbb{P}\Big[\bigcap_{j=1,\ldots,M} \Big\{\frac{1}{n} d^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}^{(j)}) \geq D + \delta\Big\}\Big] \tag{4.40}$$

$$= \mathbb{E}\Big[\mathbb{P}\Big[\bigcap_{j=1,\ldots,M} \Big\{\frac{1}{n} d^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}^{(j)}) \geq D + \delta\Big\} \,\Big|\, \boldsymbol{X}\Big]\Big] \tag{4.41}$$

$$= \mathbb{E}\Big[\Big(\mathbb{P}\Big[\frac{1}{n} d^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}') \geq D + \delta \,\Big|\, \boldsymbol{X}\Big]\Big)^M\Big] \tag{4.42}$$

$$= \mathbb{E}\Big[\Big(1 - \mathbb{P}\Big[\frac{1}{n} d^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}') < D + \delta \,\Big|\, \boldsymbol{X}\Big]\Big)^M\Big], \tag{4.43}$$

where (4.42) uses the independence of the random codewords. Observe that (4.43) bears a strong resemblance to the channel coding counterpart in (3.34).

One could proceed by characterizing the inner probability in (4.43) using the method of types, via similar steps to Section 4.5. One may also be tempted to follow the dual analysis of the GMI for channel coding in Section 2.6.4; however, the situation is different here, as we require a *lower bound* on the tail event in (4.43) to upper bound $\bar{p}_{\mathrm{e}}$, whereas the techniques of Section 2.6.4 (e.g., Markov's inequality) lead to upper bounds on the tail event.[5] Hence, we instead consider the use of the following large deviations result [28, Thm. 1], from which both the achievability of $\overline{R}_{\mathrm{iid}}$ and ensemble tightness readily follow. We state this result in a general form that permits its use even in the continuous-alphabet setting.

Lemma 4.6 (Large Deviations Result for i.i.d. Random Coding)**.** Fix a memoryless source Π_X, an auxiliary distribution $Q_{\widehat{X}}$, and a non-negative distortion measure $d(x, \widehat{x})$, and define[6]

$$D_{\min} = \mathbb{E}_{\Pi_X}[\operatorname{ess\,inf}_{Q_{\widehat{X}}} d(X, \widehat{X})], \tag{4.44}$$

$$D_{\mathrm{prod}} = \mathbb{E}_{\Pi_X \times Q_{\widehat{X}}}[d(X, \widehat{X})]. \tag{4.45}$$

[5] Note also that $d(x, \widehat{x})$ here plays the role of $-\log q(x, y)$ in channel coding, so the inner probability in (4.43) considers a lower tail event, instead of an upper tail event.

[6] As stated in Footnote 1 in Section 3.7, the *essential infimum* of a function $g(\widehat{X})$ with respect to $\widehat{X} \sim Q_{\widehat{X}}$ is defined to be supremum of $t \in \mathbb{R}$ for which $\mathbb{P}_Q[g(\widehat{X}) > t] = 1$.

Then, if $D_{\text{prod}} < \infty$, we have for any $D \in (D_{\min}, D_{\text{prod}})$ that the following holds with probability one with respect to $\boldsymbol{X}$:

$$\lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}[\log d^n(\boldsymbol{X}, \widehat{\boldsymbol{X}}') \geq nD | \boldsymbol{X}] = \overline{R}_{\text{iid}}(D, Q_{\widehat{X}}), \tag{4.46}$$

where $(\boldsymbol{X}, \widehat{\boldsymbol{X}}') \sim \Pi_X^n(\boldsymbol{x}) Q_{\widehat{X}}^n(\widehat{\boldsymbol{x}})$.

With this result in place, the desired claim in Theorem 4.5 for $D \in (D_{\min}, D_{\text{prod}})$ follows from (4.43) (with $M = \lfloor e^{nR} \rfloor$) and the fact that $(1 - \frac{1}{\alpha_n})^{\beta_n} \to 0$ whenever $\alpha_n \geq 1$ and $\lim_{n\to\infty} \frac{\alpha_n}{\beta_n} = 0$ (achievability) and $(1 - \frac{1}{\alpha_n})^{\beta_n} \to 1$ whenever $\alpha_n \geq 1$ and $\lim_{n\to\infty} \frac{\alpha_n}{\beta_n} = \infty$ (ensemble tightness).

The assumption $D \in (D_{\min}, D_{\text{prod}})$ in Lemma 4.6 is motivated in the same way as the discussion following (4.38), except that for general alphabets, strict inequality is assumed in the condition $D > D_{\min}$ in order to avoid technical issues when $D = D_{\min}$. In contrast, in the finite-alphabet setting, the case $D = D_{\min}$ simply amounts to lossless compression, which can easily be handled separately. Further discussion can be found in the text following [28, Thm. 1].

4.7.3 Example: Bernoulli Source

As a simple example illustrating the preceding achievable rates, we consider the source $X \sim \text{Bernoulli}(p)$ with $p \leq \frac{1}{2}$, and let $d(x, \widehat{x}) = \mathbb{1}\{\widehat{x} \neq x\}$ be the Hamming distortion measure. It is well-known that the matched rate-distortion function is [22, Sec. 10.3.1]

$$R^*_{\text{Matched}}(D) = [H_2(p) - H_2(D)]_+, \tag{4.47}$$

where $[\alpha]_+ = \max\{0, \alpha\}$. In addition, the optimal auxiliary distribution $Q_{\widehat{X}}$ is Bernoulli(q^*), with q^* chosen according to one of two cases: (i) if $0 \leq D \leq p$, then $q^* = \frac{p-D}{1-2D}$; (ii) if $D > p$, then $q^* = 0$.

In Figure 4.3, we compare R^*_{Matched} to $\overline{R}_{\text{iid}}$ and $\overline{R}_{\text{cc}}$ in the case that $p = \frac{1}{3}$, and $Q_{\widehat{X}}$ is chosen to be Bernoulli($\frac{1}{5}$). We observe that $\overline{R}_{\text{iid}}$ incurs a small loss compared to R^*_{Matched} across the entire range of distortion levels. While $\overline{R}_{\text{cc}}$ behaves similarly to $\overline{R}_{\text{iid}}$ at high distortion levels, it is unable to attain any distortion level below a threshold of roughly 0.13, regardless of a rate; this is due to the phenomenon discussed

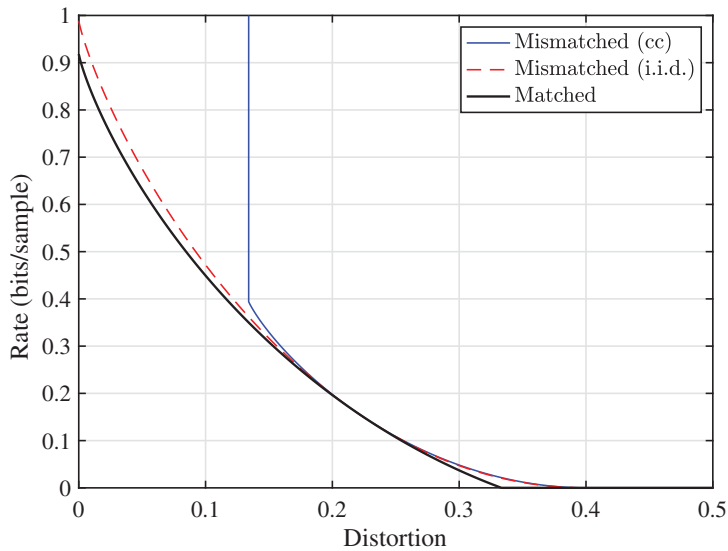


Figure 4.3: Bernoulli source example: Matched and mismatched rate-distortion curves with $X \sim \text{Bernoulli}(\frac{1}{3})$ and the Hamming distortion measure. In the mismatched case, the auxiliary distribution $Q_{\widehat{X}}$ is chosen to be $\text{Bernoulli}(\frac{1}{5})$.

in Remark 4.2. Finally, we note that all three curves coincide when $D = \frac{2}{9} \approx 0.22$; this is due to the fact that the optimal parameter $q^* = \frac{p-D}{1-2D}$ mentioned above equals $\frac{1}{5}$ under this choice, and thus coincides with the choice of $Q_{\widehat{X}}$.

4.7.4 Further Related Work

We conclude this subsection by briefly outlining some other relevant notions of mismatch and universality in the literature on rate-distortion theory. This outline is far from exhaustive, and further references can be found in [28, 50].

We first briefly mention that the techniques and results surveyed in this subsection are related to (and in some cases directly used in) studies of universal rate-distortion codes [28, 50, 98, 110]. Along similar lines, the early works of Sakrison [73, 74] introduced the problem of constructing rate-distortion codes that simultaneously achieve low distortion for a class of sources, as a natural analog of the compound channel in channel

coding [16]. This setup is also closely related to the use of Gaussian codes for non-Gaussian sources [54], which we survey in Section 4.8.

In other works such as [42, 43], the goal is also to characterize the degradation in performance when a rate-distortion codebook is designed for some source Π_X but then applied to a different source Π'_X. This is analogous to the problem considered in this subsection, but is studied in [42, 43] beyond the random coding setting, with various results including generalizations of the well-known loss of $D(\Pi_X \| \Pi'_X)$ for Huffman coding [22, Sec. 5.6].

4.8 Gaussian Compression for Non-Gaussian Sources

In the previous subsection, we focused on discrete memoryless sources with a mismatched random coding distribution. In this section, we turn to a particularly important special case of the continuous-alphabet counterpart of this problem, namely, Gaussian coding for non-Gaussian sources under the squared-error distortion criterion [54]. This can be viewed as an analogous problem to that of mismatched Gaussian codes for channel coding (*cf.*, Section 3.8.2). Before formally describing the problem setting, we provide some background on rate-distortion theory with Gaussian sources.

4.8.1 Background

The compression of Gaussian sources subject to a squared-error distortion constraint is one of the most widely-considered rate-distortion settings. In this setting, the source $\boldsymbol{X}$ is assumed to be i.i.d. on $\mathrm{N}(0, \sigma^2)$ for some $\sigma^2 > 0$, and the goal is to construct an encoder (mapping $\mathbb{R}^n$ to $\{1, \ldots, M\}$) and decoder (mapping $\{1, \ldots, M\}$ to $\mathbb{R}^n$) such that $d^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = \|\boldsymbol{x} - \widehat{\boldsymbol{x}}\|_2^2$ is small, where $\widehat{\boldsymbol{x}}$ is the reconstruction.

Under the matched counterpart of Definition 4.1 with $d_0 = d_1 = d$, the rate-distortion function is given by [22, Sec. 10.3.2]

$$R^*_{\mathrm{Matched}}(D) = \Big[\frac{1}{2} \log \frac{\sigma^2}{D}\Big]_+, \tag{4.48}$$

where $[\alpha]_+ = \max\{0, \alpha\}$. The achievability part can be proved using Gaussian codebooks (to be made precise below), along with an encoder

that maps a given source sequence $\boldsymbol{x} \in \mathbb{R}^n$ to the nearest codeword in the codebook:

$$\widehat{\boldsymbol{x}} = \operatorname*{arg\,min}_{j=1,\ldots,M} \|\boldsymbol{x} - \widehat{\boldsymbol{x}}^{(j)}\|_2^2. \tag{4.49}$$

4.8.2 Non-Gaussian Sources

Suppose that we have an arbitrary memoryless source on $\mathbb{R}$, i.e., $\boldsymbol{X} \sim \Pi_X^n$ for some Π_X. For concreteness, we assume that the underlying distribution is continuous, so that Π_X is a density function. Continuing with the assumption of squared-error distortion $d(x, \widehat{x}) = (x - \widehat{x})^2$, the matched rate-distortion function in (4.29) specializes as follows:

$$R^*_{\text{Matched}}(D) = \min_{\substack{\widetilde{P}_{X\widehat{X}}:\ \widetilde{P}_X = \Pi_X, \\ \mathbb{E}_{\widetilde{P}}[(X-\widehat{X})^2] \le D}} I(X; \widehat{X}). \tag{4.50}$$

However, achieving this rate may be difficult – the source distribution may be unknown, and even if it is known, the required codebook structure may be prohibitively complex.

In this subsection, we address these difficulties by considering the question of how well a Gaussian codebook performs when applied to a non-Gaussian source. Since we are considering the squared-error distortion $d(x, \widehat{x}) = (x - \widehat{x})^2$, we still adopt the encoding rule in (4.49). We only assume that the source has a fixed second moment:

$$\mathbb{E}_\Pi[X^2] = \sigma^2. \tag{4.51}$$

One may expect a rate increase due to the suboptimality of the codebook construction, but on the other hand, one may expect a rate decrease due to the fact that Gaussian sources are the hardest to compress for a given second moment under the squared-error distortion criterion [22, Ex. 10.8]. Analogously to mismatched channel coding with non-Gaussian noise (*cf.*, Section 3.8.2), we will see that these effects cancel each other, and all sources lead to the same rate-distortion trade-off as the Gaussian source.

While our presentation follows that of Lapidoth [54], an achievability result was also given in the early work of Sakrison [73] under the additional technical assumption that the $(2 + \epsilon)$-th moment of X exists for some $\epsilon > 0$.

4.8.3 Gaussian Codebooks

We consider two different notions of Gaussian codebooks, which we refer to as the *Gaussian i.i.d. ensemble* and the *shell ensemble*. Fixing $\Gamma > 0$, the Gaussian i.i.d. ensemble considers codewords that are independently distributed according to

$$\widehat{\boldsymbol{X}} \sim \mathrm{N}(\mathbf{0}, \Gamma \boldsymbol{I}_n) \tag{4.52}$$

where $\mathbf{0}$ is the zero vector and $\boldsymbol{I}_n$ is the $n \times n$ identity matrix. In contrast, the shell ensemble considers codewords that are independently distributed according to

$$\widehat{\boldsymbol{X}} \sim \mathrm{Uniform}(S_n(\Gamma)), \tag{4.53}$$

where $S_n(\Gamma) = \{\widehat{\boldsymbol{x}} \in \mathbb{R}^n \colon \|\widehat{\boldsymbol{x}}\|_2^2 = n\Gamma\}$ is the shell of radius $\sqrt{n\Gamma}$ in n-dimensional space.

The parameter Γ represents the "power" of codewords in the codebook, but in contrast with channel coding, this is not constrained as part of the problem statement. The optimal choice will turn out to be $\Gamma = \sigma^2 - D$ whenever $D \leq \sigma^2$, whereas the case $D > \sigma^2$ trivially yields a rate of zero by always outputting the all-zero sequence.

4.8.4 Achievability Result

The above-outlined achievability result is formally stated as follows.

Theorem 4.7 (Gaussian Codes for Non-Gaussian Sources)**.** For any memoryless source Π_X with $\mathbb{E}_\Pi[X^2] = \sigma^2$, under the Gaussian i.i.d. ensemble or shell ensemble with parameter $\Gamma > 0$, along with nearest-neighbor encoding according to (4.49), we have the following for any pair (R, D) and $\delta > 0$:

- If $\Gamma = [\sigma^2 - D]_+$ and $R \geq [\frac{1}{2} \log \frac{\sigma^2}{D}]_+ + \delta$, then as $n \to \infty$, we have
$$\mathbb{P}[\|\boldsymbol{X} - \widehat{\boldsymbol{X}}\|_2^2 \geq n(D + \delta)] \to 0. \tag{4.54}$$
- For any choice of Γ, if $R \leq [\frac{1}{2} \log \frac{\sigma^2}{D}]_+ - \delta$, then as $n \to \infty$, we have
$$\mathbb{P}[\|\boldsymbol{X} - \widehat{\boldsymbol{X}}\|_2^2 \geq n(D - \delta)] \to 1. \tag{4.55}$$

As with Theorem 3.9, this result can be viewed both positively and negatively. On the positive side, we see that Gaussian codes form a compression scheme that is robust in handling non-Gaussian sources. On the other hand, Gaussian sources are the hardest to compress for a given second moment [22, Ex. 10.8], and Theorem 3.9 states that we always attain this worst-case performance and no better.

The proof of Theorem 3.9 is outlined as follows; further details can be found in [54]:

- By the circular symmetry of the distributions in (4.52)–(4.53), and similar symmetry in the nearest-neighbor encoding rule (4.49), the conditional error probability $\mathbb{P}[\text{error}|\boldsymbol{x}]$ depends on the source realization $\boldsymbol{x}$ only through $\|\boldsymbol{x}\|_2^2$.

- From existing analyses of Gaussian codes for Gaussian sources, one can deduce that (i) $\mathbb{P}[\text{error}|\boldsymbol{x}] \to 0$ as $n \to \infty$ with $\frac{1}{n}\|\boldsymbol{x}\|_2^2 \to \gamma$ for any γ such that $R > [\frac{1}{2}\log\frac{\gamma}{D}]_+$; (ii) $\mathbb{P}[\text{error}|\boldsymbol{x}] \to 1$ as $n \to \infty$ with $\frac{1}{n}\|\boldsymbol{x}\|_2^2 \to \gamma$ for any γ such that $R < [\frac{1}{2}\log\frac{\gamma}{D}]_+$.

- Hence, the theorem follows from the fact that, for any memoryless source with $\mathbb{E}[X^2] = \sigma^2$, the law of large numbers implies that $\frac{1}{n}\|\boldsymbol{X}\|_2^2$ converges in probability to σ^2.

A complete proof can be found in [54], and an alternative proof can be found in [111] that also leads to refined asymptotic bounds. For the i.i.d. ensemble, it can also be shown that $\overline{R}_{\text{iid}}$ in (4.32) evaluates to $[\frac{1}{2}\log\frac{\sigma^2}{D}]_+$ when $Q_{\widehat{X}}$ is chosen as $\mathrm{N}(0, [\sigma^2 - D]_+)$ [28, Examples 1–2], meaning that (4.54)–(4.55) follows from the first part of Theorem 4.5.

5

Multiple-Access Channels

5.1 Introduction

The main focus of this monograph, and of the relevant literature on mismatched decoding, is on the point-to-point channel coding problem. As outlined in Section 1, this is in itself a very challenging problem with many important applications.

In this section, however, we turn the problem of coding for the *multiple-access channel* (MAC), in which two (or more) users send information to a common receiver. While this is a harder problem than the already-challenging single-user problem, its study is of interest for a number of reasons:

- The analysis of random coding comes with interesting new challenges not present in the single-user problem, both in terms of the achievability and ensemble tightness.
- We will find that multiple-access coding techniques can be beneficial even in the single-user setting, leading to rates strictly higher than the LM rate of Section 2. We will treat this perspective rather briefly in this section, focusing on the parallel BSC example of

Section 2.4.5, and then explore multi-user coding techniques for single-user channels in detail in Section 6.

The results of this section are mostly due to Lapidoth [58], but we will generally adopt the more recent proof techniques of Scarlett *et al.* [81].

5.2 Problem Setup

The problem setting that we consider is in direct analogy with the single-user setting of Section 1.1. We focus on the two-user MAC, which is sufficient to convey the key ideas, but the analysis techniques that we present can similarly be applied to any fixed number of users.

The input alphabets are denoted by $\mathcal{X}_1$ and $\mathcal{X}_2$, and the output alphabet is denoted by $\mathcal{Y}$. Except where stated otherwise, these are assumed to be finite, so that we are considering a discrete memoryless MAC (DM-MAC). The conditional transition law for a single channel use is denoted by $W(y|x_1, x_2)$, and the n-letter transition law is given by $W^n(\boldsymbol{y}|\boldsymbol{x}_1, \boldsymbol{x}_2) \triangleq \prod_{i=1}^n W(y_i|x_{1,i}, x_{2,i})$.

Encoder $\nu = 1, 2$ takes as input a message m_ν uniformly distributed on the set $\{1, \ldots, M_\nu\}$, and transmits the corresponding codeword $\boldsymbol{x}_\nu^{(m_\nu)}$ from a codebook $\mathcal{C}_\nu = \{\boldsymbol{x}_\nu^{(1)}, \ldots, \boldsymbol{x}_\nu^{(M_\nu)}\}$. Given the output sequence $\boldsymbol{y}$ generated according to $W^n(\,\cdot|\boldsymbol{x}_1, \boldsymbol{x}_2)$, the decoder forms an estimate $(\hat{m}_1, \hat{m}_2)$ of the message pair, given by

$$(\hat{m}_1, \hat{m}_2) = \operatorname*{arg\,max}_{(i,j):\ i\in\{1,\ldots,M_1\},\, j\in\{1,\ldots,M_2\}} q^n(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y}), \tag{5.1}$$

where $q^n(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) \triangleq \prod_{i=1}^n q(x_{1,i}, x_{2,i}, y_i)$ for some decoding metric $q(x_1, x_2, y)$. The method of tie-breaking has no impact on any of the results that we present, so we assume that ties are broken as errors.

The error probability for a given codebook pair $(\mathcal{C}_1, \mathcal{C}_2)$ is given by

$$p_{\mathrm{e}}(\mathcal{C}_1, \mathcal{C}_2) = \mathbb{P}[(\hat{m}_1, \hat{m}_2) \neq (m_1, m_2)]. \tag{5.2}$$

For any given codebook pair $(\mathcal{C}_1, \mathcal{C}_2)$, the decoding rule minimizing p_{e} is the maximum-likelihood (ML) rule, corresponding to (5.1) with $q(x_1, x_2, y) = W(y|x_1, x_2)$.

We consider a random coding scheme in which the codewords for the two users are independently drawn from some distributions $P_{\boldsymbol{X}_1}$

and $P_{\boldsymbol{X}_2}$:

$$(\{\boldsymbol{X}_1^{(i)}\}_{i=1}^{M_1}, \{\boldsymbol{X}_2^{(j)}\}_{j=1}^{M_2}) \sim \prod_{i=1}^{M_1} P_{\boldsymbol{X}_1}(\boldsymbol{x}_1^{(i)}) \prod_{j=1}^{M_2} P_{\boldsymbol{X}_2}(\boldsymbol{x}_2^{(j)}). \tag{5.3}$$

The error probability averaged over the random codebooks is denoted by $\bar{p}_{\mathrm{e}}(n, M_1, M_2)$.

Definition 5.1 (Mismatch Capacity Region)**.** For a given mismatched DM-MAC (W, q), a rate pair (R_1, R_2) is said to be achievable if, for all $\delta > 0$, there exist sequences of codebooks $(\mathcal{C}_{1,n}, \mathcal{C}_{2,n})$ of sizes $M_1 \geq e^{n(R_1-\delta)}$ and $M_2 \geq e^{n(R_2-\delta)}$ such that $p_{\mathrm{e}}(\mathcal{C}_{1,n}, \mathcal{C}_{2,n}) \to 0$ as $n \to \infty$. The *mismatch capacity region* $\mathcal{R}_{\mathrm{M}}^*$ is defined to be the set of all achievable rates.

To put our results in context, we recall the *matched* capacity region corresponding to $q(x_1, x_2, y) = W(y|x_1, x_2)$ [30, Ch. 4]:

$$\mathcal{R}_{\mathrm{Matched}}^* = \mathrm{conv}\Big(\bigcup_{Q_1, Q_2} \mathcal{R}(Q_1, Q_2) \Big), \tag{5.4}$$

where $\mathrm{conv}(\cdot)$ denotes the closure of the convex hull, the union is over all input distributions $Q_1 \in \mathcal{P}(\mathcal{X}_1)$ and $Q_2 \in \mathcal{P}(\mathcal{X}_2)$, and the set $\mathcal{R}(Q_1, Q_2)$ contains all pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2), \\ R_2 &\leq I(X_2; Y|X_1), \\ R_1 + R_2 &\leq I(X_1, X_2; Y), \end{aligned} \tag{5.5}$$

where $(X_1, X_2, Y) \sim Q_1 \times Q_2 \times W$. These rates are associated with three different types of errors, corresponding to the cases that (i) only the estimate of m_1 is incorrect; (ii) only the estimate of m_2 is incorrect, and (iii) both are incorrect. The achievability part is proved by applying random coding with input distributions Q_1 and Q_2, and using a time-sharing argument [30, Ch. 4] to achieve the convex hull.

5.3 Achievable Rate Region

In the single-user setting, we compared i.i.d. random coding and constant-composition random coding for DMCs, and found the latter to

provide better achievable rates. The same is true for the DM-MAC, so to simplify the exposition, we focus exclusively on constant-composition random coding.

For $\nu = 1, 2$, we fix $Q_\nu \in \mathcal{P}(\mathcal{X}_\nu)$ and let $P_{\boldsymbol{X}_\nu}$ be the uniform distribution on the type class $\mathcal{T}^n(Q_{\nu,n})$, where $Q_{\nu,n} \in \mathcal{P}_n(\mathcal{X}_\nu)$ is a type with the same support as Q_ν such that $\|Q_{\nu,n} - Q_\nu\|_\infty \leq \frac{1}{n}$. This yields

$$P_{\boldsymbol{X}_\nu}(\boldsymbol{x}_\nu) = \frac{1}{|\mathcal{T}^n(Q_{\nu,n})|} \mathbb{1}\{\boldsymbol{x}_\nu \in \mathcal{T}^n(Q_{\nu,n})\} \tag{5.6}$$

for $\nu = 1, 2$.

We first state and discuss the achievable rate region [58]. We then discuss a number of examples, before proceeding with the proof.

Theorem 5.1 (MAC Achievable Rate Region)**.** For any mismatched DM-MAC (W, q), and any input distributions $Q_1 \in \mathcal{P}(\mathcal{X}_1)$ and $Q_2 \in \mathcal{P}(\mathcal{X}_2)$, we have $\mathcal{R}_{\mathrm{LM}}(Q_1, Q_2) \subseteq \mathcal{R}^*_{\mathrm{M}}$, where $\mathcal{R}_{\mathrm{LM}}(Q_1, Q_2)$ is the set of (R_1, R_2) pairs satisfying

$$R_1 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1}=P_{X_1},\, \widetilde{P}_{X_2Y}=P_{X_2Y} \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} I_{\widetilde{P}}(X_1; X_2, Y), \tag{5.7}$$

$$R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_2}=P_{X_2},\, \widetilde{P}_{X_1Y}=P_{X_1Y} \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} I_{\widetilde{P}}(X_2; X_1, Y), \tag{5.8}$$

$$R_1 + R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1}=P_{X_1},\, \widetilde{P}_{X_2}=P_{X_2},\, \widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)], \\ I_{\widetilde{P}}(X_1;Y) \leq R_1,\, I_{\widetilde{P}}(X_2;Y) \leq R_2}} D(\widetilde{P}_{X_1X_2Y} \| Q_1 \times Q_2 \times P_Y), \tag{5.9}$$

where $P_{X_1X_2Y} = Q_1 \times Q_2 \times W$. That is, any rate pair (R_1, R_2) satisfying these conditions is achievable for the mismatched DM-MAC.

The proof is given in Section 5.5. Note that here and subsequently, the minimizations are over all joint distributions $\widetilde{P}_{X_1X_2Y} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ satisfying the specified constraints, but we let the inclusion in $\mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ remain implicit.

Although the rate conditions in (5.7)–(5.9) appear complex, they are natural generalizations of the primal expression for the LM rate

(*cf.*, Section 2.3). The condition in (5.7) arises by considering a typical realization of $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$, and studying the probability that some $(\overline{\boldsymbol{X}}_1, \boldsymbol{X}_2, \boldsymbol{Y})$ has a higher decoding metric for some incorrect $\overline{\boldsymbol{X}}_1$. Thinking of $P_{X_1X_2Y}$ and $\widetilde{P}_{X_1X_2Y}$ as the corresponding joint types of these triplets, the condition $\mathbb{E}_{\widetilde{P}}[\log q(X_1, X_2, Y)] \geq \mathbb{E}_P[\log q(X_1, X_2, Y)]$ corresponds to the event $q(\overline{\boldsymbol{X}}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \geq q(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$. Moreover, we have $\widetilde{P}_{X_2Y} = P_{X_2Y}$ because the two triplets use the same pair $(\boldsymbol{X}_2, \boldsymbol{Y})$, and $\widetilde{P}_{X_1} = P_{X_1}$ because we are considering constant-composition random coding. A similar discussion applies to the condition (5.8).

The condition (5.8) corresponds to some triplet $(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{Y})$ having a higher metric than $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$, and has some features that are somewhat less standard. In particular, the minimization problem contains the additional constraints $I_{\widetilde{P}}(X_1; Y) \leq R_1$ and $I_{\widetilde{P}}(X_2; Y) \leq R_2$. These constraints correspond to the fact that a random codebook $\mathcal{C}_\nu$ $(\nu = 1, 2)$ of size e^{nR_ν} is unlikely to produce *any* non-transmitted $\overline{\boldsymbol{X}}_\nu$ such that the empirical mutual information of $(\overline{\boldsymbol{X}}_\nu, \boldsymbol{Y})$ satisfies $I_{\widetilde{P}}(X_\nu; Y) > R_\nu$. We will see that these constraints are crucial for the ensemble tightness claim (see Section 5.6), in that the weaker achievable rate region with these constraints removed can be strictly smaller. In more detail, removing the constraints yields a pentagonal region similar to the matched MAC (5.5), whereas the region $\mathcal{R}_{\mathrm{LM}}(Q_1, Q_2)$ given above may be non-pentagonal and contain curved segments; see Section 5.4.2 for an example. Since the right-hand side of (5.9) is decreasing in R_1 and R_2, it is not even immediately obvious that $(R_1, R_2) \in \mathcal{R}_{\mathrm{LM}}(Q_1, Q_2) \implies (R_1', R_2') \in \mathcal{R}_{\mathrm{LM}}(Q_1, Q_2)$ when $R_1' \leq R_1$ and $R_2' \leq R_2$, but this property is indeed guaranteed, and can be seen by noting that (5.9) holds with strict inequality if and only if

$$\min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1}=P_{X_1}, \widetilde{P}_{X_2}=P_{X_2}, \widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} \min\{I_{\widetilde{P}}(X_1; Y) - R_1,$$
$$I_{\widetilde{P}}(X_2; Y) - R_2, D(\widetilde{P}_{X_1X_2Y} \| Q_1 \times Q_2 \times P_Y) - (R_1 + R_2)\} > 0. \tag{5.10}$$

This equivalence will arise in the achievability proof in Section 5.5.

The standard time-sharing argument for multiple-access channels [30, Ch. 4] applies without change for mismatched maximum-metric

decoding, and hence, we deduce from Theorem 5.1 that

$$\operatorname{conv}\Big(\bigcup_{Q_1,Q_2} \mathcal{R}_{\mathrm{LM}}(Q_1,Q_2) \Big) \subseteq \mathcal{R}_{\mathrm{M}}^{*}, \tag{5.11}$$

where $\operatorname{conv}(\cdot)$ denotes the closure of the convex hull. For the most part, however, we will focus on random coding under a fixed pair (Q_1, Q_2). A more detailed investigation of time-sharing for the mismatched MAC can be found in [77], where the distinction between *explicit time-sharing* and *coded time-sharing* is studied.

Since the achievable rate region of Theorem 5.1 is ensemble-tight (see Section 5.6 for a formal statement), it must recover the matched MAC conditions (5.5) in the case that $q(x_1, x_2, y) = W(y|x_1, x_2)$. As a sanity check, this can be checked directly via similar steps to (2.54)–(2.61) [58], and by doing so, one finds that the constraints $I_{\widetilde{P}}(X_1;Y) \leq R_1$ and $I_{\widetilde{P}}(X_2;Y) \leq R_2$ in (5.9) have no impact in the matched case.

5.4 Examples

5.4.1 Parallel Binary Symmetric Channels

We return to the parallel binary symmetric channel (BSC) example studied in the single-user setting in Section 2.4.5 [58]. We now study the channel as a MAC, with inputs X_1 and X_2, and output $Y = (Y_1, Y_2)$. Recall that the channels from X_ν to Y_ν ($\nu = 1, 2$) are independent BSCs with crossover probabilities δ_1 and δ_2, and the metric is

$$\log q(x_1, x_2, (y_1, y_2)) = \frac{1}{2}(\mathbb{1}\{x_1 = y_1\} + \mathbb{1}\{x_2 = y_2\}). \tag{5.12}$$

We let the input distributions Q_1, Q_2 be equiprobable on $\{0, 1\}$.

Analysis of the Achievable Rate Region. The conditions (5.7)–(5.8) are straightforward to analyze, reducing to $R_\nu \leq 1 - H_2(\delta_\nu)$ bits/use for $\nu = 1, 2$. We therefore focus our attention on (5.9), which is the most interesting of the three. The main step is to show that this constraint is satisfied by the pair $(R_1, R_2) = (1 - H_2(\delta_1), 1 - H_2(\delta_2))$.

Fixing any $\widetilde{P}_{X_1X_2Y}$ satisfying the constraints in (5.9), we have

$$\begin{aligned} D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times P_Y) & \\ &= \mathbb{E}_{\widetilde{P}}\left[\log\frac{\widetilde{P}_{X_1X_2Y}(X_1,X_2,Y)}{\widetilde{P}_{X_1}(X_1)\widetilde{P}_{X_2}(X_2)\widetilde{P}_Y(Y)}\right] && (5.13)\\ &= I_{\widetilde{P}}(X_1,X_2;Y)+I_{\widetilde{P}}(X_1;X_2) && (5.14)\\ &\geq I_{\widetilde{P}}(X_1,X_2;Y), && (5.15)\end{aligned}$$

where (5.13) makes use of the marginal constraints, and (5.14) follows by multiplying and dividing by $\widetilde{P}_{X_1X_2}$ in (5.13). The mutual information $I_{\widetilde{P}}(X_1,X_2;Y)$ was already analyzed in Section 2.4.5, where it was shown that

$$I_{\widetilde{P}}(X_1,X_2;Y)\geq(1-H_2(\widetilde{\delta}_1))+(1-H_2(\widetilde{\delta}_2)), \tag{5.16}$$

where $\widetilde{\delta}_1$ and $\widetilde{\delta}_2$ are crossover probabilities corresponding to $\widetilde{P}_{Y_1|X_1}$ and $\widetilde{P}_{Y_2|X_2}$, both of which must take the form of a BSC.

The analysis now departs from that of Section 2.4.5 by incorporating the constraints $I_{\widetilde{P}}(X_1;Y)\leq R_1$ and $I_{\widetilde{P}}(X_2;Y)\leq R_2$. Since we are considering the rate pair with $R_\nu=1-H_2(\delta_\nu)$ for $\nu=1,2$, and since $I_{\widetilde{P}}(X_\nu;Y_\nu)=1-H_2(\widetilde{\delta}_\nu)$, the constraints reduce to $H_2(\delta_\nu)\leq H_2(\widetilde{\delta}_\nu)$, and we deduce from (5.16) that

$$I_{\widetilde{P}}(X_1,X_2;Y)\geq(1-H_2(\delta_1))+(1-H_2(\delta_2)), \tag{5.17}$$

and as a result, the condition (5.9) is indeed satisfied by $(R_1,R_2)=(1-H_2(\delta_1),1-H_2(\delta_2))$. This, in turn, implies the achievability of all rate pairs within the rectangle with corners $(0,0)$ and $(1-H_2(\delta_1),1-H_2(\delta_2))$, which is the capacity region of the matched MAC, i.e., the capacity of two parallel BSCs.

Discussion. Due to the parallel structure of the codebooks and the channel, if we let $\log q$ be *any* positive weighted sum of $\mathbb{1}\{x_1=y_1\}$ and $\mathbb{1}\{x_2=y_2\}$ (*cf.*, (5.12)), we obtain a decoding rule that separately minimizes the Hamming distance between $(\boldsymbol{X}_1,\boldsymbol{Y}_1)$ and $(\boldsymbol{X}_2,\boldsymbol{Y}_2)$, which in turn is equivalent to maximum-likelihood decoding. Hence, it should not be surprising that we achieve the matched capacity region.

Nevertheless, the above analysis highlights the importance of the constraints $I_{\widetilde{P}}(X_\nu;Y_\nu)\leq R_\nu$; the analysis of Section 2.4.5 shows that

in the absence of these constraints, the right-hand side of (5.9) is equal to the suboptimal value $2(1 - H_2(\frac{\delta_1+\delta_2}{2}))$.

This example also highlights an interesting phenomenon that we will return to in detail in Section 6: Multi-user coding techniques can provide improved rates even for single-user mismatched channels. Indeed, we saw that the LM rate fails to achieve the capacity $C = (1 - H_2(\delta_1)) + (1 - H_2(\delta_2))$ under a uniform input distribution, and if we deviate from a uniform input, then even optimal decoding fails to achieve capacity. In contrast, we are always free to generate two codebooks in parallel and treat the single-user channel as if it were a MAC, and the above analysis reveals that doing so achieves the matched capacity (and hence also the mismatch capacity).

5.4.2 Non-Convexity of $\mathcal{R}_{\mathrm{LM}}(Q_1, Q_2)$

It is well-known that even in the matched case, the union of the region (5.5) over all pairs (Q_1, Q_2) can be non-convex [30, Ch. 4], meaning that the convex hull operation in (5.4) is not redundant. However, the rate region (5.5) for a given pair (Q_1, Q_2) is always convex, and typically pentagonal; in certain special cases, triangular and quadrilateral shapes are also possible.

In contrast, in the mismatched setting, due to the constraints $I_{\widetilde{P}}(X_1;Y) \leq R_1$ and $I_{\widetilde{P}}(X_2;Y) \leq R_2$ in (5.9), the achievable rate region $\mathcal{R}_{\mathrm{LM}}(Q_1, Q_2)$ corresponding to fixed (Q_1, Q_2) can be non-convex [77], and can have curved parts on the boundary rather than only straight lines. To demonstrate this, we consider the mismatched MAC with $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$, $\mathcal{Y} = \{0, 1, 2\}$, and

$$W(y|x_1, x_2) = \begin{cases} 1 - 2\delta_{x_2} & y = x_1 + x_2 \\ \delta_{x_2} & \text{otherwise,} \end{cases} \tag{5.18}$$

$$q(x_1, x_2, y) = \begin{cases} 1 - 2\delta & y = x_1 + x_2 \\ \delta & \text{otherwise,} \end{cases} \tag{5.19}$$

where we choose $\delta_0 = 0.25$ and $\delta_1 = 0.01$, and let δ equal an arbitrary value in $(0, \frac{1}{3})$, any of which yields an equivalent metric. Hence,

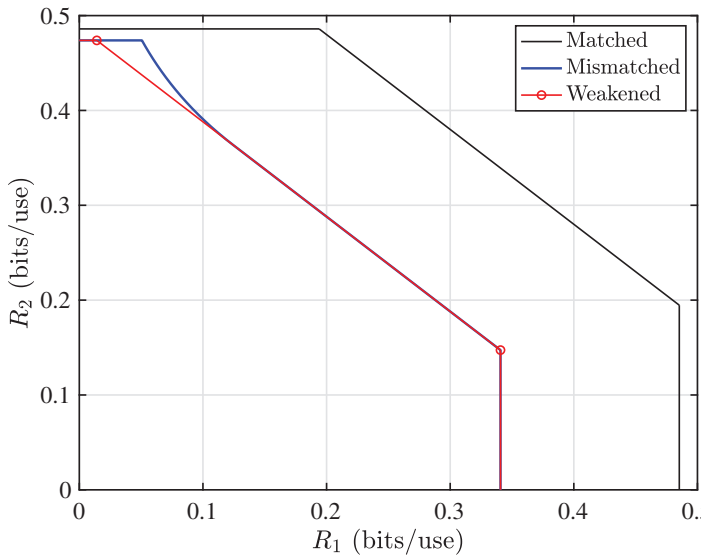


Figure 5.1: Achievable rate regions for the MAC in (5.18)–(5.19) in the matched and mismatched settings, with input distributions $Q_1 = Q_2 = (\frac{1}{2}, \frac{1}{2})$. The weakened region removes the two mutual information constraints in (5.9).

the mismatched decoder incorrectly assumes both of the transition probabilities δ_{x_2} are equal.

Figure 5.1 plots the achievable rate region $\mathcal{R}_{\mathrm{LM}}(Q_1, Q_2)$ corresponding to $Q_1 = Q_2 = (\frac{1}{2}, \frac{1}{2})$. We also plot the weakened region obtained when the constraints $I_{\widetilde{P}}(X_1; Y) \leq R_1$ and $I_{\widetilde{P}}(X_2; Y) \leq R_2$ are removed from (5.9), as well as the matched region (5.5).

We see that while the weakened and matched regions take the usual pentagonal shape, the ensemble-tight region $\mathcal{R}_{\mathrm{LM}}(Q_1, Q_2)$ curves upward near its upper "corner point", and forms a non-convex shape. This is due to the fact that the constraint $I_{\widetilde{P}}(X_1; Y) \leq R_1$ is active in this part of the curve. In this example, the constraint $I_{\widetilde{P}}(X_2; Y) \leq R_2$ has no effect on the rate region.

5.5 Proof of Achievable Rate Region

In this subsection, we present the proof of Theorem 5.1. The original proof of Lapidoth [58] and the more recent proof of Scarlett *et al.*, [81]

share several common ideas. In the following, we focus on the latter, presenting the analysis in several steps.

5.5.1 Separation into Three Error Events

We study the random-coding error probability by considering the following three error types, where we assume without loss of generality that the true messages are $(m_1, m_2) = (1, 1)$:

$$\begin{aligned}
&\textit{(Type 1)} \quad q^n(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \text{ for some } i \neq 1; \\
&\textit{(Type 2)} \quad q^n(\boldsymbol{X}_1, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \text{ for some } j \neq 1; \\
&\textit{(Type 12)} \quad q^n(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \text{ for some } i \neq 1,\ j \neq 1.
\end{aligned}$$

The corresponding probabilities of the error events are denoted by $\bar{p}_{\mathrm{e},1}(n, M_1)$, $\bar{p}_{\mathrm{e},2}(n, M_2)$, and $\bar{p}_{\mathrm{e},12}(n, M_1, M_2)$, respectively. Writing the overall random-coding error probability as $\bar{p}_{\mathrm{e}}(n, M_1, M_2)$, we immediately obtain from the union bound that

$$\bar{p}_{\mathrm{e}}(n, M_1, M_2) \leq \bar{p}_{\mathrm{e},1}(n, M_1) + \bar{p}_{\mathrm{e},2}(n, M_2) + \bar{p}_{\mathrm{e},12}(n, M_1, M_2). \quad (5.20)$$

Thus, we have $\bar{p}_{\mathrm{e}} \to 0$ as long as all three of $\bar{p}_{\mathrm{e},1}$, $\bar{p}_{\mathrm{e},2}$ and $\bar{p}_{\mathrm{e},12}$ vanish, and we can consider the three error types separately.

The analysis of all three error events is similar, but $\bar{p}_{\mathrm{e},12}$ is the only one that requires new tools compared to the single-user setting, studied in Section 2.6. We therefore focus only on $\bar{p}_{\mathrm{e},12}$ in the remainder of the proof.

5.5.2 Separation into Joint Types

The type-12 error probability $\bar{p}_{\mathrm{e},12}(n, M_1, M_2)$ admits the following exact expression:

$$\bar{p}_{\mathrm{e},12} = \mathbb{E}\left[\mathbb{P}\left[\bigcup_{i\neq 1, j\neq 1} \left\{\frac{q^n(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y})}{q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})} \geq 1\right\} \,\middle|\, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}\right]\right]. \quad (5.21)$$

In analogy with the primal-domain analysis of the single-user setting (Section 2.6.3), we will rewrite (5.21) in terms of joint types. To lighten

the notion, we introduce the sets

$$\mathcal{S}_n(\boldsymbol{Q}_n) \triangleq \{P_{X_1X_2Y} \in \mathcal{P}_n(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})\colon P_{X_1} = Q_{1,n},\, P_{X_2} = Q_{2,n}\}, \tag{5.22}$$

$$\mathcal{S}'_{12,n}(P_{X_1X_2Y}) \triangleq \{\widetilde{P}_{X_1X_2Y} \in \mathcal{P}_n(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})\colon \widetilde{P}_{X_1} = P_{X_1},$$
$$\widetilde{P}_{X_2} = P_{X_2}, \widetilde{P}_Y = P_Y, \mathbb{E}_{\widetilde{P}}[\log q(X_1, X_2, Y)] \geq \mathbb{E}_P[\log q(X_1, X_2, Y)]\Big\}, \tag{5.23}$$

where $\mathcal{S}_n(\boldsymbol{Q}_n)$ can be thought of as the set of possible joint types of $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$ when each user has a constant-composition codebook, and $\mathcal{S}'_{12,n}(P_{X_1X_2Y})$ can be thought of as the set of joint types of $(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{Y})$ that lead to a decoding error when $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in \mathcal{T}^n(P_{X_1X_2Y})$.

Fixing $P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)$ and letting $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})$ be an arbitrary triplet such that $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) \in \mathcal{T}^n(P_{X_1X_2Y})$, we find that the event in (5.21) can be written as

$$\bigcup_{i\neq 1, j\neq 1} \bigcup_{\widetilde{P}_{X_1X_2Y} \in \mathcal{S}'_{12,n}(P_{X_1X_2Y})} \Big\{(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})\Big\}. \tag{5.24}$$

Writing the probability and expectation in (5.21) as summations over joint types, substituting (5.24), and interchanging the order of the unions, we obtain the following:

$$\bar{p}_{\mathrm{e},12} = \sum_{P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)} \mathbb{P}[(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in \mathcal{T}^n(P_{X_1X_2Y})]$$
$$\times \mathbb{P}\Bigg[\bigcup_{\widetilde{P}_{X_1X_2Y} \in \mathcal{S}'_{12,n}(P_{X_1X_2Y})} \bigcup_{i\neq 1, j\neq 1} \Big\{(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})\Big\}\Bigg] \tag{5.25}$$

$$\leq \sum_{P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)} \mathbb{P}\big[(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in \mathcal{T}^n(P_{X_1X_2Y})\big] \min\Bigg\{1,$$
$$\times \sum_{\widetilde{P}_{X_1X_2Y} \in \mathcal{S}'_{12,n}(P_{X_1X_2Y})} \mathbb{P}\Bigg[\bigcup_{i\neq 1, j\neq 1} \Big\{(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})\Big\}\Bigg]\Bigg\}, \tag{5.26}$$

where $\boldsymbol{y}$ is an arbitrary element of $\mathcal{T}^n(P_Y)$ (hence depending implicitly on $P_{X_1X_2Y}$), and (5.26) follows from the truncated union bound (i.e., the minimum of one and the union bound).

5.5.3 A Refined Union Bound

It is tempting, particularly based on the primal single-user analysis in Section 2.6, to apply the truncated union bound to the inner probability in (5.26) to obtain

$$\mathbb{P}\Bigg[\bigcup_{i\neq 1, j\neq 1}\Big\{(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})\Big\}\Bigg] \leq \min\Big\{1, (M_1-1)(M_2-1)\mathbb{P}[(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})]\Big\}, \tag{5.27}$$

where $(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2) \sim P_{\boldsymbol{X}_1} \times P_{\boldsymbol{X}_2}$. However, proceeding with this standard union bound would lead to a condition of the form (5.9) without the constraints $I_{\widetilde{P}}(X_1;Y) \leq R_1$ and $I_{\widetilde{P}}(X_2;Y) \leq R_2$, and we know that these constraints can strictly improve the achievable rate region.

Fortunately, a straightforward solution is available. Observe that in order to have $(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})$, it must be the case that $(\boldsymbol{X}_1^{(i)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1Y})$. This means that the left-hand side of (5.27) is upper bounded by

$$\mathbb{P}\Bigg[\bigcup_{i\neq 1}\Big\{(\boldsymbol{X}_1^{(i)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1Y})\Big\}\Bigg] \leq (M_1-1)\mathbb{P}[(\overline{\boldsymbol{X}}_1, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1Y})]. \tag{5.28}$$

A similar bound follows by noting that $(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})$ implies $(\boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_2Y})$, and putting these bounds together with (5.27) yields

$$\mathbb{P}\Bigg[\bigcup_{i\neq 1, j\neq 1}\Big\{(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})\Big\}\Bigg] \leq \min\Big\{1, (M_1-1)\mathbb{P}[(\overline{\boldsymbol{X}}_1, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1Y})],$$

$$(M_2 - 1)\mathbb{P}\big[(\overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_2Y})\big],$$
$$(M_1 - 1)(M_2 - 1)\mathbb{P}\big[(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})\big]\Big\}. \tag{5.29}$$

This refined bound turns out to be sufficient, and in fact, we will see in Section 5.6 that it is tight to within a factor of four.

5.5.4 Deducing the Rate Condition

The remainder of the analysis is analogous to the single-user setting, so we provide a relatively concise treatment. In the following, we let $\delta > 0$ be an arbitrarily small positive constant.

By standard properties of types similar to Section 2.6.1, the three probabilities in (5.29) respectively behave as $e^{-nI_{\widetilde{P}}(X_1;Y)}$, $e^{-nI_{\widetilde{P}}(X_2;Y)}$, and $e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times P_Y)}$ times sub-exponential factors. In addition, the number of terms in the inner summation of (5.26) is polynomial in n, so the summation can be upper bounded by a polynomial times the corresponding maximum. It follows that for a given joint type $P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)$ corresponding to the outer summation of (5.26), the $\min\{1, \cdot\}$ term vanishes as $n \to \infty$ as long as

$$\min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{S}'_{12,n}(P_{X_1X_2Y})} \min\Big\{I_{\widetilde{P}}(X_1;Y) - R_1, I_{\widetilde{P}}(X_2;Y) - R_2,$$
$$D(\widetilde{P}_{X_1X_2Y}\|Q_1 \times Q_2 \times P_Y) - (R_1 + R_2)\Big\} \geq \delta. \tag{5.30}$$

Using the definition of $\mathcal{S}'_{12,n}$ in (5.23) and lower bounding the minimum over joint types by a minimum over all joint distributions, we find that (5.30) holds as long as

$$\min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1}=P_{X_1},\widetilde{P}_{X_2}=P_{X_2},\widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} \min\Big\{I_{\widetilde{P}}(X_1;Y) - R_1, I_{\widetilde{P}}(X_2;Y) - R_2,$$
$$D(\widetilde{P}_{X_1X_2Y}\|Q_1 \times Q_2 \times P_Y) - (R_1 + R_2)\Big\} \geq \delta. \tag{5.31}$$

Next, recall that the outer summation of (5.26) sums over the possible joint types $P_{X_1X_2Y}$ of $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$. Again using standard properties of

types similar to Section 2.6.1, we have

$$\mathbb{P}[(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in \mathcal{T}^n(P_{X_1X_2Y})] \leq e^{-n(D(P_{X_1X_2Y}\|Q_1\times Q_2\times W)-\delta')} \quad (5.32)$$

for arbitrarily small $\delta' > 0$ and sufficiently large n. In particular, since the divergence is only zero when the arguments are identical, the contribution of joint types with $\|P_{X_1X_2Y} - Q_1 \times Q_2 \times W\|_\infty > \delta$ is asymptotically negligible. This means that the error probability vanishes as long as (5.31) holds for all $P_{X_1X_2Y}$ such that $\|P_{X_1X_2Y} - Q_1 \times Q_2 \times W\|_\infty \leq \delta$.

Following a similar continuity argument to the single-user setting (*cf.*, Appendix A.1), one finds that the left-hand side of (5.31) is continuous as a function of $P_{X_1X_2Y}$. Hence, and since δ can be arbitrarily small, it is sufficient for (5.31) to hold under the specific choice $P_{X_1X_2Y} = Q_1 \times Q_2 \times W$. Finally, since (5.31) trivially holds when $R_1 > I_{\widetilde{P}}(X_1;Y) + \delta$ or $R_2 > I_{\widetilde{P}}(X_2;Y) + \delta$, we can limit the minimum to distributions satisfying $I_{\widetilde{P}}(X_1;Y) \leq R_1 + \delta$ and $I_{\widetilde{P}}(X_2;Y) \leq R_2 + \delta$, yielding

$$\begin{aligned} &R_1 + R_2 + \delta \\ &\quad \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1}=P_{X_1},\widetilde{P}_{X_2}=P_{X_2},\widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)]\geq\mathbb{E}_P[\log q(X_1,X_2,Y)] \\ I_{\widetilde{P}}(X_1;Y)\leq R_1+\delta,\ I_{\widetilde{P}}(X_2;Y)\leq R_2+\delta}} D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times P_Y). \end{aligned} \quad (5.33)$$

Note that the first two terms of the $\min\{\cdot\}$ in (5.31) become redundant under the new constraints, and are therefore not included here. Finally, a continuity argument reveals that the right-hand side of (5.33) is continuous with respect to $\delta \geq 0$ [58, p. 1447], and taking $\delta \to 0$, we obtain the final rate condition (5.9).

5.6 Ensemble Tightness

In Section 2.5.2, we introduced the notion of *ensemble tightness* of the single-user GMI and LM rate, showing that although these do not equal the mismatch capacity in general, no higher rates can be attained using the respective random coding ensembles. The ensemble tightness proofs

were based on the fact that the properties of types used in deriving the primal expressions are exponentially tight.

For the achievable rate region of the MAC in Theorem 5.1, establishing ensemble tightness requires additional effort. As with the achievability part, the difficulty lies in the rate condition (5.9), and more specifically, establishing the tightness of the refined union bound (5.29). To do this, we follow the approach of Scarlett *et al.* [81] and make use of a general lower bound on the probability of a union due to de Caen [27]. An alternative proof based on the second moment method was given in the earlier work of Lapidoth [58].

Before proceeding, we formally state the ensemble tightness result.

Theorem 5.2 (Ensemble Tightness for the MAC)**.** For any DM-MAC (W, q), under constant-composition random coding (*cf.*, Section (5.6)) with input distributions (Q_1, Q_2), for any rate pair (R_1, R_2) failing to satisfy (5.7)–(5.9), we have

$$\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR_1} \rfloor, \lfloor e^{nR_2} \rfloor) \to 1 \tag{5.34}$$

as $n \to \infty$.

The remainder of the subsection is devoted to the proof.

5.6.1 Separation into Three Error Events

As with the achievability part, we consider the separation into three error types in Section 5.5.1, and note that

$$\bar{p}_{\mathrm{e}}(n, M_1, M_2) \geq \max\{\bar{p}_{\mathrm{e},1}(n, M_1), \bar{p}_{\mathrm{e},2}(n, M_2), \bar{p}_{\mathrm{e},12}(n, M_1, M_2)\}, \tag{5.35}$$

where $M_\nu = \lfloor e^{nR_\nu} \rfloor$ for $\nu = 1, 2$. The single-user analysis of Section 2.5.2 readily extends to the analysis of $\bar{p}_{\mathrm{e},1}$ and $\bar{p}_{\mathrm{e},2}$, revealing that the respective error probabilities tend to one when (5.7) or (5.8) fail to hold. We therefore focus our attention on the most difficult error event; we will show that $\bar{p}_{\mathrm{e},12} \to 1$ whenever (5.9) fails to hold.

5.6.2 Separation into Joint Types

Recall that the exact type-12 error probability can be written as in (5.25). Lower bounding the union over $\widetilde{P}_{X_1X_2Y}$ therein by a maximum, we

obtain

$$\bar{p}_{\mathrm{e},12} \geq \sum_{P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)} \mathbb{P}\big[(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in \mathcal{T}^n(P_{X_1X_2Y})\big] \times \max_{\widetilde{P}_{X_1X_2Y} \in \mathcal{S}'_{12,n}(P_{X_1X_2Y})} \bar{p}_{\mathrm{e},12}(\widetilde{P}_{X_1X_2Y}), \tag{5.36}$$

where, with a slight abuse of notation, we define

$$\bar{p}_{\mathrm{e},12}(\widetilde{P}_{X_1X_2Y}) \triangleq \mathbb{P}\Big[\bigcup_{i\neq 1, j\neq 1} \{\mathcal{E}_{ij}(\widetilde{P}_{X_1X_2Y})\}\Big] \tag{5.37}$$

in terms of the event

$$\mathcal{E}_{ij}(\widetilde{P}_{X_1X_2Y}) \triangleq \Big\{(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})\Big\}. \tag{5.38}$$

5.6.3 Lower Bound on the Doubly-Indexed Union

The key step in the analysis is to lower bound (5.37) using the following result of de Caen [27].

Lemma 5.3 (de Caen's Bound [27])**.** For any finite sequence of events $\mathcal{A}_1, \ldots, \mathcal{A}_N$ on a probability space, we have

$$\mathbb{P}\Big[\bigcup_{l=1}^{N} \mathcal{A}_l\Big] \geq \sum_{l=1}^{N} \frac{\mathbb{P}[\mathcal{A}_l]^2}{\sum_{l'=1}^{N} \mathbb{P}[\mathcal{A}_l \cap \mathcal{A}_{l'}]}. \tag{5.39}$$

To apply this result to (5.37), we identify l with (i, j) and N with $(M_1-1)(M_2-1)$, and distinguish between four cases for the denominator term $\mathbb{P}[\mathcal{E}_{ij}(\widetilde{P}_{X_1X_2Y}) \cap \mathcal{E}_{i'j'}(\widetilde{P}_{X_1X_2Y})]$:

- If $i = i'$ and $j = j'$, then the probability of the intersection is simply
$$\Psi_{00} \triangleq \mathbb{P}[(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})], \tag{5.40}$$
where here and subsequently, $(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2) \sim P_{\boldsymbol{X}_1}(\boldsymbol{x}_1)P_{\boldsymbol{X}_2}(\boldsymbol{x}_2)$ denote generic random codewords drawn independently from $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$.

- If $i \neq i'$ and $j = j'$, then the probability of the intersection simplifies to

$$\Psi_{10} \triangleq \mathbb{P}[(\overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_2Y})] \times \mathbb{P}[(\overline{\boldsymbol{X}}_1, \boldsymbol{x}_2, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})]^2, \tag{5.41}$$

  where in the second probability, $\boldsymbol{x}_2$ is an arbitrary sequence such that $(\boldsymbol{x}_2, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_2Y})$. This is because the two relevant $(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y})$ triplets share the same $\boldsymbol{X}_2$ sequence, which must have joint type $\widetilde{P}_{X_2Y}$, whereas the two different $\boldsymbol{X}_1$ sequences (which are independent of each other) both need to induce the joint type $\widetilde{P}_{X_1X_2Y}$.

- Similarly, if $i \neq i'$ and $j = j'$, then the probability of the intersection simplifies to

$$\Psi_{01} \triangleq \mathbb{P}\big[(\overline{\boldsymbol{X}}_1, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1Y})\big] \times \mathbb{P}\big[(\boldsymbol{x}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{X_1X_2Y})\big]^2. \tag{5.42}$$

- If $i \neq i'$ and $j \neq j'$, then the two events in the intersection are independent, and hence their joint probability is $\Psi_{11} \triangleq \Psi_{00}^2$.

For any given pair (i, j), the number of (i', j') pairs falling into these four categories is given by 1, $M_1 - 2$, $M_2 - 2$ and $(M_1 - 2)(M_2 - 2)$, respectively. As a result, Lemma 5.3 yields the following:

$$\overline{p}_{\mathrm{e},12}(\widetilde{P}_{X_1X_2Y}) \geq \frac{(M_1 - 1)(M_2 - 1)\Psi_{00}^2}{\Psi_{00} + (M_1 - 2)\Psi_{10} + (M_1 - 2)\Psi_{01} + (M_1 - 2)(M_2 - 2)\Psi_{00}^2} \tag{5.43}$$

$$\geq \frac{(M_1 - 1)(M_2 - 1)\Psi_{00}^2}{\Psi_{00} + (M_1 - 1)\Psi_{10} + (M_1 - 1)\Psi_{01} + (M_1 - 1)(M_2 - 1)\Psi_{00}^2} \tag{5.44}$$

$$\geq \frac{(M_1-1)(M_2-1)\Psi_{00}^2}{4\max\{\Psi_{00},(M_1-1)\Psi_{10},(M_1-1)\Psi_{01},(M_1-1)(M_2-1)\Psi_{00}^2\}} \tag{5.45}$$

$$= \frac{1}{4}\min\left\{(M_1-1)(M_2-1)\Psi_{00},(M_2-1)\frac{\Psi_{00}^2}{\Psi_{10}},(M_1-1)\frac{\Psi_{00}^2}{\Psi_{01}},1\right\}. \tag{5.46}$$

Next, observe that Ψ_{00} in (5.40) can be rewritten as

$$\Psi_{00} = \mathbb{P}[(\overline{\boldsymbol{X}}_2,\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{X_2Y})]\mathbb{P}[(\overline{\boldsymbol{X}}_1,\boldsymbol{x}_2,\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{X_1X_2Y})] \tag{5.47}$$

$$= \mathbb{P}[(\overline{\boldsymbol{X}}_1,\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{X_1Y})]\mathbb{P}[(\boldsymbol{x}_1,\overline{\boldsymbol{X}}_2,\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{X_1X_2Y})], \tag{5.48}$$

which implies that

$$\frac{\Psi_{00}}{\Psi_{10}} = \mathbb{P}\big[(\overline{\boldsymbol{X}}_1,\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{X_1Y})\big], \tag{5.49}$$

$$\frac{\Psi_{00}}{\Psi_{01}} = \mathbb{P}\big[(\overline{\boldsymbol{X}}_2,\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{X_2Y})\big]. \tag{5.50}$$

Putting everything back together into (5.46), we deduce that the upper bound in (5.29) is tight to within a factor of four.

5.6.4 Deducing Ensemble Tightness

In the following, let $\delta > 0$ be arbitrarily small. As we mentioned in Section 5.5.4, the three probabilities in (5.29) behave as $e^{-nI_{\widetilde{P}}(X_1;Y)}$, $e^{-nI_{\widetilde{P}}(X_2;Y)}$, and $e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times P_Y)}$ times sub-exponential factors, and the joint type $P_{X_1X_2Y}$ being summed over in (5.36) satisfies $\|P_{X_1X_2Y}-Q_1\times Q_2\times W\|_\infty\leq\delta$ with probability approaching one. From these facts, we deduce that $p_{\mathrm{e},12}\to 1$ whenever the following condition holds for all such $P_{X_1X_2Y}$:

$$\min_{\widetilde{P}_{X_1X_2Y}\in\mathcal{S}'_{12,n}(P_{X_1X_2Y})}\min\Big\{I_{\widetilde{P}}(X_1;Y)-R_1,I_{\widetilde{P}}(X_2;Y)-R_2,$$
$$D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times P_Y)-(R_1+R_2)\Big\}\geq\delta. \tag{5.51}$$

The main challenge now is to upper bound the minimum over joint types (see (5.23)) in terms of the corresponding minimum over all joint distributions. This is done in a similar way to the single-user setting, which itself was rather technical; further details are given in Appendix A.1. Once this is done, we obtain a matching converse bound to (5.31), from which the necessity of the condition (5.9) follows.

5.7 Dual Expressions and Continuous Alphabets

The rate region of Theorem 5.1 is analogous to the primal expression for the single-user LM rate. It is therefore natural to ask whether there also exists a dual expression, written as a maximization over real-valued parameters rather than a minimization over joint distributions. The answer is affirmative, as the following theorem shows [81]. Here and subsequently, we write $\mathbb{E}_Z[\cdot]$ to denote averaging over some random variable Z while all other random variables are held fixed.

Theorem 5.4 (Dual Expression for the Achievable Rate Region)**.** The region $\mathcal{R}_{\mathrm{LM}}(Q_1, Q_2)$ described by (5.7)–(5.9) can be expressed as the set of rate pairs (R_1, R_2) satisfying

$$R_1 \leq \sup_{s\geq 0, a_1(\cdot)} \mathbb{E}\left[\log \frac{q(X_1, X_2, Y)^s e^{a_1(X_1)}}{\mathbb{E}\big[q(\overline{X}_1, X_2, Y)^s e^{a_1(\overline{X}_1)} \big| X_2, Y\big]}\right], \tag{5.52}$$

$$R_2 \leq \sup_{s\geq 0, a_2(\cdot)} \mathbb{E}\left[\log \frac{q(X_1, X_2, Y)^s e^{a_2(X_2)}}{\mathbb{E}\big[q(X_1, \overline{X}_2, Y)^s e^{a_2(\overline{X}_2)} \big| X_1, Y\big]}\right], \tag{5.53}$$

and at least one of the following:

$$\begin{aligned} R_1 \leq \sup_{\rho_2\in[0,1], s\geq 0, a_1(\cdot), a_2(\cdot)} & -\rho_2 R_2 \\ & + \mathbb{E}\left[\log \frac{\big(q(X_1, X_2, Y)^s e^{a_2(X_2)}\big)^{\rho_2} e^{a_1(X_1)}}{\mathbb{E}_{\overline{X}_1}\Big[\Big(\mathbb{E}_{\overline{X}_2}\big[q(\overline{X}_1, \overline{X}_2, Y)^s e^{a_2(\overline{X}_2)}\big]\Big)^{\rho_2} e^{a_1(\overline{X}_1)}\Big]}\right], \end{aligned} \tag{5.54}$$

$$R_2 \leq \sup_{\rho_1 \in [0,1], s \geq 0, a_1(\cdot), a_2(\cdot)} -\rho_1 R_1 + \mathbb{E}\left[\log \frac{\left(q(X_1, X_2, Y)^s e^{a_1(X_1)}\right)^{\rho_1} e^{a_2(X_2)}}{\mathbb{E}_{\overline{X}_2}\left[\left(\mathbb{E}_{\overline{X}_1}\left[q(\overline{X}_1, \overline{X}_2, Y)^s e^{a_1(\overline{X}_1)}\right]\right)^{\rho_1} e^{a_2(\overline{X}_2)}\right]}\right]. \tag{5.55}$$

where $(X_1, X_2, Y, \overline{X}_1, \overline{X}_2) \sim Q_1(x_1)Q_2(x_2)W(y|x_1, x_2)Q_1(\overline{x}_1)Q_2(\overline{x}_2)$.

The most interesting feature of this result is the presence of the additional optimization variables ρ_1 and ρ_2 in (5.54)–(5.55). The presence of these variables is closely related to the presence of the mutual information constraints in (5.9): If we were to remove these primal-domain constraints, then the equivalent dual expression would only allow $\rho_1 = \rho_2 = 1$ in (5.54)–(5.55), in which case the two conditions would be identical.

5.7.1 Overview of the Analysis

The equivalence to the primal expression is proved using Lagrange duality, but with a non-trivial additional step to handle the constraints $I_{\widetilde{P}}(X_1; Y) \leq R_1$ and $I_{\widetilde{P}}(X_2; Y) \leq R_2$ in (5.9). Indeed, forming a simple dual expression directly from (5.9) via Lagrange duality appears to be difficult. However, the following result from [81] establishes that at most one of the two mutual information constraints is active at a time, leading to a formulation more amenable to forming a Lagrange dual.

Lemma 5.5 (Equivalent Formulation of Sum Rate Condition)**.** For any mismatched DM-MAC (W, q) and input distributions (Q_1, Q_2), the condition (5.9) holds if and only if at least one of the following holds:

$$R_1 + R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1}=P_{X_1}, \widetilde{P}_{X_2}=P_{X_2}, \widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)], \\ I_{\widetilde{P}}(X_1;Y) \leq R_1}} D(\widetilde{P}_{X_1X_2Y} \| P_{X_1X_2} \times P_Y), \tag{5.56}$$

$$R_1 + R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1}=P_{X_1},\widetilde{P}_{X_2}=P_{X_2},\widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)]\geq \mathbb{E}_P[\log q(X_1,X_2,Y)], \\ I_{\widetilde{P}}(X_2;Y)\leq R_2}} D(\widetilde{P}_{X_1X_2Y}\|P_{X_1X_2}\times P_Y), \tag{5.57}$$

where $P_{X_1X_2Y} = Q_1 \times Q_2 \times W$.

The expressions in (5.56)–(5.57) constitute the primal forms of the dual expressions in (5.54)–(5.55), with the equivalence proved using Lagrange duality. Much like the single-user setting, the dual expressions also permits a direct derivation using cost-constrained random coding, which again extends to continuous-alphabet settings. The conditions (5.52)–(5.53) are attained using a near-identical analysis to the single-user setting (Section 2.6). While deriving the conditions (5.54)–(5.55) is less straightforward, the steps are still standard once the right counterpart of the non-asymptotic bound (2.81) is in place. For a fixed triplet $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})$, we are interested in the probability

$$\mathbb{P}\left[\bigcup_{i\neq 1, j\neq 1}\left\{\frac{q^n(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y})}{q^n(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})} \geq 1\right\}\right]. \tag{5.58}$$

The key idea is to avoid applying the truncated union bound $\mathbb{P}[\cup_{ij}\mathcal{E}_{ij}] \leq \min\{1, \sum_{i,j}\mathbb{P}[\mathcal{E}_{ij}]\}$ over all $(M_1-1)(M_2-1)$ incorrect message pairs simultaneously, and to instead apply it to one user's indices at a time. There are two possible orders in which this can be done; one leads to an upper bound on (5.58) of the form

$$\min\left\{1, (M_1-1)\mathbb{E}_{\overline{\boldsymbol{X}}_1}\left[\min\left\{1, (M_2-1)\mathbb{P}_{\overline{\boldsymbol{X}}_2}\left[\frac{q^n(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y})}{q^n(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})} \geq 1\right]\right\}\right]\right\} \tag{5.59}$$

and the other leads to a similar bound with the roles of $\overline{\boldsymbol{X}}_1$ and $\overline{\boldsymbol{X}}_2$ reversed. Here the subscript of $\mathbb{E}$ indicates which random variable the averaging is with respect to, and similarly for $\mathbb{P}$. Once (5.59) has been obtained, the direct derivation uses the same tools as the single-user setting (e.g., Markov's inequality), along with the standard inequality $\min\{1, \alpha\} \leq \alpha^\rho$ for $\rho \in [0, 1]$.

Another feature that distinguishes the analysis from the single-user setting is that we use cost-constrained coding with multiple auxiliary costs, as introduced in Section 3.5. Specifically, each user's codeword distribution takes the form (3.7), with two auxiliary costs per user. By doing so, we can introduce two different functions $a_1(x_1)$ in (5.52) and (5.54)–(5.55), and similarly for $a_2(x_2)$ in (5.53) and (5.54)–(5.55). It suffices to let the choices of a_1 and a_2 in (5.54)–(5.55) be identical, since the theorem statement only requires one of the two to hold.

The interested reader is referred to [81] for the details of the above-outlined techniques.

6

Multi-User Coding Techniques for Single-User Channels

6.1 Introduction

The parallel channel example of Section 5.4.1 reveals the somewhat surprising fact that multiple-access coding techniques can lead to improved achievable rates when applied to single-user channels. This suggests that other multi-user coding techniques may also be beneficial for the single-user setting, i.e., that certain types of structure in the codebook can be preferable to letting all of the codewords be independent. In this section, we explore this phenomenon in detail, with a particular emphasis on superposition coding techniques. We consider the point-to-point setup described in Section 1.1, assuming finite alphabets except where stated otherwise.

The improvements of constant-composition and cost-constrained random coding over i.i.d. random coding (*cf.*, Sections 2–3) can also be interpreted as being due to a form of structure in the code. The difference is that such improvements are a result of introducing dependencies between the symbols of individual codewords, whereas the techniques of this section introduce dependencies between the different codewords of the codebook.

The random coding constructions of this section, as well as those of Section 2, can be summarized as follows, in non-increasing order of achievable rate:

1. Refined superposition coding (Theorem 6.6);
2. Standard superposition coding (Theorem 6.3);
3. Expurgated parallel coding (Theorem 6.1);
4. Constant-composition or cost-constrained coding with independent codewords (Theorem 2.3);
5. i.i.d. coding with independent codewords (Theorem 2.2).

We will see that the gap between (1) and (2) can be strict for a given input distribution; no examples are known where the gap between (2) and (3) is strict; and the gaps between the remaining three can be strict even for an optimized input distribution.

Throughout the section, we will focus primarily on proving the achievability of primal-domain expressions. Similarly to the previous sections, the rates will exhibit ensemble tightness and permit equivalent dual expressions and continuous-alphabet extensions. However, many of these results become increasingly cumbersome to state and prove, so for the most part, we will refer the interested reader to [77, 81] for the details.

This section is predominantly based on the works of Lapidoth [58], Scarlett *et al.* [77, 81], and Somekh-Baruch [93].

6.2 Expurgated Parallel Coding

We saw in Section 5 that by applying the achievable rate of the mismatched MAC to the single-user setting, it is possible to improve on the LM rate. In particular, we saw an example where such an approach proves that the mismatch capacity equals the matched capacity, but the LM rate is strictly smaller.

While we only demonstrated this phenomenon in a single example, one can use the same idea to establish a general achievable rate: Fix any finite alphabets $\mathcal{X}_1, \mathcal{X}_2$, input distributions Q_1, Q_2, and a function

ψ: $\mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}$, and apply Theorem 5.1 to the mismatched DM-MAC $W(y|\psi(x_1, x_2))$ with metric $q(\psi(x_1, x_2), y)$. It follows that any rate $R = R_1 + R_2$ satisfying the conditions therein is achievable for the mismatched single-user channel $W(y|x)$ with metric $q(x, y)$.

It was shown by Lapidoth [58] that we can in fact do better by considering the collection of random codeword pairs $(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)})$, and keeping only the pairs whose empirical distribution is close to $Q_1 \times Q_2$. This *expurgation of non-typical codeword pairs* has a negligible impact on the coding rates, but can lead to improved conditions for vanishing error probability. One cannot apply this argument to the mismatched MAC itself, as it breaks the requirement of having two distinct codebooks (one per user). In the single-user setting, there is no such requirement, so we are free to pick subsets of the pairs $(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)})$ to our liking.

The resulting single-user achievable rate is stated as follows; we omit the proof, since in Section 6.3 we will use a simpler superposition coding technique to derive an achievable rate that can be weakened to this one.

Theorem 6.1 (Expurgated Parallel Coding Rate)**.** Consider a mismatched DMC (W, q), and fix the finite alphabets $(\mathcal{X}_1, \mathcal{X}_2)$, input distributions (Q_1, Q_2), and function ψ: $\mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}$. Then, the rate

$$R = R_1 + R_2 \tag{6.1}$$

is achievable for (W, q) provided that the pair (R_1, R_2) satisfies

$$R_1 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}\colon\ \widetilde{P}_{X_1X_2}=P_{X_1X_2},\, \widetilde{P}_{X_2Y}=P_{X_2Y} \\ \mathbb{E}_{\widetilde{P}}[\log q(\psi(X_1,X_2),Y)] \geq \mathbb{E}_P[\log q(\psi(X_1,X_2),Y)]}} I_{\widetilde{P}}(X_1; Y|X_2), \tag{6.2}$$

$$R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}\colon\ \widetilde{P}_{X_1X_2}=P_{X_1X_2},\, \widetilde{P}_{X_1Y}=P_{X_1Y} \\ \mathbb{E}_{\widetilde{P}}[\log q(\psi(X_1,X_2),Y)] \geq \mathbb{E}_P[\log q(\psi(X_1,X_2),Y)]}} I_{\widetilde{P}}(X_2; Y|X_1), \tag{6.3}$$

$$R_1 + R_2 \leq \min_{\substack{\widetilde{P}_{X_1X_2Y}\colon\ \widetilde{P}_{X_1X_2}=P_{X_1X_2},\, \widetilde{P}_{Y}=P_{Y} \\ \mathbb{E}_{\widetilde{P}}[\log q(\psi(X_1,X_2),Y)] \geq \mathbb{E}_P[\log q(\psi(X_1,X_2),Y)], \\ I_{\widetilde{P}}(X_1;Y) \leq R_1,\, I_{\widetilde{P}}(X_2;Y) \leq R_2}} I_{\widetilde{P}}(X_1, X_2; Y), \tag{6.4}$$

where $P_{X_1X_2Y}(x_1, x_2, y) = Q_1(x)Q_2(x)W(y|\psi(x_1, x_2))$.

The above rate conditions are identical to those of the achievable rate region for the MAC in Theorem 5.1, except that each minimization is further subject to $\widetilde{P}_{X_1X_2} = P_{X_1X_2}$. This results from the fact that each codeword pair is constructed to have empirical distribution close to $P_{X_1X_2} = Q_1 \times Q_2$. Since the minimization problems in the rate conditions are more constrained, the resulting rate is at least as high as that attained directly from Theorem 5.1. We also recall that the constraints $I_{\widetilde{P}}(X_1;Y) \leq R_1$ and $I_{\widetilde{P}}(X_2;Y) \leq R_2$ in (6.4) are non-standard; their main implications discussed following Theorem 5.1 also apply here.

A disadvantage of the achievable rate in Theorem 6.1 is that it has numerous auxiliary parameters, namely, $\mathcal{X}_1$, $\mathcal{X}_2$, Q_1, Q_2, and ψ. While natural choices arose in the parallel channel example of Section 5.4, it is generally difficult to jointly optimize them. This will be another advantage of superposition coding in Section 6.3, which will have fewer auxiliary parameters.

We briefly mention that equivalent dual expressions for (6.2)–(6.4) can be found in [77]; these are nearly identical to those of Theorem 5.4, except that the separate functions $a_1(x_1)$ and $a_2(x_2)$ are replaced by *joint* functions $a(x_1, x_2)$.

6.2.1 Comparison to the LM Rate

The easiest way to recover the LM rate (in its primal form (2.15)) from Theorem 6.1 is to trivially choose $|\mathcal{X}_2| = 1$, $R_2 = 0$, and $\psi(x_1, x_2) = x_1$. By doing so, conditions (6.2) and (6.4) both weaken to $R_1 \leq I_{\text{LM}}(Q_{X_1})$. In fact, under the preceding choices, we not only recover the LM rate, but the random-coding ensemble itself reduces to constant-composition random coding with independent codewords.

It is also worth noting the following general statement: Upon removing the two mutual information constraints from (6.4), the right-hand side reduces to the LM rate with input distribution Q_X equal to the marginal distribution of $(X_1, X_2, X) \sim Q_1(x)Q_2(x)\mathbb{1}\{x = \psi(x_1, x_2)\}$. This is formalized as follows.

Lemma 6.2 (Expurgated Parallel Coding and the LM Rate)**.** Under the setup of Theorem 6.1, we have

$$\min_{\substack{\widetilde{P}_{X_1X_2Y}\colon\ \widetilde{P}_{X_1X_2}=P_{X_1X_2},\,\widetilde{P}_Y=P_Y\\ \mathbb{E}_{\widetilde{P}}[\log q(\psi(X_1,X_2),Y)]\geq\mathbb{E}_P[\log q(\psi(X_1,X_2),Y)]}} I_{\widetilde{P}}(X_1,X_2;Y)$$

$$= \min_{\substack{\widetilde{P}_{XY}\colon\ \widetilde{P}_X=P_X,\widetilde{P}_Y=P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)]\geq\mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y), \tag{6.5}$$

where $P_{X_1X_2XY}(x_1,x_2,x,y)$ is given by the joint distribution $Q_1(x)Q_2(x)\mathbb{1}\{x=\psi(x_1,x_2)\}\}W(y|\psi(x_1,x_2))$.

Proof. Fixing any feasible $\widetilde{P}_{X_1X_2Y}$ on the left-hand side of (6.5), and letting $\widetilde{P}_{X_1X_2XY}$ be an arbitrary joint distribution consistent with $\widetilde{P}_{X_1X_2Y}$ such that $X=\psi(X_1,X_2)$, we have

$$I_{\widetilde{P}}(X_1,X_2;Y) = I_{\widetilde{P}}(X_1,X_2,X;Y) \tag{6.6}$$

$$= I_{\widetilde{P}}(X;Y) + I_{\widetilde{P}}(X_1,X_2;Y|X) \tag{6.7}$$

$$\geq I_{\widetilde{P}}(X;Y). \tag{6.8}$$

We now minimize both sides of this inequality over all $\widetilde{P}_{X_1X_2XY}$ such that $\widetilde{P}_{X_1X_2Y}$ is feasible on the left-hand side of (6.5), and $X=\psi(X_1,X_2)$. The left-hand side is as in (6.5), and the right-hand side yields the minimization problem

$$\min_{\substack{\widetilde{P}_{X_1X_2XY}\colon\ \widetilde{P}_{X_1X_2}=P_{X_1X_2},\,X=\psi(X_1,X_2),\,\widetilde{P}_Y=P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)]\geq\mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y). \tag{6.9}$$

Since the inequality constraint and objective depend on $\widetilde{P}_{X_1X_2XY}$ only through $\widetilde{P}_{XY}$, the first two constraints only amount to constraining $\widetilde{P}_X=P_X$, and we therefore recover the right-hand side of (6.5).

We have established that the right-hand side of (6.5) is lower-bounded by the right-hand side. However, the only step containing an inequality is (6.8), and this holds with equality when $\widetilde{P}_{X_1X_2XY}$ is chosen such that $(X_1,X_2)\to X\to Y$ forms a Markov chain. Hence, the minimizer must have this property, implying (6.5). □

Lemma 6.2 further highlights the importance of the mutual information constraints in (6.4): For all mismatched DMCs, if these constraints are removed, then we cannot hope to improve on the LM rate.

6.3 Superposition Coding

One of the most fundamental multi-user coding techniques in network information theory is superposition coding, which has been applied to broadcast channels, interference channels, and more [30]. The idea is to generate a number of independent "cloud centers" $\{\boldsymbol{U}^{(i)}\}$, and then for each corresponding index i, generate a number of codewords $\{\boldsymbol{X}^{(i,j)}\}$ that are *conditionally independent* given $\boldsymbol{U}^{(i)}$. In the degraded broadcast channel model [30, Ch. 5], this is done in a manner such that one user can identify the cloud center, while the other user can recover the precise codeword within the cloud.

In this subsection, we derive an achievable rate based on superposition coding with constant-composition codewords [81, 93]. We will show in Section 6.3.5 that this rate is at least as high as that of expurgated parallel coding (Theorem 6.1). We will also see that superposition coding requires fewer auxiliary parameters, and hence, even when a strict improvement over Theorem 6.1 is not possible, the difficult task of finding a good set of parameters may be simplified.

6.3.1 Codebook Construction

The random-coding ensemble depends on an auxiliary alphabet $\mathcal{U}$, an auxiliary codeword distribution $P_{\boldsymbol{U}} \in \mathcal{P}(\mathcal{U}^n)$, and a conditional codeword distribution $P_{\boldsymbol{X}|\boldsymbol{U}} \in \mathcal{P}(\mathcal{X}^n|\mathcal{U}^n)$. We fix a pair (R_0, R_1) and generate the codewords in two steps:

1. An auxiliary codebook $\{\boldsymbol{U}^{(i)}\}_{i=1}^{M_0}$ with $M_0 \triangleq \lfloor e^{nR_0} \rfloor$ codewords of length n is generated, with each auxiliary codeword independently drawn from $P_{\boldsymbol{U}}$.

2. For each $i = 1, \ldots, M_0$, a codebook $\{\boldsymbol{X}^{(i,j)}\}_{j=1}^{M_1}$ with $M_1 \triangleq \lfloor e^{nR_1} \rfloor$ codewords of length n is generated, with each codeword conditionally independently drawn from $P_{\boldsymbol{X}|\boldsymbol{U}}(\,\cdot\,|\boldsymbol{U}^{(i)})$.

The message m at the input to the encoder is indexed as (m_0, m_1), and for any such pair, the corresponding transmitted codeword is $\boldsymbol{X}^{(m_0,m_1)}$.

Stated concisely, we have

$$\left\{\left(\boldsymbol{U}^{(i)}, \{\boldsymbol{X}^{(i,j)}\}_{j=1}^{M_1}\right)\right\}_{i=1}^{M_0} \sim \prod_{i=1}^{M_0} P_{\boldsymbol{U}}(\boldsymbol{u}^{(i)}) \prod_{j=1}^{M_1} P_{\boldsymbol{X}|\boldsymbol{U}}(\boldsymbol{x}^{(i,j)}|\boldsymbol{u}^{(i)}). \quad (6.10)$$

We consider constant-composition coding with an input distribution $Q_{UX} \in \mathcal{P}(\mathcal{U} \times \mathcal{X})$. Specifically, letting $Q_{UX,n} \in \mathcal{P}_n(\mathcal{U} \times \mathcal{X})$ be a joint type with the same support as Q_{UX} such that $\|Q_{UX,n} - Q_{UX}\|_\infty \leq \frac{1}{n}$, we set

$$P_{\boldsymbol{U}}(\boldsymbol{u}) = \frac{1}{|\mathcal{T}^n(Q_{U,n})|} \mathbb{1}\{\boldsymbol{u} \in \mathcal{T}^n(Q_{U,n})\}, \quad (6.11)$$

$$P_{\boldsymbol{X}|\boldsymbol{U}}(\boldsymbol{x}|\boldsymbol{u}) = \frac{1}{|\mathcal{T}^n_{\boldsymbol{u}}(Q_{UX,n})|} \mathbb{1}\{(\boldsymbol{u}, \boldsymbol{x}) \in \mathcal{T}^n(Q_{UX,n})\}, \quad (6.12)$$

where $\mathcal{T}^n_{\boldsymbol{u}}(Q_{UX,n})$ is a *conditional type class*, i.e., the set of all $\boldsymbol{x}$ such that $(\boldsymbol{u}, \boldsymbol{x})$ has the specified joint type. This construction is referred to as *constant-composition superposition coding*.

One could also consider i.i.d. superposition coding, which is the standard choice in matched multi-user settings (e.g., the broadcast channel) [30]. Similarly to Section 2, this choice leads to achievable rates that are similar to the constant-composition ensemble, but generally weaker. For brevity, we focus our attention on the constant-composition case. We briefly discuss a cost-constrained ensemble for continuous-alphabet channels in Section 6.3.6.

6.3.2 Statement of Achievable Rate

We proceed by stating the achievable rate, first in the primal form and then in the dual form. The proofs are discussed in Section 6.3.6.

Theorem 6.3 (Superposition Coding Rate)**.** Consider a DMC (W, q). For any finite auxiliary alphabet $\mathcal{U}$ and input distribution $Q_{UX} \in \mathcal{P}(\mathcal{U} \times \mathcal{X})$, the rate

$$R = R_0 + R_1 \quad (6.13)$$

is achievable provided that (R_0, R_1) satisfy

$$R_1 \le \min_{\substack{\widetilde{P}_{UXY}\colon\ \widetilde{P}_{UX}=P_{UX},\, \widetilde{P}_Y=P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)]\ge \mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y|U) \tag{6.14}$$

$$R_0 + R_1 \le \min_{\substack{\widetilde{P}_{UXY}\colon\ \widetilde{P}_{UX}=P_{UX},\, \widetilde{P}_{UY}=P_{UY},\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)]\ge \mathbb{E}_P[\log q(X,Y)],\\ I_{\widetilde{P}}(U;Y)\le R_0}} I_{\widetilde{P}}(U,X;Y). \tag{6.15}$$

Moreover, this rate is ensemble-tight for constant-composition superposition coding.

We notice that these conditions bear a strong resemblance to those of Theorem 6.1. In fact, the two constructions are not as different as they may seem; both can be viewed as constructing a 2D grid of codewords, one indexed by $(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)})$, and the other by $(\boldsymbol{U}^{(i)}, \boldsymbol{X}^{(i,j)})$. The main difference is that under superposition coding, the rows are independent; this simplifies the construction and analysis, as it allows us to let all $(\boldsymbol{U}, \boldsymbol{X})$ pairs have a fixed *joint* composition without any need for expurgation.

It may appear that this extra independence of superposition coding makes it "less structured" than expurgated parallel coding, or that it weakens the rate due to a single mutual information constraint in (6.15) in contrast with the two in (6.4). However, we will see in Section 6.3.5 that Theorem 6.3 can be weakened to Theorem 6.1, and this will crucially rely on the fact that only one of the two mutual information constraints in (6.4) can be active at a time (similarly to Lemma 5.5).

The preceding rate conditions admit a similar dual formulation to Theorem 5.4, as stated in the following result. Recall that $\mathbb{E}_Z[\cdot]$ denotes averaging over some random variable Z with all other random variables held fixed.

Theorem 6.4 (Dual Form of Superposition Coding Rate)**.** For any mismatched DMC (W, q) and input distribution Q_{UX}, the rate conditions

in (6.14)–(6.15) can be written as

$$R_1 \le \sup_{s\ge 0, a(\cdot,\cdot)} \mathbb{E}\left[\log \frac{q(X,Y)^s e^{a(U,X)}}{\mathbb{E}[q(\widetilde{X},Y)^s e^{a(U,\widetilde{X})}|U,Y]}\right], \tag{6.16}$$

$$R_0 \le \sup_{\rho\in[0,1], s\ge 0, a(\cdot,\cdot)} \mathbb{E}\left[\log \frac{(q(X,Y)^s e^{a(U,X)})^\rho}{\mathbb{E}_{\overline{U}}\left[\left(\mathbb{E}_{\overline{X}}[q(\overline{X},Y)^s e^{a(\overline{U},\overline{X})}]\right)^\rho\right]}\right] - \rho R_1, \tag{6.17}$$

where $(U, X, Y, \widetilde{X}, \overline{U}, \overline{X}) \sim Q_{UX}(u,x)W(y|x)Q_{X|U}(\widetilde{x}|u)Q_{UX}(\overline{u},\overline{x})$.

We will briefly discuss a direct derivation of this dual form in Section 6.3.6, as well as its extension to continuous-alphabet channels.

6.3.3 Example 1: Zero-Undetected Error Capacity

We return to the example of Section 2.4.3, where $\mathcal{X} = \mathcal{Y} = \{0,1,2\}$, and the channel and metric are described by the entries of the matrices

$$\boldsymbol{W} = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.75 & 0.25 \\ 0.25 & 0 & 0.75 \end{bmatrix}, \qquad \boldsymbol{q} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \tag{6.18}$$

where x indexes the rows, and y indexes the columns. We saw in Section 2.4.3 that the optimized LM rate is $C_{\text{LM}} = 0.599$ bits/use, and is attained using the input distribution $Q_X = (0.449, 0.551, 0)$.

It was stated in [2] that the rate $C_{\text{LM}}^{(2)}$ obtained by considering the second-order product of the channel and metric (see Section 2.7) is equal to $C_{\text{LM}}^{(2)} = 0.616$ bits/use. Using local optimization techniques, we verified that this rate is achieved by the input distribution $(0, 0.250, 0, 0.319, 0, 0, 0, 0.181, 0.250)$, where the order of the inputs is $(0,0), (0,1), (0,2), (1,0), \ldots, (2,2)$.

The global optimization of (6.14)–(6.15) over $\mathcal{U}$ and Q_{UX} appears to be difficult. However, setting $|\mathcal{U}| = 2$ and applying local optimization techniques using a number of starting points, we obtained an achievable rate of $R_{\text{SC}}^* = 0.695$ bits/use, with $Q_U = (0.645, 0.355)$, $Q_{X|U}(\cdot|1) = (0.3, 0.7, 0)$, and $Q_{X|U}(\cdot|2) = (0,0,1)$. Hence, even if this is not the globally optimal choice of parameters, it reveals the interesting fact

that superposition coding not only yields a strict improvement over the single-letter LM rate, but also over the two-letter version.

6.3.4 Example 2: The LM Rate is Not Tight for Binary-Input DMCs

It was claimed in an early work of Balakirsky [14] that for binary-input DMCs, the LM rate with an optimized input distribution is tight, i.e., $C_{\mathrm{M}} = C_{\mathrm{LM}}$. However, a recent work of the present authors provided a counter-example to this claim [76], which we describe in the following. We let $\mathcal{X} = \{0, 1\}$ and $\mathcal{Y} = \{0, 1, 2\}$, and consider the channel and metric described by the entries of the $|\mathcal{X}| \times |\mathcal{Y}|$ matrices

$$\boldsymbol{W} = \begin{bmatrix} 0.97 & 0.03 & 0 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \quad \boldsymbol{q} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0.5 & 1.36 \end{bmatrix}. \tag{6.19}$$

LM Rate. Using numerical evaluations and a brute force search over Q_X, along with theoretical results quantifying the approximation errors incurred, it was shown in [76] that the LM rate with an optimized input distribution, $C_{\mathrm{LM}} = \max_{Q_X} I_{\mathrm{LM}}(Q_X)$, satisfies

$$0.19746 \leq C_{\mathrm{LM}} \leq 0.19751 \quad \text{bits/use}. \tag{6.20}$$

The optimal input distribution, rounded to five decimal places, is given by $Q_X = (0.75597, 0.24403)$.

Superposition Coding Rate. When we applied the superposition coding rate of Theorem 6.3 directly to (W, q), we only managed to obtain a rate coinciding with C_{LM}. However, the technique of passing to the k-th order product alphabet, introduced in Section 2.7, is equally valid for other random-coding rates, including the superposition coding rate of Theorem 6.3. We consider the case $k = 2$, and denote the resulting rate with an optimized input distribution Q_{UX^2} as $C_{\mathrm{SC}}^{(2)}$. The counter-example to [14] follows by combining (6.20) with the following:

$$C_{\mathrm{SC}}^{(2)} \geq 0.19908 \quad \text{bits/use}, \tag{6.21}$$

which clearly implies $C_{\mathrm{M}} > C_{\mathrm{LM}}$.

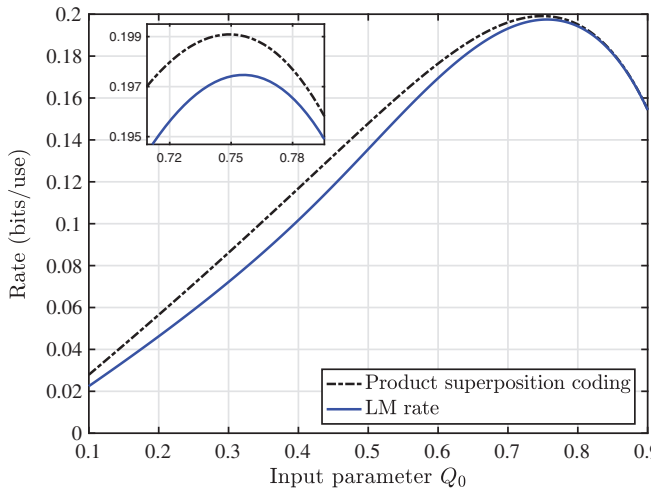


Figure 6.1: Binary-input example: LM rate and second-order superposition coding rate as a function of Q_0 for the mismatched channel described by (6.19).

To establish (6.21), we set $|\mathcal{U}| = 2$, and choose the input distribution Q_{UX^2} on $\mathcal{U} \times \mathcal{X}^2$ as follows:

$$Q_U = \begin{bmatrix} 1 - Q_1^2 & Q_1^2 \end{bmatrix} \tag{6.22}$$

$$Q_{X^2|U=0} = \frac{1}{1 - Q_1^2} \begin{bmatrix} Q_0^2 & Q_0 Q_1 & Q_0 Q_1 & 0 \end{bmatrix} \tag{6.23}$$

$$Q_{X^2|U=1} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.24}$$

with $Q_0 = 0.749$ and $Q_1 = 0.251$. Here, the alphabet $\mathcal{X}^2$ is listed in the order $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$. Note that this choice of Q_{UX^2} ensures that the marginal distribution Q_{X^2} is the two-fold product of $Q_X = (Q_0, Q_1)$.

Other Input Distributions. The construction in (6.24) is valid for any pair (Q_0, Q_1) summing to one, and it is interesting to compare the resulting rate with the LM rate applied to $Q_X = (Q_0, Q_1)$. The rates are shown in Figure 6.1, where we observe that superposition coding consistently improves on the LM rate for the values of Q_0 shown.

6.3.5 Comparison to Expurgated Parallel Coding

The following lemma formally states that Theorem 6.3 provides an achievable rate at least as high as that of Theorem 6.1. No examples are known where a strict improvement is attained, suggesting that the two rates could be equivalent. Superposition coding, however, has the advantage of having fewer parameters (namely, $(\mathcal{U}, Q_{UX})$ instead of $(\mathcal{X}_1, \mathcal{X}_2, \psi, Q_1, Q_2)$), and in the authors' experience, this usually makes it easier to find good choices (e.g., based on intuition, or local optimization techniques).

Lemma 6.5 (Comparison of Superposition Coding and Expurgated Parallel Coding)**.** For any mismatched DMC (W, q), we have

$$\sup_{\mathcal{X}_1,\mathcal{X}_2,\psi,Q_1,Q_2} \max_{(R_1,R_2)\in\mathcal{R}_{\text{MAC}}} R_1 + R_2 \le \sup_{\mathcal{U},Q_{UX}} \max_{(R_0,R_1)\in\mathcal{R}_{\text{SC}}} R_0 + R_1, \quad (6.25)$$

where $\mathcal{R}_{\text{MAC}} = \mathcal{R}_{\text{MAC}}(\mathcal{X}_1, \mathcal{X}_2, \psi, Q_1, Q_2)$ is the set of rate pairs satisfying (6.2)–(6.4), and $\mathcal{R}_{\text{SC}} = \mathcal{R}_{\text{SC}}(\mathcal{U}, Q_{UX})$ is the set of rate pairs satisfying (6.14)–(6.15).

We only outline the proof here, and refer the reader to [77] for the full details. First, analogously to Lemma 5.5, the condition in (6.4) holds if and only if at least one of the following hold:

$$R_1 + R_2 \le \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1X_2}=P_{X_1X_2},\widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(\psi(X_1,X_2),Y)]\ge\mathbb{E}_P[\log q(\psi(X_1,X_2),Y)], \\ I_{\widetilde{P}}(X_1;Y)\le R_1}} I_{\widetilde{P}}(X_1, X_2; Y) \quad (6.26)$$

$$R_1 + R_2 \le \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1X_2}=P_{X_1X_2},\widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(\psi(X_1,X_2),Y)]\ge\mathbb{E}_P[\log q(\psi(X_1,X_2),Y)], \\ I_{\widetilde{P}}(X_2;Y)\le R_2}} I_{\widetilde{P}}(X_1, X_2; Y). \quad (6.27)$$

In other words, the two mutual information constraints in (6.4) can be considered one at a time without affecting the rates.

Once this is established, the proof proceeds in two parts:

1. Show that by identifying $U^{\text{sc}} = X_2^{\text{ex}}$, $R_1^{\text{sc}} = R_1^{\text{ex}}$ and $R_0^{\text{sc}} = R_2^{\text{ex}}$, we can weaken (6.14) to (6.2), and (6.15) to (6.27).

2. Show that by identifying $U^{\mathrm{sc}} = X_1^{\mathrm{ex}}$, $R_1^{\mathrm{sc}} = R_2^{\mathrm{ex}}$ and $R_0^{\mathrm{sc}} = R_1^{\mathrm{ex}}$, we can weaken (6.14) to (6.3), and (6.15) to (6.26).

Here, the superscript explicitly denotes which coding technique the parameters correspond to. The above two claims are proved using identical steps, so we consider the former. We choose the input distribution in terms of the expurgated parallel coding parameters as

$$Q_{UX}(u,x) = \sum_{x_1} Q_1(x_1)Q_2(u)\mathbb{1}\{x = \psi(x_1,u)\}. \tag{6.28}$$

By doing so, simple manipulations to the objective functions and constraints in (6.14) and (6.15) allow us to recover the conditions (6.2) and (6.27). For example, if we substitute this choice of Q_{UX} into the superposition coding metric constraint

$$\sum_{u,x,y} Q_{UX}(u,x)\widetilde{P}_{Y|UX}(y|u,x)\log q(x,y) \geq \sum_{x,y} Q_X(x)W(y|x)\log q(x,y), \tag{6.29}$$

then we get the following upon renaming u as x_2:

$$\sum_{x_1,x_2,y} Q_1(x_1)Q_2(x_2)\widetilde{P}_{Y|X_1X_2}(y|x_1,x_2)\log q(\psi(x_1,x_2),y) \geq \sum_{x_1,x_2,y} Q_1(x_1)Q_2(x_2)W(y|\psi(x_1,x_2))\log q(\psi(x_1,x_2),y), \tag{6.30}$$

where $\widetilde{P}_{Y|X_1X_2}(y|x_1,x_2) = \widetilde{P}_{Y|X_1X}(y|x_1,\psi(x_1,x_2))$. This is precisely the metric constraint for expurgated parallel coding. The desired claim follows by treating the other constraints and objective functions similarly, and then minimizing over $\widetilde{P}_{Y|X_1X_2}$.

We note that a similar argument is also given [93, Prop. 1], in which a form of superposition coding is considered with MAC-like parameters $\mathcal{X}_1$, $\mathcal{X}_2$, and ψ, as opposed to the parameters $\mathcal{U}$ and Q_{UX} considered here. The parameters $(\mathcal{X}_1, \mathcal{X}_2, \psi)$ are viewed as a forming a *cognitive* MAC in which one user knows both messages, meaning that superposition coding can be used to induce general joint input distributions $Q_{X_1X_2}$, rather than only product distributions of the form $Q_{X_1} \times Q_{X_2}$. An advantage of this perspective is that it permits the direct comparison of superposition coding and expurgated parallel coding for *fixed* parameters $(\mathcal{X}_1, \mathcal{X}_2, \psi)$.

In particular, the following argument from [93] highlights that the superposition coding rates can indeed be strictly higher with fixed parameters: Fix $(\mathcal{X}_1, \mathcal{X}_2)$ and $\mathcal{X}$ with $|\mathcal{X}| = |\mathcal{X}_1| \cdot |\mathcal{X}_2|$, and let ψ be any one-to-one mapping from $\mathcal{X}_1 \times \mathcal{X}_2$ to $\mathcal{X}$. Then, varying $Q_{X_1X_2}$ can induce any distribution on X, but varying Q_{X_1} and Q_{X_2} can only induce a narrower range of distributions (e.g., any symbol having zero probability forces others to also be zero, due to the product structure). This restriction lowers the achievable rate, for example, when matched decoding ($q = W$) is used and the capacity-achieving input distribution does not conform to the product structure.

6.3.6 Proof Techniques for the Achievable Rate

The proof of Theorem 6.3 is comparatively simpler than that of the multiple-access channel (Theorem 5.1), so we focus mostly on the ideas and intuition rather than the full details.

Separation into Two Error Events. Recalling the codebook construction in Section 6.3.1, we assume without loss of generality that the transmitted codeword corresponds to $(m_0, m_1) = (1, 1)$, and we define the following two error events:

$$\begin{aligned} &\text{(Type 0)} \quad q^n(\boldsymbol{X}^{(i,j)}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y}) \text{ for some } i \neq 1,\ j; \\ &\text{(Type 1)} \quad q^n(\boldsymbol{X}^{(1,j)}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y}) \text{ for some } j \neq 1. \end{aligned}$$

Denoting the probabilities of these events by $\bar{p}_{\mathrm{e},0}(n, M_0, M_1)$ and $\bar{p}_{\mathrm{e},1}(n, M_1)$ respectively, it follows that the overall random-coding error probability $\bar{p}_{\mathrm{e}}(n, M_0, M_1)$ satisfies

$$\max\{\bar{p}_{\mathrm{e},0}, \bar{p}_{\mathrm{e},1}\} \leq \bar{p}_{\mathrm{e}} \leq \bar{p}_{\mathrm{e},0} + \bar{p}_{\mathrm{e},1}. \tag{6.31}$$

Non-Asymptotic Bounds. The first step of the analysis is to establish the following non-asymptotic bounds:

$$\bar{p}_{\mathrm{e},0} \leq \mathbb{E}_{\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y}}\Big[\min\Big\{1, (M_0 - 1)\mathbb{E}_{\overline{\boldsymbol{U}}}\Big[\min\{1, M_1 \mathbb{P}_{\overline{\boldsymbol{X}}}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y})]\}\Big]\Big\}\Big], \tag{6.32}$$

$$\bar{p}_{\mathrm{e},1} \le \mathbb{E}\Big[\min\Big\{1, (M_1-1)\mathbb{P}[q^n(\widetilde{\boldsymbol{X}},\boldsymbol{Y}) \ge q^n(\boldsymbol{X},\boldsymbol{Y}) \,|\, \boldsymbol{U},\boldsymbol{X},\boldsymbol{Y}]\Big\}\Big], \tag{6.33}$$

where

$$\begin{aligned}(\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y},\widetilde{\boldsymbol{X}},\overline{\boldsymbol{U}},\overline{\boldsymbol{X}}) \sim & P_{\boldsymbol{U}}(\boldsymbol{u})P_{\boldsymbol{X}|\boldsymbol{U}}(\boldsymbol{x}|\boldsymbol{u})W^n(\boldsymbol{y}|\boldsymbol{x}) \\ & \times P_{\boldsymbol{X}|\boldsymbol{U}}(\widetilde{\boldsymbol{x}}|\boldsymbol{u})P_{\boldsymbol{U}}(\overline{\boldsymbol{u}})P_{\boldsymbol{X}|\boldsymbol{U}}(\overline{\boldsymbol{x}}|\overline{\boldsymbol{u}}).\end{aligned} \tag{6.34}$$

For the type-1 error probability, (6.33) follows by conditioning on $(\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y})$ and applying the truncated union bound (i.e., the minimum of one and the standard union bound) over the error events. On the other hand, more care is needed for the type-0 error probability. We first write

$$\bar{p}_{\mathrm{e},0} = \mathbb{E}\Bigg[\mathbb{P}\Big[\bigcup_{i\neq 1, j\neq 1}\Big\{q^n(\boldsymbol{X}^{(i,j)},\boldsymbol{Y}) \ge q^n(\boldsymbol{X},\boldsymbol{Y})\Big\}\,\Big|\,\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y}\Big]\Bigg], \tag{6.35}$$

where we assume without loss of generality that $(m_0,m_1)=(1,1)$. The upper bound (6.32) then follows by applying the truncated union bound first to the union over i, and then to the union over j, similarly to the analysis of the DM-MAC in Section 5.7.

Primal Analysis Techniques. Once the above bounds are in place, we can apply similar techniques to the single-user setting (*cf.*, Section 2). Focusing on the more interesting type-0 error event, we write (6.32) in terms of joint types as follows:

$$\begin{aligned}\bar{p}_{\mathrm{e},0} \le & \sum_{\substack{P_{UXY}\in\mathcal{P}_n(\mathcal{U}\times\mathcal{X}\times\mathcal{Y}):\\ P_{UX}=Q_{UX,n}}} \mathbb{P}\Big[(\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y})\in\mathcal{T}^n(P_{UXY})\Big] \\ & \times \min\Bigg\{1, (M_0-1)\sum_{\substack{\widetilde{P}_{UXY}\in\mathcal{P}_n(\mathcal{U}\times\mathcal{X}\times\mathcal{Y}):\\ P_{UX}=Q_{UX,n},\,\widetilde{P}_Y=P_Y\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)]\ge\mathbb{E}_P[\log q(X,Y)]}} \mathbb{P}\Big[(\overline{\boldsymbol{U}},\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{UY})\Big] \\ & \qquad \times \min\Big\{1, M_1\mathbb{P}\Big[(\overline{\boldsymbol{u}},\overline{\boldsymbol{X}},\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{UXY})\,|\,\overline{\boldsymbol{U}}=\overline{\boldsymbol{u}}\Big]\Big\}\Bigg\},\end{aligned} \tag{6.36}$$

where $(\overline{\boldsymbol{u}}, \boldsymbol{y})$ denotes an arbitrary pair such that $\boldsymbol{y} \in \mathcal{T}^n(P_Y)$ and $(\overline{\boldsymbol{u}}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{UY})$. Here the constraints on (U, X) arise due the constant-composition construction, and the constraint $\widetilde{P}_Y = P_Y$ arises since $(\boldsymbol{U}, \boldsymbol{X}, \boldsymbol{Y})$ and $(\overline{\boldsymbol{U}}, \overline{\boldsymbol{X}}^{(j)}, \boldsymbol{Y})$ share the same $\boldsymbol{Y}$ sequence.

The proof now proceeds as in the previous sections by applying standard properties of types, using the law of large numbers to establish $\|P_{UXY} - Q_{UX} \times W\|_\infty \le \delta$, and applying a continuity argument similar to Appendix A.1. The probabilities in the second and third lines of (6.36) give rise to the mutual information terms $I_{\widetilde{P}}(U;Y)$ and $I_{\widetilde{P}}(U;Y|X)$ respectively, and the resulting rate condition is

$$R_0 \le \min_{\substack{\widetilde{P}_{UXY}\colon\ \widetilde{P}_{UX}=P_{UX},\, \widetilde{P}_{UY}=P_{UY},\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \ge \mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(U;Y) + [I_{\widetilde{P}}(X;Y|U) - R_1]_+. \quad (6.37)$$

This condition is trivially satisfied when $I_{\widetilde{P}}(U;Y) > R_0$, so we can further constrain $I_{\widetilde{P}}(U;Y) \le R_0$ in the minimization, and then lower bound the function $[\cdot]^+$ by its argument to obtain (6.15).

Ensemble Tightness. The ensemble tightness proof for the type-1 error probability (i.e., the condition (6.14)) directly follows that of the single-user setting. For the type-0 error probability (i.e., the condition (6.15)), we need to avoid the two union bounds used in weakening (6.35) to (6.32). However, this is also straightforward: For the union over i we instead resort to the exact expression similarly to the single-user setting (*cf.*, (2.109)), and for the union over j, it suffices to note that for independent events the truncated union bound is tight to within a factor of $\frac{1}{2}$ (e.g., see [90, Lemma A.2]).

Dual Analysis Techniques. To obtain the dual expression of Theorem 6.4, we can weaken (6.32) and (6.33) using standard tools such as Markov's inequality and $\min\{1, z\} \le z^\rho$, as well as the fact that for any function $a(u, x)$, the quantity $\sum_{i=1}^n a(u_i, x_i)$ takes the same value for any $(\boldsymbol{U}, \boldsymbol{X})$ occurring with positive probability. Beyond handling the unions one at a time to obtain (6.32), no new tools are needed compared to the single-user dual analysis for constant-composition random coding (*cf.*, Section 2).

Similarly to Section 3, it is of interest to extend the dual analysis to channels with continuous alphabets. To do so, we need to define $P_{\boldsymbol{U}}$ and $P_{\boldsymbol{X}|\boldsymbol{U}}$ in terms of an input distribution Q_{UX} and auxiliary cost $a(u,x)$ in such a way that (i) $\sum_{i=1}^{n} a(U_i, X_i)$ is almost surely close to its mean, and (ii) $P_{\boldsymbol{U}}$ and $P_{\boldsymbol{X}|\boldsymbol{U}}$ can be upper bounded by i.i.d. distributions on Q_U and $Q_{X|U}$ times sub-exponential pre-factors. It was shown in [77] that such distributions exist, requiring only simple tricks following the cost-constrained codeword distributions used in Section 3.

6.4 Refined Superposition Coding

Superposition coding leads to a structured codebook in which the codewords $\{\boldsymbol{X}^{(i,j)}\}_{j=1}^{M_1}$ are correlated through their dependence on $\boldsymbol{U}^{(i)}$, and we have seen that this can lead to strict improvements over independent codewords. In this subsection, we present a variation of this idea that adds another level of structure, namely, the codewords $\{\boldsymbol{X}^{(i,j)}\}_{j=1}^{M_1}$ are dependent even conditioned on $\boldsymbol{U}^{(i)}$. We call the technique *refined superposition coding* [81], and we will see that it yields a rate at least as good as the standard version, with strict improvements being possible (at least for fixed $(\mathcal{U}, Q_{UX})$).

6.4.1 Codebook Construction

As with superposition coding, the main parameters to the ensemble are the finite alphabet $\mathcal{U}$ and input distribution $Q_{UX} \in \mathcal{P}(\mathcal{U} \times \mathcal{X})$. In addition, we fix the rates R_0 and $\{R_{1u}\}_{u \in \mathcal{U}}$, and accordingly define $M_0 \triangleq \lfloor e^{nR_0} \rfloor$ and $M_{1u} \triangleq \lfloor e^{nR_{1u}} \rfloor$.

The codewords are generated in three steps:

- Generate the length-n auxiliary codewords $\{\boldsymbol{U}^{(i)}\}_{i=1}^{M_0}$ independently according to the following constant-composition distribution $P_{\boldsymbol{U}}$:

$$P_{\boldsymbol{U}}(\boldsymbol{u}) = \frac{1}{|\mathcal{T}^n(Q_{U,n})|} \mathbb{1}\Big\{\boldsymbol{u} \in \mathcal{T}^n(Q_{U,n})\Big\}, \tag{6.38}$$

  where $Q_{U,n} \in \mathcal{P}_n(\mathcal{X})$ is a type with the same support as Q_U such that $\|Q_{U,n} - Q_U\|_\infty \le \frac{1}{n}$.

- For each $u \in \mathcal{U}$, we define

$$n_u \triangleq Q_{U,n}(u)n \tag{6.39}$$

and fix a partial codeword distribution $P_{\boldsymbol{X}_u} \in \mathcal{P}(\mathcal{X}^{n_u})$. For each $i = 1, \ldots, M_0$ and $u \in \mathcal{U}$, we draw the length-n_u partial codewords $\{\boldsymbol{X}_u^{(i,j_u)}\}_{j_u=1}^{M_{1u}}$ independently from $P_{\boldsymbol{X}_u}$.

- A given message m is indexed as $(m_0, m_{11}, \ldots, m_{1|\mathcal{U}|})$, and the corresponding codeword $\boldsymbol{X}_{(m)}$ is constructed by treating $\boldsymbol{U}^{(m_0)}$ as a time-sharing sequence: At the indices where $\boldsymbol{U}^{(m_0)}$ equals u, we transmit the symbols of $\boldsymbol{X}_u^{(m_0,m_{1u})}$.

The first two steps are summarized as follows:

$$\left\{\left(\boldsymbol{U}^{(i)}, \{\boldsymbol{X}_1^{(i,j_1)}\}_{j_1=1}^{M_{11}}, \ldots, \{\boldsymbol{X}_{|\mathcal{U}|}^{(i,j_{|\mathcal{U}|})}\}_{j_{|\mathcal{U}|}=1}^{M_{1|\mathcal{U}|}}\right)\right\}_{i=1}^{M_0} \sim \prod_{i=1}^{M_0}\left(P_{\boldsymbol{U}}(\boldsymbol{u}^{(i)}) \prod_{u\in\mathcal{U}}\left(\prod_{j_u=1}^{M_{1u}} P_{\boldsymbol{X}_u}(\boldsymbol{x}_u^{(i,j_u)})\right)\right). \tag{6.40}$$

From the third step, there are $M = M_0 \times \prod_{u=1}^{|\mathcal{U}|} M_{1u}$ codewords, and hence the overall rate is

$$R = R_0 + \sum_{u=1}^{|\mathcal{U}|} Q_{U,n}(u) R_{1u}. \tag{6.41}$$

An example of the construction of the codeword $\boldsymbol{x}$ from the auxiliary sequence $\boldsymbol{u}$ and partial codewords $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$ is shown in Figure 6.2, where we have $\mathcal{U} = \{1, 2, 3\}$ and $\mathcal{X} = \{a, b, c\}$.

It remains to specify the partial codeword distributions $P_{\boldsymbol{X}_u}$. Letting $Q_{UX,n}$ be a joint type with the same support as Q_{UX} satisfying $\|Q_{UX,n} - Q_{UX}\|_\infty \le \frac{1}{n}$, we choose $P_{\boldsymbol{X}_u}$ to be the uniform distribution on the type class $\mathcal{T}^{n_u}(Q_{X|U,n}(\cdot|u))$:

$$P_{\boldsymbol{X}_u}(\boldsymbol{x}_u) = \frac{1}{|\mathcal{T}^{n_u}(Q_{X|U,n}(\cdot|u))|} \mathbb{1}\{\boldsymbol{x}_u \in \mathcal{T}^{n_u}(Q_{X|U,n}(\cdot|u))\}. \tag{6.42}$$

Combining this with (6.38), we have by symmetry that each pair $(\boldsymbol{U}^{(i)}, \boldsymbol{X}^{(i,j_1,\ldots,j_{|\mathcal{U}|})})$ is uniformly distributed on $\mathcal{T}^n(Q_{UX})$. Hence, the

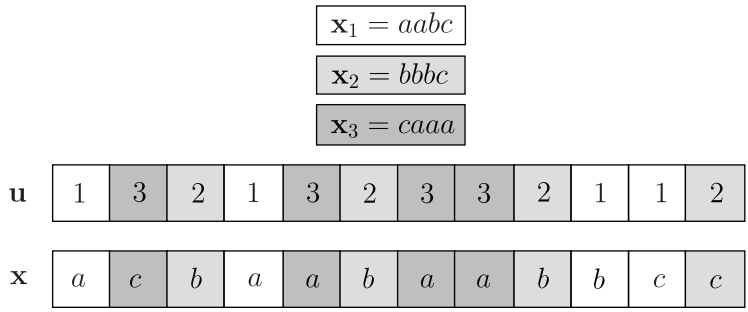


Figure 6.2: The construction of the codeword from the auxiliary sequence $\boldsymbol{u}$ and the partial codewords $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$ for refined SC. Here we have $\mathcal{U} = \{1, 2, 3\}$, $\mathcal{X} = \{a, b, c\}$, $n_1 = n_2 = n_3 = 4$, and $n = 12$.

marginal distribution of each $(\boldsymbol{U}, \boldsymbol{X})$ pair is the same as superposition coding, but the joint codebook distribution has some additional structure. We refer to the overall construction as *constant-composition refined superposition coding*.

Considering the case $|\mathcal{U}| = 2$, one can picture the structure of the codewords $\{\boldsymbol{X}^{(i,j_1,j_2)}\}_{j_1,j_2}$ by imagining them arranged in an $M_{11} \times M_{12}$ grid. Given $\{\boldsymbol{U}^{(i)}\}_{i=1}^{M_0}$, refined superposition coding generates two separate codebooks of partial codewords $\{\boldsymbol{X}_1\}_{j_1=1}^{M_{11}}$ and $\{\boldsymbol{X}_2\}_{j_2=1}^{M_{12}}$ and fills in the grid accordingly, whereas standard superposition coding fills every entry of the grid in a conditionally independent manner given $\{\boldsymbol{U}^{(i)}\}_{i=1}^{M_0}$. As a result, the added structure is analogous to that of using multiple-access coding (Section 5) in place of mutually independent codewords (Section 2).

6.4.2 Statement of Achievable Rate

The achievable rate of constant-composition refined superposition coding is stated as follows [81]. We make use of the definition of the LM rate in (2.15).

Theorem 6.6 (Refined Superposition Coding Rate)**.** Consider a DMC (W, q). For any finite set $\mathcal{U}$ and input distribution $Q_{UX} \in \mathcal{P}(\mathcal{U} \times \mathcal{X})$,

the rate

$$R = R_0 + \sum_{u=1}^{|\mathcal{U}|} Q_U(u) R_{1u} \tag{6.43}$$

is achievable provided that R_0 and $\{R_{1u}\}_{u=1}^{|\mathcal{U}|}$ satisfy

$$R_{1u} \le I_{\text{LM}}\big(Q_{X|U}(\cdot|u)\big), \quad \forall u \in \mathcal{U}, \tag{6.44}$$

$$\begin{aligned} R_0 \le & \min_{\substack{\widetilde{P}_{UXY}:\ \widetilde{P}_{UX}=P_{UX},\, \widetilde{P}_{UY}=P_{UY},\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \ge \mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(U;Y) \\ & + \left[\max_{\mathcal{K}\subseteq\mathcal{U}, \mathcal{K}\neq\emptyset} \sum_{u\in\mathcal{K}} Q_U(u) \Big(I_{\widetilde{P}}(X;Y|U=u) - R_{1u} \Big) \right]^+ . \end{aligned} \tag{6.45}$$

Moreover, this rate is ensemble-tight for constant-composition refined superposition coding.

We will show in Section 6.4.5 that this rate is at least high as that of that of standard superposition coding (Theorem 6.3) for any pair $(\mathcal{U}, Q_{UX})$ of auxiliary parameters.

A dual formulation of Theorem 6.6 was given in [81] for the special case $|\mathcal{U}| = 2$, along with a direct derivation using cost-constrained coding that generalizes to continuous-alphabet channels. The statement and proof of the dual equivalence are rather cumbersome even in the simplest case $|\mathcal{U}| = 2$, and they are therefore omitted here.

Before comparing to standard superposition and discussing the proof of Theorem 6.6, we present some examples from [81].

6.4.3 Example 1: Sum Channel

The notions of *product channels* and *sum channels* were introduced in Shannon's original paper [87]. We have already studied product channels (i.e., parallel channels) in detail in Sections 2.4.5 and 5.4; here we provide an analogous example regarding the sum channel.

Given two channels (W_1, W_2) respectively defined on the alphabets $(\mathcal{X}_1, \mathcal{Y}_1)$ and $(\mathcal{X}_2, \mathcal{Y}_2)$, the *sum channel* is defined to be the channel $W(y|x)$ with $|\mathcal{X}| = |\mathcal{X}_1| + |\mathcal{X}_2|$ and $|\mathcal{Y}| = |\mathcal{Y}_1| + |\mathcal{Y}_2|$ such that one of the two subchannels is used on each transmission. One can similarly

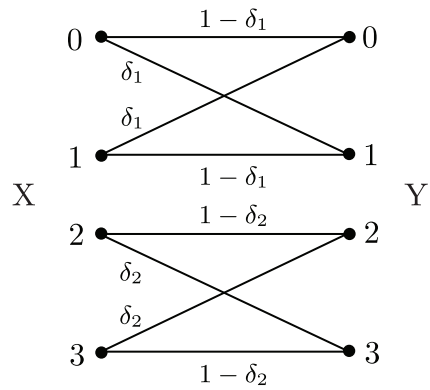


Figure 6.3: Sum of two binary symmetric channels with crossover probabilities (δ_1, δ_2).

combine two metrics $q_1(x_1, y_1)$ and $q_2(x_2, y_2)$ to form a sum metric $q(x, y)$: Assuming without loss of generality that $\mathcal{X}_1$ and $\mathcal{X}_2$ are disjoint and $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are disjoint, we have

$$q(x,y) = \begin{cases} q_1(x_1, y_1) & x_1 \in \mathcal{X}_1 \text{ and } y_1 \in \mathcal{Y}_1 \\ q_2(x_2, y_2) & x_2 \in \mathcal{X}_2 \text{ and } y_2 \in \mathcal{Y}_2 \\ 0 & \text{otherwise,} \end{cases} \tag{6.46}$$

and similarly for $W(y|x)$.

An example of a sum channel is shown in Figure 6.3, where both subchannels are binary symmetric channels (BSCs). The corresponding channel and metric matrices are given by

$$\boldsymbol{W} = \begin{bmatrix} 1-\delta_1 & \delta_1 & 0 & 0 \\ \delta_1 & 1-\delta_1 & 0 & 0 \\ 0 & 0 & 1-\delta_2 & \delta_2 \\ 0 & 0 & \delta_2 & 1-\delta_2 \end{bmatrix}. \tag{6.47}$$

While the channel diagram is similar to that of the product channel of Figure 2.3, the two are fundamentally different: Here we are not transmitting over both subchannels simultaneously, but rather, choosing one to transmit over at each time instant. To appreciate the difference, observe that we can transmit one bit across the sum channel with *zero*

error by using only the symbols $\{0, 2\}$, whereas we can never achieve an error probability of zero in the product channel.

We use Theorem 6.6 to provide an achievable rate for a general sum channel, and then specialize it to the BSC example. Let Q_1 and Q_2 be the distributions that maximize the LM rate in (2.15) on the two subchannels. We set $\mathcal{U} = \{1, 2\}$, $Q_{X|U}(\cdot|1) = (Q_1, \mathbf{0})$ and $Q_{X|U}(\cdot|2) = (\mathbf{0}, Q_2)$, where $\mathbf{0}$ denotes the zero vector of length two. We leave Q_U to be specified.

Combining the constraints $\widetilde{P}_{UX} = Q_{UX}$ and $\mathbb{E}_{\widetilde{P}}[\log q(X, Y)] \geq \mathbb{E}_P[\log q(X, Y)]$ in (6.45), it can be shown that the minimizing $\widetilde{P}_{UXY}$ only takes non-zero values for (u, x, y) such that (i) $u = 1$, $x \in \mathcal{X}_1$ and $y \in \mathcal{Y}_1$, or (ii) $u = 2$, $x \in \mathcal{X}_2$ and $y \in \mathcal{Y}_2$. It follows that U is a deterministic function of Y under the minimizing $\widetilde{P}_{UXY}$, and hence

$$I_{\widetilde{P}}(U; Y) = H(Q_U) - H_{\widetilde{P}}(U|Y) = H(Q_U). \tag{6.48}$$

Therefore, the right-hand side of (6.45) is lower bounded by $H(Q_U)$. Using (6.43), it follows that we can achieve the rate

$$R^* = H(Q_U) + Q_U(1) I_{\mathrm{LM}}(Q_1, W_1) + Q_U(2) I_{\mathrm{LM}}(Q_2, W_2) \tag{6.49}$$

$$= \log\left(e^{I_{\mathrm{LM}}(Q_1, W_1)} + e^{I_{\mathrm{LM}}(Q_2, W_2)}\right), \tag{6.50}$$

where $I_{\mathrm{LM}}(Q_u, W_u)$ is the LM rate for subchannel $u \in \{1, 2\}$, and (6.50) holds under the optimal auxiliary distribution $Q_U(1) = 1 - Q_U(2) = \frac{e^{I_{\mathrm{LM}}(Q_1, W_1)}}{e^{I_{\mathrm{LM}}(Q_1, W_1)} + e^{I_{\mathrm{LM}}(Q_2, W_2)}}$, analogously to the matched case [87, Sec. 16].

Now, suppose that the two subchannels are BSCs with different crossover probabilities $\delta_1, \delta_2 \in \left(0, \frac{1}{2}\right)$, whereas the mismatched decoder assumes a common crossover probability $\delta \in \left(0, \frac{1}{2}\right)$. By the binary example in Section 2.4.1, we have for $u = 1, 2$ that the choice $Q_u = \left(\frac{1}{2}, \frac{1}{2}\right)$ yields $I_{\mathrm{LM}}(Q_u, W_u) = 1 - H_2(\delta_u)$ bits/use, which equals the matched capacity C_u of the subchannel. As a result, from (6.50) and the sum channel capacity formula $C^* = \log\left(e^{C_1} + e^{C_2}\right)$ [87, Sec. 16], the mismatch capacity is identical to the matched capacity. In contrast, it can be shown that the LM rate is strictly less than the matched capacity whenever $\delta_1 \neq \delta_2$. We illustrate this numerically in Figure 6.4, where we plot the matched rate (achieved by refined SC), as well as the LM rate under the (matched) capacity-achieving input distribution.

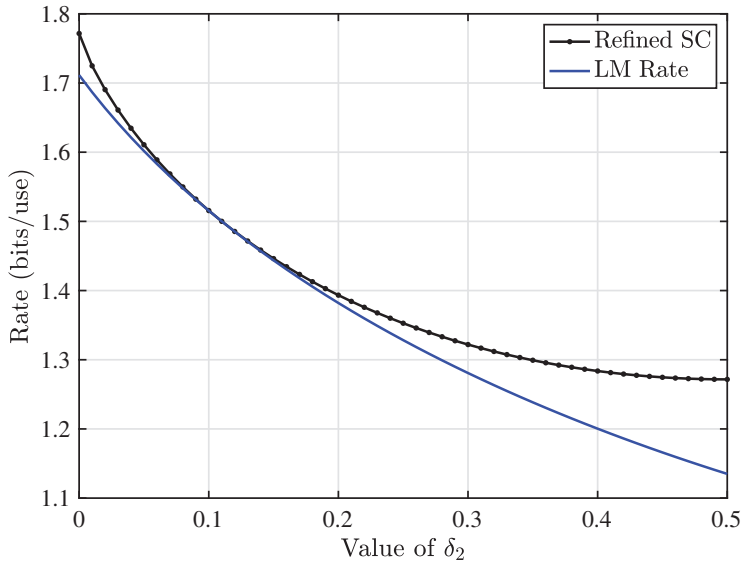


Figure 6.4: Sum channel example: Achievable rates with $\delta_1 = 0.11$, as a function of δ_2. The refined SC rate and the LM rate only coincide when $\delta_1 = \delta_2$.

6.4.4 Example 2: Comparison of Superposition Coding Methods

We now turn to an example comparing the standard superposition coding rate of Theorem 6.3 with the refined superposition coding rate of Theorem 6.6. We consider the channel and decoding metric described by the entries of the matrices

$$\boldsymbol{W} = \begin{bmatrix} 0.99 & 0.01 & 0 & 0 \\ 0.01 & 0.99 & 0 & 0 \\ 0.1 & 0.1 & 0.7 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.7 \end{bmatrix}, \qquad \boldsymbol{q} = \begin{bmatrix} 1 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0.05 & 0.15 & 1 & 0.05 \\ 0.15 & 0.05 & 0.5 & 1 \end{bmatrix}. \tag{6.51}$$

This is an admittedly artificial example, but we will see that it constitutes an interesting example for the purpose of comparison.

Using an exhaustive search to three decimal places, we obtained the following: (i) the optimized LM rate is $C_{\text{LM}} = 1.111$ bits/use, and is achieved by the input distribution $Q_X = (0.403, 0.418, 0, 0.179)$; (ii) the

Table 6.1: Achievable rates (bits/use) for the mismatched channel in (6.51). In addition, $C_{\text{LM}} = 1.111$ bits/use, and $C_{\text{GMI}} = 0.954$ bits/use

Input Distribution	Refined SC	Standard SC
$Q_{UX}^{(1)}$	1.313	1.060
$Q_{UX}^{(2)}$	1.236	1.236

optimized GMI is $C_{\text{GMI}} = 0.954$ bits/use, and is achieved by the input distribution $Q_X = (0.330, 0.331, 0.155, 0.184)$.

For refined superposition coding, setting $|\mathcal{U}| = 2$ and applying local optimization techniques using a number of starting points, we obtained an achievable rate of $R_{\text{RSC}}^* = 1.313$ bits/use, with $Q_U = (0.698, 0.302)$, $Q_{X|U}(\cdot|1) = (0.5, 0.5, 0, 0)$ and $Q_{X|U}(\cdot|u) = (0, 0, 0.528, 0.472)$. We denote the corresponding input distribution by $Q_{UX}^{(1)}$.

Applying similar local optimization techniques for standard superposition coding, we obtained an achievable rate of $R_{\text{SC}}^* = 1.236$ bits/use, with $Q_U = (0.830, 0.170)$, $Q_{X|U}(\cdot|1) = (0.435, 0.450, 0.115, 0)$ and $Q_{X|U}(\cdot|2) = (0, 0, 0, 1)$. We denote the corresponding input distribution by $Q_{UX}^{(2)}$.

The superposition coding rates are summarized in Table 6.1. While the achievable rate of Theorem 6.6 coincides with that of Theorem 6.3 under $Q_{UX}^{(2)}$, the former is significantly higher under $Q_{UX}^{(1)}$. Both types of superposition coding yield a strict improvement over the LM rate, which in turn provides a strict improvement over the GMI.

Since we did not perform a global optimization of $(\mathcal{U}, Q_{UX})$, we cannot necessarily include that refined superposition coding beats the standard version for the best possible choices of $\mathcal{U}$ and Q_{UX}. Nevertheless, at the very least, this example highlights that refined superposition coding can facilitate the search for a good set of parameters, and can provide significant improvements for a fixed set of parameters.

6.4.5 Comparison to Standard Superposition Coding

As mentioned previously, the refined superposition coding rate is always at least as high as that of standard superposition coding. This is formally stated in the following.

Lemma 6.7 (Comparison of Superposition Coding Rates)**.** For any mismatched DMC (W, q), any pair of parameters $(\mathcal{U}, Q_{UX})$ such that $|\mathcal{U}|$ is finite, we have

$$\max_{(R_0,R_1)\in\mathcal{R}_{\text{SC}}(Q_{UX})} R_0 + R_1 \leq \max_{(R_0,R_1,\ldots,R_{|\mathcal{U}|})\in\mathcal{R}_{\text{RSC}}(Q_{UX})} R_0 + \sum_u Q_U(u) R_{1u}, \tag{6.52}$$

where $\mathcal{R}_{\text{SC}}(Q_{UX})$ is the set of rate pairs satisfying (6.14)–(6.15), and $\mathcal{R}_{\text{RSC}}(Q_{UX})$ is the set of rate tuples satisfying (6.44)–(6.45).

Proof. We will show that for the tuple $(R_0, R_1, \ldots, R_{|\mathcal{U}|})$ achieving equality in the refined superposition coding conditions (6.44)–(6.45), we can lower bound R_0 to recover (6.15), and we can lower bound

$$R_1 \triangleq \sum_u Q_U(u) R_{1u} \tag{6.53}$$

to recover (6.14).

For the condition on R_0, we lower bound the right-hand side of (6.45) by replacing the maximum over $\mathcal{K}$ by the particular choice $\mathcal{K} = \mathcal{U}$, meaning the highest possible value R_0^* of R_0 satisfies

$$R_0^* \geq \min_{\substack{\widetilde{P}_{UXY}:\ \widetilde{P}_{UX}=P_{UX},\, \widetilde{P}_{UY}=P_{UY},\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)]\geq\mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(U;Y) + \Big[I_{\widetilde{P}}(X;Y|U) - R_1\Big]^+, \tag{6.54}$$

where we have used (6.53) and the definition of conditional mutual information. Since the condition in (6.54) trivially holds $I_{\widetilde{P}}(U;Y) > R_0$, this condition is unchanged when we constrain the minimum to satisfy $I_{\widetilde{P}}(U;Y) \leq R_0$. By doing so, and lower bounding the $[\cdot]^+$ function by its argument, we recover (6.15).

Next, defining $\widetilde{P}_{XY|U}(\cdot,\cdot|u)$ to be the minimizer in (2.15) corresponding to a given $u \in \mathcal{U}$ in (6.44), and recalling (6.53), we have

$$R_1 = \sum_u Q_U(u) I_{\widetilde{P}}(X;Y|U=u) \tag{6.55}$$

$$= I_{\widetilde{P}}(X;Y|U), \tag{6.56}$$

where in (6.56), we define $\widetilde{P}_{UXY}$ according to the above marginals along with $\widetilde{P}_U = Q_U$. The proof now simply amounts to showing that the constraints on the LM rate in (2.15) for each $u \in \mathcal{U}$ imply the superposition coding constraints in (6.14). This is a straightforward exercise; for instance, we have $\mathbb{E}_{\widetilde{P}}[\log q(X,Y)|U=u] \geq \mathbb{E}_P[\log q(X,Y)|U=u]$ for all $u \in \mathcal{U}$, and averaging over U gives the desired constraint $\mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]$. □

6.4.6 Proof Techniques

Here we describe the key idea behind the proof of Theorem 6.6; we omit the full details, which can be found in [81]. We focus on the case $\mathcal{U} = \{1,2\}$ for clarity, but the same arguments apply more generally.

Separation into Three Error Events. Recalling the codebook construction described in Section 6.4.1, we rewrite the maximum-metric decoding rule as follows:

$$(\hat{m}_0, \hat{m}_{11}, \hat{m}_{12}) = \arg\max_{(i,j_1,j_2)} q^n(\boldsymbol{x}^{(i,j_1,j_2)}, \boldsymbol{y}) \tag{6.57}$$

$$= \arg\max_{(i,j_1,j_2)} q^{n_1}(\boldsymbol{x}_1^{(i,j_1)}, \boldsymbol{y}_1(\boldsymbol{u}^{(i)})) q^{n_2}(\boldsymbol{x}_2^{(i,j_2)}, \boldsymbol{y}_2(\boldsymbol{u}^{(i)})), \tag{6.58}$$

where $\boldsymbol{y}_u(\boldsymbol{u})$ denotes the subsequence of $\boldsymbol{y}$ corresponding to the indices where $\boldsymbol{u}$ equals u. The objective in (6.58) follows by separating the indices where $u=1$ from those where $u=2$.

From (6.58), we see that for any given i, the pair (j_1, j_2) with the highest metric is the one for which j_1 maximizes $q^{n_1}(\boldsymbol{x}_1^{(i,j_1)}, \boldsymbol{y}_1(\boldsymbol{u}^{(i)}))$ and j_2 maximizes $q^{n_2}(\boldsymbol{x}_2^{(i,j_2)}, \boldsymbol{y}_2(\boldsymbol{u}^{(i)}))$. As a result, we can split the error

event into the following three types, assuming without loss of generality that the transmitted message corresponds to $(m_0, m_{11}, m_{12}) = (1, 1, 1)$:

$$\begin{aligned}
&\text{(Type 0)} \quad q^n(\boldsymbol{X}^{(i,j_1,j_2)}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y}) \text{ for some } i \neq 1,\ j_1,\ j_2;\\
&\text{(Type 1)} \quad q^{n_1}(\overline{\boldsymbol{X}}_1^{(1,j_1)}, \boldsymbol{Y}_1(\boldsymbol{U})) \geq q^{n_1}(\boldsymbol{X}_1, \boldsymbol{Y}_1(\boldsymbol{U})) \text{ for some } j_1 \neq 1;\\
&\text{(Type 2)} \quad q^{n_2}(\overline{\boldsymbol{X}}_2^{(1,j_2)}, \boldsymbol{Y}_2(\boldsymbol{U})) \geq q^{n_2}(\boldsymbol{X}_2, \boldsymbol{Y}_2(\boldsymbol{U})) \text{ for some } j_2 \neq 1.
\end{aligned}$$

By the union bound, the overall error probability is upper bounded by the sum over these three error types.

Observe that the type-1 error event corresponds to the error event for the standard constant-composition ensemble (*cf.*, Section 2.3.2) with rate R_{11}, length $n_1 = nQ_U(1)$, and input distribution $Q_{X|U}(\cdot|1)$. A similar statement holds for the type-2 error probability $\bar{p}_{\mathrm{e},2}$. As a result, we de not need to re-analyze these events; we immediately obtain the condition (6.44) from Theorem 2.3. In the following, we focus on the more difficult type-0 error event.

Separation into Joint Types. The error probability for the type-0 error event is given by

$$\bar{p}_{\mathrm{e},0} = \mathbb{P}\Big[\bigcup_{i\neq 1}\bigcup_{j_1,j_2}\{q^n(\boldsymbol{X}^{(i,j_1,j_2)}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y})\}\Big], \tag{6.59}$$

where $(\boldsymbol{Y}|\boldsymbol{X} = \boldsymbol{x}) \sim W^n(\cdot|\boldsymbol{x})$. Writing the probability as an expectation given $(\boldsymbol{U}, \boldsymbol{X}, \boldsymbol{Y})$ and applying the truncated union bound, we obtain

$$\begin{aligned}
\bar{p}_{\mathrm{e},0} \leq \mathbb{E}_{\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y}}\Bigg[\min\Bigg\{1, (M_0 - 1)\\
\times \mathbb{E}_{\overline{\boldsymbol{U}}}\Bigg[\mathbb{P}_{\{\overline{\boldsymbol{X}}^{(j_1,j_2)}\}}\Big[\bigcup_{j_1,j_2}\{q^n(\overline{\boldsymbol{X}}^{(j_1,j_2)}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y})\}\Big]\Bigg]\Bigg\}\Bigg],
\end{aligned} \tag{6.60}$$

where the codewords $\overline{\boldsymbol{X}}^{(j_1,j_2)}$ correspond to the auxiliary codeword $\overline{\boldsymbol{U}}$ (e.g., $\overline{\boldsymbol{U}} = \boldsymbol{U}^{(2)}$ and $\overline{\boldsymbol{X}}^{(j_1,j_2)} = \boldsymbol{X}^{(2,j_1,j_2)}$).

As with the other achievable rates in the monograph, one can write (6.60) in terms of joint types. The resulting counterpart to (6.36) is

given by

$$\bar{p}_{\mathrm{e},0} \le \sum_{\substack{P_{UXY}\in\mathcal{P}_n(\mathcal{U}\times\mathcal{X}\times\mathcal{Y}):\\ P_{UX}=Q_{UX,n}}} \mathbb{P}\Big[(\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y})\in\mathcal{T}^n(P_{UXY})\Big]$$
$$\times \min\Bigg\{1,(M_0-1)\sum_{\substack{\widetilde{P}_{UXY}\in\mathcal{P}_n(\mathcal{U}\times\mathcal{X}\times\mathcal{Y}):\\ P_{UX}=Q_{UX,n},\,\widetilde{P}_Y=P_Y\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)]\ge\mathbb{E}_P[\log q(X,Y)]}} \mathbb{P}\Big[(\overline{\boldsymbol{U}},\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{UY})\Big]$$
$$\times\,\mathbb{P}\Big[\bigcup_{j_1,j_2}\Big\{(\overline{\boldsymbol{u}},\overline{\boldsymbol{X}}^{(j_1,j_2)},\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{UXY})\Big\}\Big]\Bigg\}, \tag{6.61}$$

where $(\overline{\boldsymbol{u}},\boldsymbol{y})$ denotes an arbitrary pair such that $\boldsymbol{y}\in\mathcal{T}^n(P_Y)$ and $(\overline{\boldsymbol{u}},\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{UY})$.

Bounding the Double Union. We bound the doubly-indexed union in (6.61) using a similar approach to the multiple-access channel in Section 5.5. Recalling the definition of $\boldsymbol{y}_u$ following (6.58), we have

$$\begin{aligned}&(\overline{\boldsymbol{u}},\overline{\boldsymbol{X}}^{(j_1,j_2)},\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{UXY})\\ &\iff (\overline{\boldsymbol{X}}_u^{(j_u)},\boldsymbol{y}_u(\overline{\boldsymbol{u}}))\in\mathcal{T}^{n_u}(\widetilde{P}_{XY|U}(\cdot,\cdot|u)),\quad u=1,2,\end{aligned} \tag{6.62}$$

where $\{\overline{\boldsymbol{X}}_u\}_{u=1,2}$ are the partial codewords such that $\overline{\boldsymbol{X}}^{(j_1,j_2)}$ is constructed from $(\overline{\boldsymbol{u}},\overline{\boldsymbol{X}}_1,\overline{\boldsymbol{X}}_2)$. We then claim that

$$\begin{aligned}&\mathbb{P}\Big[\bigcup_{j_1,j_2}\Big\{(\overline{\boldsymbol{u}},\overline{\boldsymbol{X}}^{(j_1,j_2)},\boldsymbol{y})\in\mathcal{T}^n(\widetilde{P}_{UXY})\Big\}\Big]\\ &\quad\le\min\Big\{1,\min_{u=1,2}M_{1u}\mathbb{P}\Big[(\overline{\boldsymbol{X}}_u,\boldsymbol{y}_u(\overline{\boldsymbol{u}}))\in\mathcal{T}^{n_u}(\widetilde{P}_{XY|U}(\cdot,\cdot|u))\Big],\\ &\qquad M_{11}M_{12}\mathbb{P}\Big[\bigcap_{u=1,2}\Big\{(\overline{\boldsymbol{X}}_u,\boldsymbol{y}_u(\overline{\boldsymbol{u}}))\in\mathcal{T}^{n_u}(\widetilde{P}_{XY|U}(\cdot,\cdot|u))\Big\}\Big]\Big\}.\end{aligned} \tag{6.63}$$

The first term is trivial, the last term is the standard union bound, and the middle term corresponds to restricting the above condition for $u=1,2$ to only hold for *either* $u=1$ or $u=2$, rather than both. One can view (6.63) as taking the minimum over the four subsets $\mathcal{K}\subseteq\mathcal{U}$, including the empty set.

Wrapping Up. The remainder of the proof amounts to the same techniques as the previous sections, namely, applying standard properties of types, using the law of large numbers to establish $\|P_{UXY} - Q_{UX} \times W\|_\infty \leq \delta$, and applying a continuity argument similarly to Appendix A.1. Regarding the first of these, we note that up to sub-exponential pre-factors, $\mathbb{P}\big[(\overline{\boldsymbol{U}}, \boldsymbol{y}) \in \mathcal{T}^n(\widetilde{P}_{UY})\big]$ behaves as $e^{-nI_{\widetilde{P}}(U;Y)}$, and $\mathbb{P}\big[(\overline{\boldsymbol{X}}_u, \boldsymbol{y}_u(\overline{\boldsymbol{u}})) \in \mathcal{T}^{n_u}(\widetilde{P}_{XY|U}(\cdot,\cdot|u))\big]$ behaves as $e^{-n_u I_{\widetilde{P}}(X;Y|U=u)}$; these mutual information quantities arise accordingly in the rate condition (6.45).

Further Analysis Techniques. The ensemble tightness of the rate in Theorem 6.6 can be established using similar techniques to those of the multiple-access channel (*cf.*, Section 5.6). The details can be found in [81], along with the corresponding dual expressions and continuous-alphabet extensions.

7

Error Exponents

7.1 Introduction

Throughout the monograph, our focus has been on achievable rates via random coding. A notable limitation of studying achievable rates is that they only establish the first-order conditions for attaining asymptotically vanishing error probability, meaning that in principle the block length may need to be very large to achieve a given target.

In this section, we survey developments in random-coding error exponents, which provide refined asymptotic achievability bounds characterizing the speed of convergence of the error probability to zero at fixed rates. While error exponents are a classical topic in Shannon theory [24, 38], their study in the presence of mismatched decoding comes with interesting new challenges and insights, e.g., regarding the number of auxiliary costs (*cf.*, Section 3.5) required to maximize the exponent. While other refined asymptotics (e.g., speed of convergence to the first-order rate for a fixed error probability) have also been considered under mismatched decoding [79], we focus on error exponents, which are generally better understood.

Before proceeding, we formalize the notion of an achievable error exponent and ensemble tightness, focusing on the single-user mismatched decoding setting of Sections 2 and 3.

Definition 7.1 (Error Exponents)**.** For a mismatched memoryless channel described by (W, q), we say that $E(R)$ is an *achievable error exponent* at rate R if, for any $\delta > 0$, there exists a sequence of codebooks $\mathcal{C}_n$ with $M \geq e^{n(R-\delta)}$ codewords such that

$$\limsup_{n\to\infty} \frac{1}{n} \log p_{\mathrm{e}}(\mathcal{C}_n) \geq E(R). \tag{7.1}$$

In addition, for a given random coding ensemble with random-coding error probability $\bar{p}_{\mathrm{e}}(n, M)$, we say that $E(R)$ is the *ensemble-tight* error exponent at rate R if

$$\lim_{n\to\infty} \frac{1}{n} \log \bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor) = E(R). \tag{7.2}$$

As with the achievable rates, we will focus on i.i.d. random coding (Definition 2.1), constant-composition random coding (Definition 2.2), and various forms of cost-constrained random coding (Sections 3.4 and 3.5), and we will first consider independently-generated codewords before turning to multi-user coding techniques (Sections 5 and 6).

The results of this section are due to Csiszár and Körner [23], Kaplan and Shamai [48], Scarlett *et al.* [79, 81, 83], and Somekh-Baruch *et al.* [97].

7.2 Random-Coding Exponents for DMCs

We first consider discrete memoryless channels (DMCs), in which the input and output alphabets are finite. As is the case for the GMI and the LM rate (*cf.*, Section 2.3), the random-coding error exponents of i.i.d. random coding and constant-composition random coding can be written in both primal and dual forms. In fact, analogous primal and dual forms have long been known in the matched case (e.g., see [37]), with Csiszár and Körner's type-based analysis giving the primal form [24], and Gallager's analysis giving the dual form [38].

We proceed by defining the primal and dual forms of the relevant error exponents under mismatched decoding, and then formally state

their achievability and ensemble tightness. For i.i.d. random coding, the primal expression is given by

$$E_{\mathrm{r}}^{\mathrm{iid}}(Q_X, R) = \min_{P_{XY}} \min_{\substack{\widetilde{P}_{XY}:\ \widetilde{P}_Y = P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} D(P_{XY} \| Q_X \times W) + \left[D(\widetilde{P}_{XY} \| Q_X \times \widetilde{P}_Y) - R\right]^+, \tag{7.3}$$

and the dual expression is given by [48]

$$E_{\mathrm{r}}^{\mathrm{iid}}(Q_X, R) \triangleq \max_{\rho \in [0,1]} E_0^{\mathrm{iid}}(Q_X, \rho) - \rho R, \tag{7.4}$$

where

$$E_0^{\mathrm{iid}}(Q_X, \rho) \triangleq \sup_{s \geq 0} -\log \sum_x Q_X(x) \sum_y W(y|x) \left(\frac{\sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x}, y)^s}{q(x,y)^s} \right)^{\rho}. \tag{7.5}$$

For constant-composition random coding, the primal expression is given by [23]

$$E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R) = \min_{P_{XY}:\ P_X = Q_X} \min_{\substack{\widetilde{P}_{XY}:\ \widetilde{P}_X = Q_X, \widetilde{P}_Y = P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} \tag{7.6}$$

$$D(P_{XY} \| Q_X \times W) + \left[I_{\widetilde{P}}(X;Y) - R\right]^+, \tag{7.7}$$

and the dual expression is given by [79]

$$E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R) = \max_{\rho \in [0,1]} E_0^{\mathrm{cc}}(Q_X, \rho) - \rho R, \tag{7.8}$$

where

$$E_0^{\mathrm{cc}}(Q_X, \rho) = \sup_{s \geq 0, a(\cdot)} -\sum_x Q_X(x) \log \sum_y W(y|x) \left(\frac{\sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x}, y)^s e^{a(\overline{x})}}{q(x,y)^s e^{a(x)}} \right)^{\rho}. \tag{7.9}$$

Similarly to the primal-dual equivalences for achievable rates in Section 2.5.3, the proofs are based on Lagrange duality. The details can be found in [79].

Theorem 7.1 (Ensemble-Tight Error Exponents for DMCs)**.** For any mismatched DMC (W, q), input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, and rate $R > 0$, we have under i.i.d. random coding that

$$\lim_{n\to\infty} -\frac{1}{n} \log \bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor) = E_{\mathrm{r}}^{\mathrm{iid}}(Q_X, R), \tag{7.10}$$

and under constant-composition random coding that

$$\lim_{n\to\infty} -\frac{1}{n} \log \bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor) = E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R). \tag{7.11}$$

It is straightforward to show that the i.i.d. and constant-composition exponents are positive for $R < I_{\mathrm{GMI}}(Q_X)$ and $R < I_{\mathrm{LM}}(Q_X)$, respectively. We outline the proof of this claim for the constant-composition case [26], and note that similar arguments hold for the i.i.d. case:

- In the primal expression (7.7), the term $D(P_{XY} \| Q_X \times W)$ is always positive when P_{XY} is not equal to $Q_X \times W$. Hence, we can take $P_{XY} \to Q_X \times W$, and upon doing so, it only remains to apply the continuity of the LM rate (*cf.*, Lemma 2.5).

- In the dual expression, we can lower bound $E_0^{\mathrm{cc}}(Q_X, \rho)$ in (7.9) by the value $E_0^{\mathrm{cc}}(Q_X, \rho, s, a)$ corresponding to any fixed choices of $s \geq 0$ and $a(\cdot)$. Following the analysis of Gallager [38], one finds that the exponent is positive for all rates up to $\frac{\partial E_0^{\mathrm{cc}}(Q_X, \rho, s, a)}{\partial \rho}\big|_{\rho=0}$, and a direct calculation reveals that this derivative equals the objective function in (2.16). Optimizing $s \geq 0$ and $a(\cdot)$ recovers the LM rate.

The mismatched decoding error exponents also bear a strong resemblance to the above-mentioned error exponents for maximum-likelihood decoding, and in fact recover them as a special case [26, 79] by setting $q(x, y) = W(y|x)$. In this case, the minimizations in the primal expressions are achieved with $\widetilde{P}_{XY} = P_{XY}$, and the maximizations in the dual expressions are achieved with $s = \frac{1}{1+\rho}$.

In analogy with the fact that $I_{\mathrm{GMI}}(Q_X) \leq I_{\mathrm{LM}}(Q_X)$, we observe that

$$E_{\mathrm{r}}^{\mathrm{iid}}(Q_X, R) \leq E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R) \tag{7.12}$$

for any channel (W, q), input distribution Q_X, and rate R. This can be seen from the primal expressions by noting that the minimizations in (7.7) are more constrained than those in (7.3) (note also that $D(\widetilde{P}_{XY}\|Q_X \times P_Y)$ and $I_{\widetilde{P}}(X;Y)$ are identical under the constraints $\widetilde{P}_X = Q_X$ and $\widetilde{P}_Y = P_Y$). Alternatively, using the dual expressions, we see that setting $a(\cdot) = 0$ in (7.9) and applying Jensen's inequality recovers (7.5).

7.2.1 Proof Outline

Initial Non-Asymptotic Bounds. The starting point of the analysis is the following upper bound on the random-coding error probability when the M codewords are independently drawn from a common distribution $P_{\boldsymbol{X}}$ (see (2.81)):

$$\bar{p}_{\mathrm{e}} \leq \mathbb{E}\Big[\min\big\{1, (M-1)\mathbb{P}\big[q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y})|\boldsymbol{X}, \boldsymbol{Y}\big]\big\}\Big], \quad (7.13)$$

where $(\boldsymbol{X}, \boldsymbol{Y}, \overline{\boldsymbol{X}}) \sim P_{\boldsymbol{X}}(\boldsymbol{x})W^n(\boldsymbol{y}|\boldsymbol{x})P_{\boldsymbol{X}}(\overline{\boldsymbol{x}})$. We will also use the fact that

$$\bar{p}_{\mathrm{e}} \geq \frac{1}{2}\,\mathbb{E}\Big[\min\big\{1, (M-1)\mathbb{P}\big[q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y})|\boldsymbol{X}, \boldsymbol{Y}\big]\big\}\Big], \quad (7.14)$$

which follows by recalling the step from (2.80) to (2.81), and noting that the truncated union bound is tight to within a factor of $\frac{1}{2}$ for independent events [90, Lemma A.2].

With (7.13)–(7.14) in place, Theorem 7.1 is attained via suitable modifications of the achievable rate analysis of Section 2.6. We first outline the primal analysis, and then the dual analysis.

Achievability and Ensemble Tightness – Primal Expressions. For the primal expressions, the idea is to avoid using the law of large numbers to establish $\|P_{XY} - Q \times W\|_\infty \leq \delta$ (as was done in Section 2.6), and to instead use the following exponential characterization of the probability of a joint type P_{XY} occurring:

$$e^{-n(D(P_{XY}\|Q_X\times W)+\delta)} \leq \mathbb{P}[(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{T}^n(P_{XY})] \leq e^{-n(D(P_{XY}\|Q_X\times W)-\delta)} \quad (7.15)$$

for arbitrarily small $\delta > 0$ and sufficiently large n. This bound holds under both i.i.d. and constant-composition coding, with the latter requiring P_X to coincide with the input distribution Q_X (otherwise the probability in (7.15) is zero).

In more detail, we first write the expectation in (7.13) in terms of the joint type P_{XY} of $(\boldsymbol{X}, \boldsymbol{Y})$ as

$$\bar{p}_{\mathrm{e}} \leq \sum_{P_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})} \mathbb{P}[(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{T}^n(P_{XY})] \times \min\left\{1, (M-1)\mathbb{P}\big[q^n(\overline{\boldsymbol{X}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})\big]\right\}, \tag{7.16}$$

where in the inner probability, $(\boldsymbol{x}, \boldsymbol{y})$ is an arbitrary pair with joint type P_{XY}. Upon substituting (7.15) and the previously-obtained bounds on $\mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{y}) \geq q^n(\boldsymbol{x}, \boldsymbol{y})]$ (see (2.89) for the i.i.d. ensemble and (2.94) for the constant-composition ensemble) into (7.13), and upper bounding the outer summation by $(n+1)^{|\mathcal{X}| \cdot |\mathcal{Y}| - 1}$ times the maximum, we readily obtain the exponents (7.3) and (7.7). The ensemble tightness is proved similarly by starting from (7.14) and using the lower bound counterparts when applying each property of types, along with a continuity argument (*cf.*, Appendix A.1).

Achievability – Dual Expressions. The dual analysis is similarly straightforward given the corresponding rate derivations. Applying Markov's inequality and the inequality $\min\{1, z\} \leq z^{\rho}$ (for $\rho \in [0, 1]$) in (7.13), we obtain

$$\bar{p}_{\mathrm{e}} \leq \mathbb{E}\left[\left(M\mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y})}{q^n(\boldsymbol{X}, \boldsymbol{Y})}\right)^{s} \,\middle|\, \boldsymbol{X}, \boldsymbol{Y}\right]\right)^{\rho}\right]. \tag{7.17}$$

This step strengthens the step (2.97) that we used for the derivation of the rate, and leads to the presence of the free parameter $\rho \in [0, 1]$ in both (7.4) and (7.8). Since the subsequent analysis under i.i.d. or constant-composition random coding is similar to that of cost-constrained coding in the following subsection, and also straightforward given the dual analysis in Section 2.6, we omit the details.

7.3 Random-Coding Exponents with Continuous Alphabets

In this subsection, we consider the setup of Section 3, in which the channel alphabets may be continuous, and the transmission is subject to a system cost $\frac{1}{n}\sum_{i=1}^{n} c(x_i) \leq \Gamma$.

In this setting, we introduced the cost-constrained random coding ensemble in Section 3.4 with a single auxiliary cost, which is used to recover the optimization variable $a(\cdot)$ in the dual form of the LM rate. In Section 3.5, we generalized this ensemble to contain multiple auxiliary costs $a_1(\cdot), \ldots, a_L(\cdot)$, with the motivation being that one may not always have knowledge of the optimal (single) choice of $a(x)$. We also briefly mentioned in Section 5.7 that the use of multiple auxiliary costs is beneficial in the multiple-access setting. In this subsection and the subsequent subsections, we will see that yet another benefit of using multiple auxiliary costs is in attaining improved error exponents [79].

For convenience, we repeat the definition of the codeword distribution: With input distribution Q_X, system cost $c(\cdot)$, auxiliary costs $\{a_l\}_{l=1}^{L}$, and parameter $\delta > 0$, we have

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\Omega_n} \prod_{i=1}^{n} Q_X(x_i) \mathbb{1}\{\boldsymbol{x} \in \mathcal{D}_n\}, \tag{7.18}$$

where

$$\mathcal{D}_n \triangleq \Bigg\{\boldsymbol{x}\colon \left|\frac{1}{n}\sum_{i=1}^{n} c(x_i) - \phi_c\right| \leq \delta, \\ \left|\frac{1}{n}\sum_{i=1}^{n} a_l(x_i) - \phi_l\right| \leq \delta,\ \forall l = 1, \ldots, L\Bigg\}, \tag{7.19}$$

and where $\phi_c = \mathbb{E}_Q[c(X)]$, $\phi_a = \mathbb{E}_Q[a(X)]$, and Ω_n is a normalizing constant.

Two Auxiliary Costs. Here we show that two suitably optimized auxiliary costs suffice to obtain an error exponent matching that of the constant-composition ensemble. Specifically, we generalize the achievability of $E_{\mathrm{r}}^{\mathrm{cc}}$ (*cf.*, Theorem 7.1) to continuous alphabets. The analogous generalization of the achievability of $E_{\mathrm{r}}^{\mathrm{iid}}$ is similar, but is omitted here.

The generalized exponent is given by

$$E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R) = \max_{\rho \in [0,1]} E_0^{\mathrm{cc}}(Q_X, \rho) - \rho R, \tag{7.20}$$

where

$$E_0^{\mathrm{cc}}(Q_X, \rho) = \sup_{s \geq 0, a(\cdot)} \mathbb{E}\left[-\log \mathbb{E}\left[\left(\frac{\mathbb{E}\left[q(\overline{X}, Y)^s e^{a(\overline{X})} | Y \right]}{q(X,Y)^s e^{a(X)}} \right)^{\rho} \middle| X \right] \right] \tag{7.21}$$

and where $(X, Y, \overline{X}) \sim Q_X(x) W(y|x) Q_X(\overline{x})$, and the supremum is over all $a(\cdot)$ such that $\mathbb{E}_Q[a(X)] < \infty$.

While the expression (7.21) only contains a single auxiliary function $a(\cdot)$ (corresponding to one of the auxiliary costs), a second auxiliary cost is used to ensure that the average over X in (7.21) is *outside* the logarithm, which is preferable due to Jensen's inequality. To derive an exponent with the expectation inside the logarithm, a single auxiliary cost suffices [86]; such an exponent is positive for all $R < I_{\mathrm{LM}}(Q_X)$, but can be smaller than $E_{\mathrm{r}}^{\mathrm{cc}}$ in general [79].

Theorem 7.2 (Error Exponent for Cost-Constrained Coding)**.** For any mismatched memoryless channel (W, q), we have the following under cost-constrained random coding with $L = 2$ and suitably-chosen $a_1(\cdot)$, $a_2(\cdot)$, and $\delta > 0$:

$$\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor) \leq e^{-n(E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R) - \delta')} \tag{7.22}$$

for an arbitrarily small constant $\delta' > 0$ and sufficiently large n. Hence, under an input constraint (c, Γ), if $\mathbb{E}_Q[c(X)] < \Gamma$, then $E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R)$ is an achievable error exponent at rate R.

We provide an outline of the proof (see [79] for the details), highlighting the fact that the two auxiliary costs $a_1(\cdot)$ and $a_2(\cdot)$ play different roles. As in Section 7.2, the analysis is based on the bound

$$\bar{p}_{\mathrm{e}} \leq \mathbb{E}\left[\left(M \mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y})}{q^n(\boldsymbol{X}, \boldsymbol{Y})} \right)^s \middle| \boldsymbol{X}, \boldsymbol{Y} \right] \right)^{\rho} \right]. \tag{7.23}$$

We fix $a_1(\cdot)$ to be specified later, and analyze the inner expectation in the same way as Section 3.6. After doing so, we find that the exponential

decay of the error probability given $\boldsymbol{X} = \boldsymbol{x}$ is dictated by the following quantity:

$$-\frac{1}{n}\sum_{i=1}^{n}\log\mathbb{E}\left[\left(\frac{\mathbb{E}\big[q(\overline{X},y)^s e^{a_1(\overline{X})}|Y\big]}{q(X,Y)^s e^{a_1(X)}}\right)^{\rho}\,\middle|\,X=x_i\right]. \tag{7.24}$$

In order to replace this empirical average by an average over Q_X, we choose the second auxiliary cost as

$$a_2(x) = -\log\mathbb{E}\left[\left(\frac{\mathbb{E}\big[q(\overline{X},Y)^s e^{a_1(\overline{X})}|Y\big]}{q(X,Y)^s e^{a_1(X)}}\right)^{\rho}\,\middle|\,X=x\right]. \tag{7.25}$$

Since $\left|\frac{1}{n}\sum_{i=1}^{n} a_2(x_i) - \mathbb{E}_Q[a_2(X)]\right| \le \delta$ by construction in (7.19), this means that (7.24) can be lower bounded by the argument to the supremum in (7.21), up to an arbitrarily small loss of δ. The proof is completed by noting that δ can be arbitrarily small, optimizing the free parameters $s \ge 0$ and $a_1(\cdot)$, and renaming $a_1(\cdot)$ to $a(\cdot)$.

Additional Auxiliary Costs. In analogy with Theorem 3.4, one can use similar analysis techniques to deduce an achievable error exponent for a fixed and possibly suboptimal set of auxiliary costs. We state the resulting error exponent without proof, and refer the reader to [79] for the details:

$$E_{\mathrm{r}}^{\mathrm{cost}}(Q_X, R, \{a_l\}) = \max_{\rho\in[0,1]} E_0^{\mathrm{cost}}(Q_X,\rho,\{a_l\}) - \rho R, \tag{7.26}$$

where

$$E_0^{\mathrm{cost}}(Q_X,\rho,\{a_l\}) \triangleq \sup_{s\ge 0,\{r_l\},\{\bar{r}_l\}} - \log\mathbb{E}\left[\left(\frac{\mathbb{E}\big[q(\overline{X},Y)^s e^{\sum_{l=1}^{L}\bar{r}_l(a_l(\overline{X})-\phi_l)}|Y\big]}{q(X,Y)^s e^{\sum_{l=1}^{L} r_l(a_l(X)-\phi_l)}}\right)^{\rho}\right], \tag{7.27}$$

and where $\{a_l\}_{l=1}^{L}$ are the auxiliary costs, and $\phi_\ell = \mathbb{E}_Q[a_l(X)]$.

Similarly to the achievable rates, the use of $L > 2$ auxiliary costs may be useful when one does not have knowledge of the optimal $a_1(\cdot)$ and $a_2(\cdot)$ in the proof of Theorem 7.2. It is tempting to ask whether using $L > 2$ *optimized* auxiliary costs can lead to a better exponent

than that of Theorem 7.2; however, the answer turns out to be negative, since

$$E_{\mathrm{r}}^{\mathrm{iid}}(Q_X, R) \le E_{\mathrm{r}}^{\mathrm{cost}}(Q_X, R, \{a_l\}) \le E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R). \tag{7.28}$$

The left inequality follows by setting each $r_l = \bar{r}_l = 0$ in (7.27), and the right inequality follows from two steps: (i) use Jensen's inequality to take the outer expectation over X outside the logarithm in (7.27); (ii) define $a(x) = \sum_{l=1}^{L} \bar{r}_l(a_l(\overline{X}) - \phi_l)$ to recover the objective function in (7.21).

7.4 Expurgated Exponents

Expurgation is a classical technique for attaining improved error exponents at low rates [36]. The technique consists of removing, from a randomly generated codebook, a given fraction of codewords whose conditional error probability is the highest. In this subsection, we overview an extension of Gallager's expurgation technique [38, Sec. 5.7] to the mismatched setting [83]. In Section 7.5, we will discuss alternative techniques that also provide these low-rate improvements.

The derivation of primal and dual expurgated exponents follows similar steps to those of the random coding exponent once the appropriate initial non-asymptotic bound (i.e., a counterpart to (7.13)) is obtained. Such a bound is given in the following theorem.

Theorem 7.3 (Non-Asymptotic Expurgated Bound)**.** For any mismatched memoryless channel (W, q) and codeword distribution $P_{\boldsymbol{X}}$, there exists a codebook $\mathcal{C}_n$ with M codewords of length n whose error probability satisfies

$$p_{\mathrm{e}}(\mathcal{C}_n) \le \inf_{\rho \ge 1} \Big(4(M-1)\mathbb{E}\Big[\mathbb{P}\big[q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \ge q^n(\boldsymbol{X}, \boldsymbol{Y}) \,\big|\, \boldsymbol{X}, \overline{\boldsymbol{X}}\big]^{1/\rho}\Big]\Big)^{\rho}, \tag{7.29}$$

where $(\boldsymbol{X}, \boldsymbol{Y}, \overline{\boldsymbol{X}}) \sim P_{\boldsymbol{X}}(\boldsymbol{x}) W^n(\boldsymbol{y}|\boldsymbol{x}) P_{\boldsymbol{X}}(\overline{\boldsymbol{x}})$.

The proof closely follows the analysis of Gallager [38, Sec. 5.7], and is outlined as follows:

- Let C_n be a codebook with $M' = 2M - 1$ codewords drawn independently from $P_{\boldsymbol{X}}$, and let $p_{\mathrm{e},m}(\mathsf{C}_n)$ be the conditional error

probability associated with the m-th message. We have from Markov's inequality that at least half of the codewords satisfy $p_{\mathrm{e},m}(\mathsf{C}_n)^{1/\rho} \leq 2\mathbb{E}[p_{\mathrm{e},m}(\mathsf{C}_n)^{1/\rho}]$; hence, there exists a codebook $\mathcal{C}_n$ with M codewords such that

$$p_e(\mathcal{C}_n) \leq \left(2\mathbb{E}[p_{\mathrm{e},m}(\mathsf{C}_n)^{1/\rho}]\right)^{\rho}. \tag{7.30}$$

- We obtain (7.29) by using the union bound to upper bound $\mathbb{E}[p_{\mathrm{e},m}(\mathsf{C}_n)^{1/\rho}]$ in terms of $M' - 1 = 2(M-1)$ error events $\{q^n(\boldsymbol{X}^{(\overline{m})}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}^{(m)}, \boldsymbol{Y})\}_{\overline{m} \neq m}$, and then applying the inequality $\left(\sum_i \alpha_i\right)^{1/\rho} \leq \sum_i \alpha_i^{1/\rho}$ for $\rho \geq 1$.

With Theorem 7.3 in place, we can consider the same choices of $P_{\boldsymbol{X}}$ as those of Sections 7.2 and 7.3: i.i.d., constant-composition, and cost-constrained. We focus on the latter two, as they attain higher exponents. Starting with the discrete memoryless setting, the primal form of the constant-composition exponent is given by

$$E_{\mathrm{ex}}^{\mathrm{cc}}(Q_X, R) \triangleq \min_{\substack{P_{X\widetilde{X}Y}:\ P_X = P_{\widetilde{X}} = Q_X, \\ \mathbb{E}_P[\log q(\widetilde{X},Y)] \geq \mathbb{E}_P[\log q(X,Y)] \\ I_P(X;\widetilde{X}) \leq R}} D(P_{X\widetilde{X}Y} \| Q_X \times Q_X \times W) - R, \tag{7.31}$$

and the dual expression is given by

$$E_{\mathrm{ex}}^{\mathrm{cc}}(Q_X, R) = \sup_{\rho \geq 1} E_{\mathrm{x}}^{\mathrm{cc}}(Q_X, \rho) - \rho R, \tag{7.32}$$

where

$$E_{\mathrm{x}}^{\mathrm{cc}}(Q_X, \rho) = \sup_{s \geq 0, a(\cdot)} -\rho \sum_x Q_X(x) \times \log \sum_{\overline{x}} Q_X(\overline{x}) \left(\sum_y W(y|x) \left(\frac{q(\overline{x}, y)}{q(x, y)} \right)^s \frac{e^{a(\overline{x})}}{e^{a(x)}} \right)^{1/\rho}. \tag{7.33}$$

Once again, the equivalence between the two is proved via Lagrange duality [83].

Theorem 7.4 (Expurgated Error Exponent for DMCs)**.** For any mismatched DMC (W, q), input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, and rate $R > 0$,

there exists a sequence of codebooks $\mathcal{C}_n$ with $M = \lfloor e^{nR} \rfloor$ messages of length n such that

$$\lim_{n\to\infty} -\frac{1}{n} \log p_{\mathrm{e}}(\mathcal{C}_n) \geq E_{\mathrm{ex}}^{\mathrm{cc}}(Q_X, R). \tag{7.34}$$

The primal expression is derived from (7.29) using the method of types. The dual expression is derived by fixing $s \geq 0$ and weakening (7.29) as follows via Markov's inequality:

$$p_{\mathrm{e}}(\mathcal{C}_n) \leq \left(4(M-1)\mathbb{E}\left[\mathbb{E}\left[\left(\frac{q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y})}{q^n(\boldsymbol{X}, \boldsymbol{Y})} \right)^s \,\middle|\, \boldsymbol{X}, \overline{\boldsymbol{X}} \right]^{1/\rho} \right] \right)^{\rho}. \tag{7.35}$$

With this bound in place, the desired error exponent is obtained via further bounding techniques similar to the dual derivation of the random-coding exponent.

The extension of Theorem 7.4 to continuous alphabets follows similar steps to Section 7.3 via cost-constrained random coding with $L = 2$: We let one auxiliary cost $a_1(\cdot)$ correspond to $a(\cdot)$ in (7.33), and we choose the other auxiliary cost as

$$a_2(x) = -\log \mathbb{E}\left[\left(\mathbb{E}\left[\left(\frac{q(\overline{X}, Y)}{q(x, Y)} \right)^s \frac{e^{a_1(\overline{X})}}{e^{a_1(x)}} \,\middle|\, \overline{X}, X = x \right] \right)^{1/\rho} \right], \tag{7.36}$$

in analogy with (7.25). The details can be found in [83].

7.5 Attaining the Best of Both Exponents

In addition to the constant-composition expurgated exponent $E_{\mathrm{ex}}^{\mathrm{cc}}$ in (7.31), recall from (7.7) that the primal form of the constant-composition random-coding exponent $E_{\mathrm{r}}^{\mathrm{cc}}$ is given by

$$E_{\mathrm{r}}^{\mathrm{cc}}(Q_X, R) = \min_{P_{XY}\colon\, P_X = Q_X} \; \min_{\substack{\widetilde{P}_{XY}\colon\, \widetilde{P}_X = Q_X, \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} D(P_{XY} \| Q_X \times W) + \left[I_{\widetilde{P}}(X;Y) - R \right]^+. \tag{7.37}$$

With these two achievable error exponents established, we highlight the following limitations of Theorem 7.4:

- Despite the low-rate improvement, the exponent $E_{\mathrm{ex}}^{\mathrm{cc}}$ is loose at high rates, in particular falling to zero at a rate strictly below $I_{\mathrm{LM}}(Q_X)$, in contrast with $E_{\mathrm{r}}^{\mathrm{cc}}$ (see Section 7.6 for an example). While one could overcome this by simply using the fact that $\max\{E_{\mathrm{r}}^{\mathrm{cc}}, E_{\mathrm{ex}}^{\mathrm{cc}}\}$ is clearly an achievable exponent, it is also of interest to find a construction that directly attains the best of both exponents simultaneously.

- While a random codebook C_n is considered as a step towards constructing the codebook $\mathcal{C}_n$, the structure of $\mathcal{C}_n$ is less clear (i.e., it is difficult to characterize properties of the removed vs. retained codewords). Hence, the standard expurgation argument is, in a sense, less constructive than standard random-coding methods.

- In contrast with the second part of Theorem 7.1 concerning $E_{\mathrm{r}}^{\mathrm{cc}}$, Theorem 7.4 does not make any claim of ensemble tightness, so it is unclear to what extent the exponent could be improved via refined analysis techniques.

The first of these limitations was addressed in an early work of Csiszár and Körner [23], who derived the following error exponent:

$$E_{\mathrm{CK}}(Q_X, R) = \min_{\substack{P_{X\widetilde{X}Y}:\ P_X = P_{\widetilde{X}} = Q_X,\\ \mathbb{E}_P[\log q(\widetilde{X},Y)] \geq \mathbb{E}_P[\log q(X,Y)],\\ I_P(X;\widetilde{X}) \leq R}} D(P_{Y|X} \| W | Q_X) + [I(\widetilde{X}; Y, X) - R]_+. \tag{7.38}$$

This can be weakened to $E_{\mathrm{ex}}^{\mathrm{cc}}$ by lower bounding $[\cdot]_+$ by its argument, or can be weakened to $E_{\mathrm{r}}^{\mathrm{cc}}$ by dropping the constraint $I_P(X;\widetilde{X}) \leq R$ and lower bounding $I_P(\widetilde{X}; Y, X) \geq I_P(\widetilde{X}; Y)$. In the latter case, the resulting optimization problem in (7.38) depends on $P_{X\widetilde{X}Y}$ only through $(P_{XY}, \widetilde{P}_{XY})$, and the minimization can be split into two minimizations of the form (7.37) accordingly.

The derivation of E_{CK} in [23] uses combinatorial arguments to establish the existence of a constant-composition code with certain properties, but the proof is highly non-constructive in the sense discussed above. In addition, the analysis in [23] appears to be unsuitable for channels with continuous alphabets, and leaves open the question of ensemble tightness. In the following, we outline a recent alternative derivation of (7.38) that addresses both of these limitations [97].

The analysis in [97] is based on a sequential random coding construction resembling the Gilbert-Varshamov construction for binary codes [41, 103]. Fixing an input distribution Q_X, a generic "distance" function $d(\boldsymbol{x}, \boldsymbol{x}')$, and a constant Δ, the codebook is generated as follows:

1. The first codeword, $\boldsymbol{X}_1$, is drawn uniformly from the type class $\mathcal{T}^n(Q_{X,n})$, where $Q_{X,n}$ is a type approximating Q_X;

2. Given $\boldsymbol{X}_1 = \boldsymbol{x}_1$, the second codeword $\boldsymbol{X}_2$ is drawn uniformly from the set

$$\mathcal{T}^n(Q_{X,n}, \boldsymbol{x}_1) \triangleq \{\overline{\boldsymbol{x}} \in \mathcal{T}^n(Q_{X,n})\colon d(\overline{\boldsymbol{x}}, \boldsymbol{x}_1) > \Delta\} \tag{7.39}$$

$$= \mathcal{T}^n(Q_{X,n}) \backslash \{\overline{\boldsymbol{x}} \in \mathcal{T}^n(Q_{X,n})\colon d(\overline{\boldsymbol{x}}, \boldsymbol{x}_1) \leq \Delta\}\,, \tag{7.40}$$

   i.e., the set of sequences with composition $Q_{X,n}$ whose distance to $\boldsymbol{x}_1$ exceeds Δ;

3. Continuing recursively, if the first $i-1$ codewords are $\boldsymbol{x}_1^{i-1} \triangleq (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1})$, then the i-th codeword $\boldsymbol{X}_i$ is drawn uniformly from the set

$$\mathcal{T}^n(Q_{X,n}, \boldsymbol{x}_1^{i-1})$$

$$\triangleq \{\overline{\boldsymbol{x}} \in \mathcal{T}^n(Q_{X,n})\colon d(\overline{\boldsymbol{x}}, \boldsymbol{x}_j) > \Delta,\ \forall j = 1, \ldots, i-1\} \tag{7.41}$$

$$= \mathcal{T}^n(Q_{X,n}, \boldsymbol{x}_1^{i-2}) \backslash \left\{\overline{\boldsymbol{x}} \in \mathcal{T}^n(Q_{X,n}, \boldsymbol{x}_1^{i-2})\colon d(\overline{\boldsymbol{x}}, \boldsymbol{x}_{i-1}) \leq \Delta\right\}. \tag{7.42}$$

Essentially, we sequentially perform random coding while avoiding any codewords that are too close to those already selected, according to a generic distance $d(\cdot, \cdot)$. This procedure was termed the *(generalized) random Gilbert-Varshamov* (RGV) construction in [97].

While we use the terminology "distance", $d(\cdot,\cdot)$ need not be a distance function in the topological sense. However, for the analysis, it is required to be symmetric, bounded, and dependent only on the joint empirical distribution of its arguments. To highlight the latter condition, we henceforth write $d(P_{X\widetilde{X}})$ as a function on the probability simplex. For instance, the choice $d(P_{X\widetilde{X}}) = \mathbb{P}[X \neq \widetilde{X}]$ corresponds to a normalized Hamming distance, and the negative mutual information $d(P_{X\widetilde{X}}) = -I_P(X;\widetilde{X})$ will be highlighted below.

If the recursive procedure is continued for too many iterations, the entire space of codewords may be exhausted. To ensure that this does not occur, it suffices to enforce that M times the volume of the appropriate d-ball is smaller than the size of the type class $|\mathcal{T}(Q_{X,n})|$. Asymptotically, this translates into the following rate condition [97]:

$$R \leq \min_{\substack{P_{X\widetilde{X}}:\ d(P_{X\widetilde{X}})\leq\Delta,\\ P_X=P_{\widetilde{X}}=Q_X}} I_P(X;\widetilde{X}) - 2\delta \tag{7.43}$$

for arbitrarily small $\delta > 0$. Under this condition, it was shown in [97] that the above sequential random coding scheme attains the following strengthened (albeit asymptotic) variant of (7.13):

$$\begin{aligned}\bar{p}_{\mathrm{e}} \doteq & \sum_{\boldsymbol{x}\in\mathcal{T}^n(Q_{X,n}),\boldsymbol{y}} \frac{1}{|\mathcal{T}^n(Q_{X,n})|} W^n(\boldsymbol{y}|\boldsymbol{x}) \\ & \times \min\Bigg\{1, (M-1) \sum_{\substack{\boldsymbol{x}'\in\mathcal{T}^n(Q_{X,n}):\ q^n(\boldsymbol{x}',\boldsymbol{y})\geq q^n(\boldsymbol{x},\boldsymbol{y})\\ d(\boldsymbol{x},\boldsymbol{x}')\geq\Delta}} \frac{1}{|\mathcal{T}^n(Q_{X,n})|}\Bigg\},\end{aligned} \tag{7.44}$$

where $\doteq$ denotes asymptotic equality up to a sub-exponential prefactor. Note that the key difference compared to (7.13) is the constraint $d(\boldsymbol{x},\boldsymbol{x}') \geq \Delta$. With this expression in place, an analysis based on the method of types yields the following theorem [97].

Theorem 7.5 (Ensemble-Tight Exponent for the RGV Construction)**.** For any mismatched DMC (W,q), input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, bounded symmetric type-dependent distance function $d(P_{X\widetilde{X}})$, parameters Δ and $\delta > 0$, and rate R satisfying (7.43), we have under the RGV construction

that

$$\liminf_{n\to\infty} -\frac{1}{n}\log \overline{p}_{\mathrm{e}}(n, \lfloor e^{nR}\rfloor) \geq E_{\mathrm{RGV}}(Q_X, R, \Delta), \tag{7.45}$$

$$\limsup_{n\to\infty} -\frac{1}{n}\log \overline{p}_{\mathrm{e}}(n, \lfloor e^{nR}\rfloor) \leq E_{\mathrm{RGV}}(Q_X, R, \Delta+\epsilon), \tag{7.46}$$

where $\epsilon > 0$ is arbitrarily small, and

$$\begin{aligned} &E_{\mathrm{RGV}}(Q_X, R, \Delta) \\ &= \min_{\substack{P_{X\widetilde{X}Y}:\ P_X = P_{\widetilde{X}} = Q_X,\\ \mathbb{E}_P[\log q(\widetilde{X},Y)] \geq \mathbb{E}_P[\log q(X,Y)],\\ d(P_{X\widetilde{X}}) \geq \Delta}} D(P_{Y|X}\|W|Q_X) + [I_P(\widetilde{X};Y,X) - R]_+. \end{aligned} \tag{7.47}$$

In addition, (7.45) holds even in the case that $d(\cdot)$ is non-symmetric.

Observe that (7.45) and (7.46) establish ensemble tightness under the technical condition that E_{RGV} is continuous in Δ at the specified input. The final claim of the theorem concerning non-symmetric $d(\boldsymbol{x}, \boldsymbol{x}')$ is proved by considering a symmetrized distance of the form $\min\{d(\boldsymbol{x}, \boldsymbol{x}'), d(\boldsymbol{x}', \boldsymbol{x})\}$ and applying the first part of the theorem [97].

Setting $d(P_{X\widetilde{X}}) = -I_P(X;\widetilde{X})$ in (7.47) readily recovers E_{CK} in (7.38), but the more general form here also provides insight into other distance functions (though it can be shown that no choice provides a better exponent than E_{CK}), e.g., see (7.51) below.

In the case of additive distance functions, i.e., $d(\boldsymbol{x}, \boldsymbol{x}') = \frac{1}{n}\sum_{i=1}^n d(x_i, x_i')$,[1] one can use Lagrange duality to establish an equivalent dual form of (7.47), given by [97]

$$E_{\mathrm{RGV}}(Q_X, R, \Delta) = \sup_{\rho\in[0,1]} E_{0,\mathrm{RGV}}(Q_X, \rho, \Delta) - \rho R, \tag{7.48}$$

where

$$\begin{aligned} E_{0,\mathrm{RGV}}(Q_X, \rho, \Delta) = &\sup_{r\geq 0, s\geq 0, a(\cdot)} -\sum_x Q_X(x) \\ &\times \log \sum_y W(y|x)\left(\frac{\sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x}, y)^s e^{a(\overline{x})} e^{r(d(x,\overline{x})-\Delta)}}{q(x,y)^s e^{a(x)}}\right)^{\rho}. \end{aligned} \tag{7.49}$$

[1] Here we normalize by $\frac{1}{n}$ in contrast with the previous sections; this is for consistency with the assumption that the distance can be expressed in terms of the joint type: $d(P_{X\widetilde{X}}) = \mathbb{E}_P[d(X,\widetilde{X})]$ is equivalent to $d(\boldsymbol{x}, \boldsymbol{x}') = \frac{1}{n}\sum_{i=1}^n d(x_i, x_i')$.

Similarly, one can obtain the following dual expression for the rate condition (7.43):

$$R \leq \sup_{r\geq 0, a(\cdot)} \sum_x Q_X(x) \log \sum_{\overline{x}} Q_X(\overline{x}) e^{a(\overline{x})-\phi_a} e^{-r(d(x,\overline{x})-\Delta)} - 2\delta, \quad (7.50)$$

where $\phi_a = \mathbb{E}_Q[a(X)]$. Setting $r = 0$ in (7.49) readily recovers $E_{\mathrm{r}}^{\mathrm{cc}}$ (*cf.*, (7.37)). With some additional effort, we can also show that (7.48) recovers $E_{\mathrm{ex}}^{\mathrm{cc}}$ (*cf.*, (7.32)). The idea is to use the final claim of Theorem 7.5 concerning non-symmetric distance functions, setting $\rho = 1$ in (7.49) and choosing

$$d(x, \overline{x}) = -\log \sum_y W(y|x) \Big(\frac{q(\overline{x}, y)}{q(x, y)} \Big)^s \quad (7.51)$$

for some $s \geq 0$. With a suitable choice of Δ, it can be shown that (7.49) recovers $E_{\mathrm{x}}^{\mathrm{cc}}$ in (7.33), while also maintaining the rate condition (7.50) [97].

Finally, we briefly discuss the generalization of (7.48)–(7.50) to continuous alphabets, which can be achieved (for additive distance functions) by suitably modifying the recursive random coding construction. The idea is to recursively draw from a cost-constrained codeword distribution *conditioned* on the distance to each previously-selected codeword being at least Δ, and to derive an upper bound analogous to (7.44), with the cost-constrained codeword distribution in place of $\frac{1}{|\mathcal{T}^n(Q_{X,n})|}$. The interested reader is referred to [97] for the details.

7.6 Numerical Example

Here we present a numerical example from [79] that serves to compare the exponents attained by i.i.d. vs. constant-composition coding, as well as random coding vs. expurgated exponents. We consider an asymmetric channel, as we found that this better highlights these differences. Specifically, the channel and metric are defined by the entries of the $|\mathcal{X}| \times |\mathcal{Y}|$ matrices

$$\boldsymbol{W} = \begin{bmatrix} 1-2\delta_0 & \delta_0 & \delta_0 \\ \delta_1 & 1-2\delta_1 & \delta_1 \\ \delta_2 & \delta_2 & 1-2\delta_2 \end{bmatrix}, \quad \boldsymbol{q} = \begin{bmatrix} 1-2\delta & \delta & \delta \\ \delta & 1-2\delta & \delta \\ \delta & \delta & 1-2\delta \end{bmatrix} \quad (7.52)$$

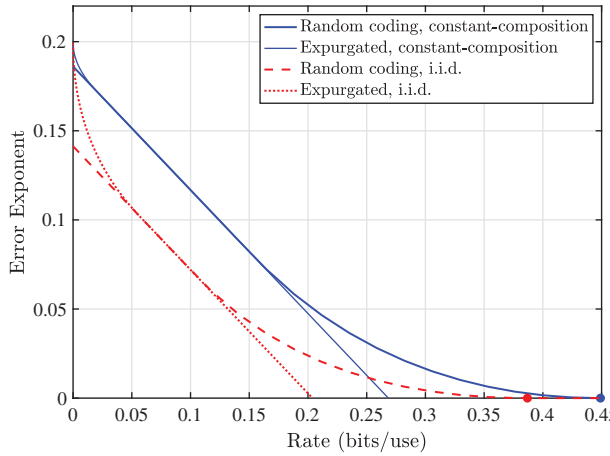


Figure 7.1: Mismatched random-coding and expurgated exponents under constant-composition and i.i.d. random coding. The GMI and LM rate are marked on the horizontal axis.

with $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$, and with δ_0, δ_1, δ_2, and δ taking values in $(0, \frac{1}{3})$. We set $\delta_0 = 0.01$, $\delta_1 = 0.05$, $\delta_2 = 0.25$, and $Q_X = (0.1, 0.3, 0.6)$, and note that any choice of $\delta \in (0, \frac{1}{3})$ yields an equivalent decoder (namely, minimum-distance decoding). Under these parameters, we have $I_{\text{GMI}}(Q_X) = 0.387$, $I_{\text{LM}}(Q_X) = 0.449$, and $I(X;Y) = 0.471$ bits/use. The resulting error exponents are shown in Figure 7.1.

As expected, constant-composition random coding consistently yields a higher exponent than i.i.d. random coding. For both ensembles, the random-coding exponent is positive for all rates below the corresponding achievable rate, whereas this fails to hold for the expurgated exponents. On the other hand, the expurgated exponents are better at low rates.

In the limit of zero rate, the i.i.d. and constant-composition expurgated exponents approach the same value, i.e., $E_{\text{ex}}^{\text{cc}}(0) = E_{\text{ex}}^{\text{iid}}(0)$. In fact, this behavior is not specific to this example, but rather a general phenomenon for any channel, decoding metric, and input distribution [83].

7.7 Error Exponents for Multi-User Coding Techniques

The analysis leading to the achievable rates for the multiple-access channel (Section 5) and superposition coding (Section 6) can also be

adapted to obtain achievable error exponents. The modifications are similar to those above for the single-user channel, but somewhat more technical [77].

Starting with the multiple-access channel, the main modification required is to use the exponential upper bound (5.32) directly (as well as a matching lower bound for the ensemble tightness part), rather than only using it to infer that $\|P_{X_1X_2Y} - Q_1 \times Q_2 \times W\|_\infty \leq \delta$. By doing so, we find that the three rate conditions in (5.7)–(5.9) each have a corresponding error exponent containing the divergence term $D(P_{X_1X_2Y}\|Q_1 \times Q_2 \times W)$ from (5.32), and the overall error exponent is the minimum of the three. For instance, the error exponent associated with (5.9) is given by

$$\begin{aligned} &E_{\mathrm{r},12}^{\mathrm{cc}}(\boldsymbol{Q}, R_1, R_2) \\ &= \min_{P_{X_1X_2Y}:\ P_{X_1}=Q_1, P_{X_2}=Q_2} \ \min_{\substack{\widetilde{P}_{X_1X_2Y}:\ \widetilde{P}_{X_1}=P_{X_1}, \widetilde{P}_{X_2}=P_{X_2}, \widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} \\ &\quad D(P_{X_1X_2Y}\|Q_1 \times Q_2 \times W) + \Big[\max\Big\{I_{\widetilde{P}}(X_1;Y) - R_1, \\ &\quad I_{\widetilde{P}}(X_2;Y) - R_2, D(\widetilde{P}_{X_1X_2Y}\|Q_1 \times Q_2 \times P_Y) - R_1 - R_2\Big\}\Big]^+. \end{aligned} \tag{7.53}$$

Dual expressions for the error exponents, along with direct derivations, can be obtained using analogous (albeit more cumbersome) techniques to those for the achievable rate region. As was the case in the single-user setting, the number of auxiliary costs used is higher when it comes to attaining the error exponents; it was shown in [81] that five per user is sufficient.

Similar ideas apply in the context of superposition coding (*cf.*, Section 6) [77]. While we saw that refined superposition coding (*cf.*, Section 6.4) provides an achievable rate at least as high as all other known achievable rates, it is worth noting that the error exponent may be small due to the use of shorter "sub-codes" of length $n_u = nQ_U(u)$ for each auxiliary symbol $u \in \mathcal{U}$. If such a sub-code has error probability decaying as $e^{-n_u E_u}$ for some $E_u > 0$, then its error exponent with respect to the full block length n is $Q_U(u)E_u$, which may be low if $Q_U(u)$ is

small. In contrast, for standard superposition coding (*cf.*, Section 6.3), all of the error events directly correspond to a block length of n.

To demonstrate this difference, and also compare against the exponents of standard random coding (i.e., independent codewords), we revisit the example from Section 6.4.4, in which the refined superposition coding rate was strictly higher than the standard version. We compare the exponents as follows:

- We let Q^*_{UX} be the optimized input distribution attaining the highest refined superposition coding rate, and let $(R^*_0, R^*_{11}, R^*_{12})$ be the corresponding triplet attaining equality in each rate condition of Theorem 6.6. For a range of $\alpha \in [0, 1]$, we compute the error exponent for the rate triplets $(R_0, R_{11}, R_{12}) = (\alpha R^*_0, \alpha R^*_{11}, \alpha R^*_{12})$, with rate $R = R_0 + Q^*_U(1)R_{11} + Q^*_U(2)R_{12}$.

- Similarly, for the same Q^*_{UX}, letting (R^*_0, R^*_1) be the pair attaining equality in the rate conditions for standard superposition coding (Theorem 6.3), we plot the exponent for various pairs $(R_0, R_1) = (\alpha R^*_0, \alpha R^*_1)$, with rate $R = R_0 + R_1$.

The resulting error exponents are shown in Figure 7.2, along with the error exponents for i.i.d. and constant-composition random coding with independent codewords and an input distribution chosen to maximize the respective rate (see Section 6.4.4).

Although refined superposition coding has the highest achievable rate, we see that the corresponding error exponent is much smaller than the standard version at lower rates, likely as a result of using shorter sub-codes. Interestingly, a similar phenomenon also occurs when moving from standard superposition coding to independent codewords: The superposition coding rate is comparable to the LM rate and higher than the GMI, yet its exponent is smaller at lower rates. A possible reason for this is that superposition coding yields the minimum of two error exponents – one for each error type – whereas with independent codewords we only have a single error type.

This example indicates a trend of ensembles with higher rates tending to yield lower exponents at smaller rates (with the exception of

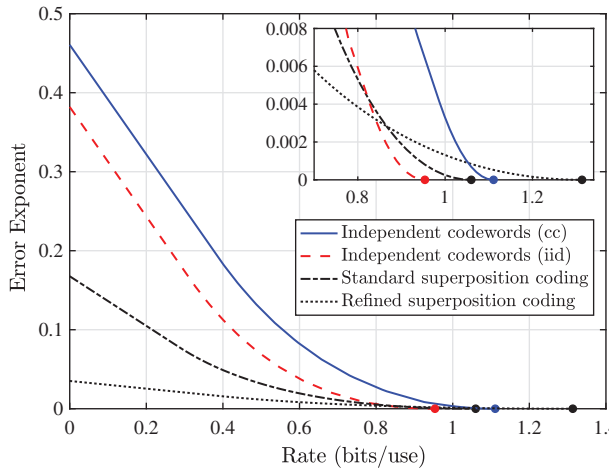


Figure 7.2: Error exponents of superposition coding for the example in Section 6.4.4, using the input distribution optimized for the refined superposition coding rate. The exponents under independent codewords are also shown, in which the input distribution is chosen to optimize the corresponding rate (GMI or LM rate). The achievable rates are marked on the horizontal axis.

i.i.d. vs. constant-composition random coding), and suggests an interesting open problem of finding an ensemble that naturally attains the best of both worlds.

We conclude this discussion by mentioning another example highlighting that the error exponents attained by multi-user coding techniques may have no general ordering (i.e., neither exponent can be weakened to one another). In [93], a *cognitive* variant of the multiple-access channel was considered, in which one user has access to both messages. Under this setting, the standard superposition coding technique was compared to a random binning technique, and it was shown that neither of the two dominate each other in general. This suggests that unlike most achievable rates, one may not typically expect to be able to weaken one mismatched decoding error exponent to another without imposing additional assumptions.

8

Upper Bounds on the Mismatch Capacity

8.1 Introduction

In contrast with the extensive results on achievable rates that we surveyed in the preceding sections, studies of mismatch capacity upper bounds (i.e., converse results) have produced relatively fewer results. In this section, we survey the main upper bounds that are known to date. We focus on discrete memoryless channels (DMCs), as considered in Section 2.

In Section 8.2, we overview some early observations that were made having a converse flavor. We then return to the Csiszár–Narayan conjecture on the multi-letter extension of the LM rate (*cf.*, Section 2.7) in Section 8.3, and give some partial results suggesting its validity. In Section 8.4, we provide related results for other decoders beyond the standard maximum-metric decoder. Additional multi-letter upper bounds are presented in Section 8.5, and we conclude by presenting a recent single-letter upper bound in Section 8.6.

Before proceeding, we recall some of the notation from Sections 1 and 2. The mismatch capacity (Definition 1.1) is denoted by C_{M}. We will frequently make use of the LM rate $I_{\text{LM}}(Q_X)$, whose primal expression

(2.15) is repeated as follows:

$$I_{\mathrm{LM}}(Q_X) = \min_{\substack{\widetilde{P}_{XY}\in\mathcal{P}(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_X=Q_X,\widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)]\geq\mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y), \tag{8.1}$$

where $P_{XY} = Q_X \times W$. The input-optimized LM rate (*cf.*, (2.17)) is defined as

$$C_{\mathrm{LM}} = \max_{Q_X\in\mathcal{P}(\mathcal{X})} I_{\mathrm{LM}}(Q_X), \tag{8.2}$$

and its k-letter extension (*cf.*, (2.121)) is given by

$$C_{\mathrm{LM}}^{(k)} = \frac{1}{k}\max_{Q_{X^k}} I_{\mathrm{LM}}(Q_{X^k}, W^k, q^k). \tag{8.3}$$

We will often make the dependence of the preceding quantities on the channel W and decoding metric q explicit by writing $I_{\mathrm{LM}}(Q_X, W, q)$, $C_{\mathrm{LM}}(W, q)$, $C_{\mathrm{M}}(W, q)$, $C_{\mathrm{LM}}^{(k)}(W, q)$, and so on.

This section is predominantly based on the works of Csiszár and Narayan [26], Somekh-Baruch [91, 92], and Asadi Kangarshahi and Guillén i Fàbregas [11].

8.2 Initial Results

Perhaps the most elementary known result with a converse flavor is that of [26, Thm. 2], which gives a necessary and sufficient condition for the positivity of the mismatch capacity. This result was stated in Lemma 2.6, and is repeated as follows in a slightly different form.

Lemma 8.1 (Condition for Positive Mismatch Capacity)**.** For any mismatched DMC (W, q), we have

$$C_{\mathrm{M}} > 0 \iff C_{\mathrm{LM}} > 0, \tag{8.4}$$

with both inequalities holding if and only if there exists $Q_X \in \mathcal{P}(\mathcal{X})$ such that

$$\mathbb{E}_{Q_X\times W}[\log q(X, Y)] > \mathbb{E}_{Q_X\times P_Y}[\log q(X, Y)], \tag{8.5}$$

where P_Y is the Y-marginal of $P_{XY} = Q_X \times W$.

While this result precisely establishes the conditions under which $C_{\mathrm{M}} = 0$ vs. $C_{\mathrm{M}} > 0$, it does not provide any further information about how high C_{M} may be when it is positive; for a given value $C > 0$ of the matched capacity, all we can immediately deduce is that $C_{\mathrm{M}} \in (0, C]$.

As we saw in Section 2.4.1, the mismatch capacity is completely characterized in the special case of binary channels, i.e., $|\mathcal{X}| = |\mathcal{Y}| = 2$ [26]. We formalize this statement in the following lemma, in which we let C denote the matched capacity and write the sign function as

$$\operatorname{sign}(z) = \begin{cases} 1 & z > 0 \\ 0 & z = 0 \\ -1 & z < 0, \end{cases} \tag{8.6}$$

where we also allow $\operatorname{sign}(\infty) = 1$ and $\operatorname{sign}(-\infty) = -1$.

Lemma 8.2 (Mismatch Capacity of Binary Channels)**.** For any mismatched DMC (W, q) with $|\mathcal{X}| = |\mathcal{Y}| = 2$, we have

$$C_{\mathrm{M}} = C_{\mathrm{LM}} = \begin{cases} C & \operatorname{sign}\left(\log \dfrac{q(1,1)q(2,2)}{q(1,2)q(2,1)}\right) = \operatorname{sign}\left(\log \dfrac{W(1|1)W(2|2)}{W(2|1)W(1|2)}\right) \\ 0 & \text{otherwise,} \end{cases} \tag{8.7}$$

where C is the matched capacity of W.

Throughout the monograph, we have presented several ensemble tightness results [65], upper bounding the maximal rate for which the random-coding error probability $\bar{p}_{\mathrm{e}}$ tends to zero for a given random-coding ensemble. While such results do not provide upper bounds on C_{M}, they are helpful in ruling out the possibility of weaknesses in the proofs of the achievable rates. For convenience, Lemma 2.9 is re-stated as follows (see also Section 3.7 regarding the GMI for continuous-alphabet channels).

Lemma 8.3 (Ensemble Tightness)**.** For any mismatched DMC (W, q) and any given input distribution $Q_X \in \mathcal{P}(\mathcal{X})$, we have the following:

- Under i.i.d. random coding, $\bar{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor) \to 1$ as $n \to \infty$ for any $R > I_{\mathrm{GMI}}(Q_X)$;

- Under constant-composition random coding, $\overline{p}_{\mathrm{e}}(n, \lfloor e^{nR} \rfloor) \to 1$ as $n \to \infty$ for any $R > I_{\mathrm{LM}}(Q_X)$.

Another result having a converse flavor is that of [65, Lemma 1], which we briefly outline as follows. Consider the special case that $q(x, y)$ represents a conditional distribution, written as $V(y|x)$ to highlight this distinction. Consider a constant-composition codebook whose codewords have type Q_X,[1] let P_Y be the induced output distribution, and for a small constant $\epsilon > 0$, let $\mathcal{T}_\epsilon^n(P_Y)$ be the set of sequences whose empirical distribution is ϵ-close to P_Y in the ℓ_∞ sense. That is, $\mathcal{T}_\epsilon^n(P_Y)$ represents a typical (high-probability) set of output sequences. For a fixed codeword $\boldsymbol{x}$, the set

$$S_d(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathcal{T}_\epsilon^n(P_Y) \colon -\log V^n(\boldsymbol{y}|\boldsymbol{x}) \leq nd\} \tag{8.8}$$

can be interpreted as a decoding sphere with normalized radius d. By a concentration argument, the transmitted codeword will, with high probability, lie near the surface of this sphere when $d = \mathbb{E}_{Q_X \times W}[\log V(Y|X)]$. In [65, Lemma 1], it is shown that if there exist e^{nR} *disjoint* decoding spheres $\{S_d(\boldsymbol{x}_j)\}_{j=1}^{e^{nR}}$ in $\mathcal{T}_\epsilon^n(P_Y)$ for some R and the preceding choice of d, then it must be the case that $R < C_{\mathrm{LM}}$. A strengthened statement is also given in [65, Lemma 2], characterizing the amount of overlap in the case that $R > C_{\mathrm{LM}}$. In any case, overlapping decoding spheres does not imply an upper bound on C_{M}, since it may be the case that the overlapping parts have low probability with respect to W^n.

Finally, we recall that an early work of Balakirsky [14] reported that $C_{\mathrm{LM}} = C_{\mathrm{M}}$ for general binary-input channels, which would considerably generalize the fact that this holds for $|\mathcal{X}| = |\mathcal{Y}| = 2$ (*cf.*, Lemma 8.2). However, as we saw in Section 6.3.4, this claim was refuted via a counterexample in which the 2-letter extension of superposition coding attains a rate exceeding C_{LM} for a binary-input ternary-output channel [76].

8.3 The Csiszár–Narayan Conjecture

As discussed in Section 2.7, applying the LM rate to the product channel W^k from $\mathcal{X}^k$ to $\mathcal{Y}^k$ yields an achievable rate $C_{\mathrm{LM}}^{(k)}$ (see (8.3)) for the

[1] This is without loss of generality, since any code has a constant-composition sub-code with the same asymptotic rate, as outlined in Section 2.2.

original channel. It is known that even when $k = 2$, this rate can exceed the LM rate $C_{\mathrm{LM}} = C_{\mathrm{LM}}^{(1)}$ (e.g., see Section 6.3.3). Additionally, since $C_{\mathrm{LM}}^{(k)}$ is achievable for any positive integer k, we deduce that

$$C_{\mathrm{LM}}^{(\infty)} \triangleq \sup_{k \in \mathbb{Z}} C_{\mathrm{LM}}^{(k)} \tag{8.9}$$

is also an achievable rate. In other words,

$$C_{\mathrm{M}} \geq C_{\mathrm{LM}}^{(\infty)}. \tag{8.10}$$

In addition, the supremum over k in (8.10) can be replaced by the limit as k tends to infinity, yielding

$$C_{\mathrm{LM}}^{(\infty)} = \lim_{k \to \infty} C_{\mathrm{LM}}^{(k)}. \tag{8.11}$$

Indeed, $C_{\mathrm{LM}}^{(k)}$ is a uniformly bounded sequence, and for every pair of integers (k_1, k_2) we have $C_{\mathrm{LM}}^{(k_1 k_2)} \geq \max\{C_{\mathrm{LM}}^{(k_1)}, C_{\mathrm{LM}}^{(k_2)}\}$, since coding of blocks of length k_1 (or k_2) is a special case of coding over blocks of length $k_1 k_2$.

Csiszár and Narayan [26] conjectured that the mismatch capacity is equal to the multi-letter extension of the LM rate, in the sense that inequality (8.10) holds with equality:

$$C_{\mathrm{M}} \stackrel{?}{=} C_{\mathrm{LM}}^{(\infty)}. \tag{8.12}$$

Even if this conjecture were proved to be true, this would not directly provide a *computable* expression for the mismatch capacity, since (i) in itself, (8.12) does not indicate how close $C_{\mathrm{LM}}^{(k)}$ is to $C_{\mathrm{LM}}^{(\infty)}$ for any finite value of k, and (ii) the optimization over Q_{X^k} in the definition $C_{\mathrm{LM}}^{(k)}(W, q) = \frac{1}{k} \max_{Q_{X^k}} I_{\mathrm{LM}}(Q_{X^k}, W^k, q^k)$ becomes computationally intractable as k increases, particularly since I_{LM} is non-concave with respect to the input distribution in general (see Lemma 2.12).

Nevertheless, the conjecture would yield some interesting observations; for instance, if (8.12) is true, then all of the following statements also hold:

(i) We can attain a rate arbitrarily close to C_{M} by employing random coding over the k-letter extension W^k (with metric q^k) for sufficiently large k.

(ii) At any rate below C_{M}, there exist sequences of codes whose error probability decays exponentially in the block length.

(iii) Mismatched maximum-metric decoding and mismatched threshold-based decoding lead to the same mismatch capacity (see Section 8.4.2).

Given the (hypothetical) validity of (8.12), the first two implications follow immediately from the fact that constant-composition random coding achieves $I_{\mathrm{LM}}(Q_X)$ (see Section 2) and yields a positive error exponent (see Section 7), and the third follows since derivations of the LM rate already exist that upper bound the maximum-metric decoding error probability by that of threshold decoding (see Section 8.4.2).

At the time of writing, the Csiszár–Narayan conjecture remains unresolved in general. It was proved in [26] that the conjecture is true in the case of the erasures-only metrics, and further partial results suggesting its validity were given by Somekh-Baruch [96]. These findings are surveyed in more detail below.

8.3.1 Erasures-Only Metrics

Recall from Section 1.2.7 that the mismatched decoding problem with metric $q(x,y) = \mathbb{1}\{W(y|x) > 0\}$ yields the zero-undetected error capacity, which is a problem of independent interest. It was shown in [26, Thm. 3] that (8.12) is true not only in this special case, but also for any decoding metric $q(x,y)$ satisfying the condition

$$q(x,y) = q_{\max}, \ \forall (x,y) \text{ such that } W(y|x) > 0, \tag{8.13}$$

where $q_{\max} = \max_{x \in \mathcal{X}, y \in \mathcal{Y}} q(x,y)$. Metrics satisfying (8.13) are referred to as *erasures-only metrics*, because they lead to an error whenever there exist two or more codewords consistent with the output, which implies that all errors are detectable. Not all metrics satisfying (8.13) yield the zero-undetected error capacity, since there could also exist (x,y) pairs such that $q(x,y) = q_{\max}$ and $W(y|x) = 0$; the metric used to recover the zero-error capacity (see Section 1.2.8) is one such example.

The validity of the Csiszár–Narayan conjecture for erasures-only metrics is formally stated as follows.

Theorem 8.4 (Mismatch Capacity for Erasures-Only Metrics)**.** For any mismatched DMC (W, q), if q is an erasures-only metric (i.e., satisfies (8.13)), then we have

$$C_{\mathrm{M}} = C_{\mathrm{LM}}^{(\infty)}. \tag{8.14}$$

Proof. We know that $C_{\mathrm{M}} \geq C_{\mathrm{LM}}^{(\infty)}$, so it suffices to establish that $C_{\mathrm{M}} \leq C_{\mathrm{LM}}^{(\infty)}$. We consider a sequence of codebooks $\mathcal{C}_k$ with block length k such that the error probability satisfies $\lim_{k\to\infty} p_{\mathrm{e}}(\mathcal{C}_k) \to 0$. For brevity, we denote $\epsilon_k = p_{\mathrm{e}}(\mathcal{C}_k)$. Here we use k instead of n for the block length to highlight the fact that we consider $C_{\mathrm{LM}}^{(k)}$ in (8.3) and choose Q_{X^k} depending on $\mathcal{C}_k$.

We first consider the case that all codewords of $\mathcal{C}_k$ are distinct, and then discuss the general case at the end of the proof. Let P_{X^k} be the equiprobable distribution over the codebook $\mathcal{C}_k$, i.e., for $\boldsymbol{x} \in \mathcal{X}^k$,

$$P_{X^k}(\boldsymbol{x}) = \begin{cases} \dfrac{1}{M} & \boldsymbol{x} \in \mathcal{C}_k \\ 0 & \text{otherwise,} \end{cases} \tag{8.15}$$

and let $P_{X^kY^k} = P_{X^k} \times W^k$ be the corresponding joint distribution under the channel W^k. For a given input distribution Q_{X^k}, we define the joint distribution $Q_{X^kY^k} = Q_{X^k} \times W^k$, yielding

$$C_{\mathrm{LM}}^{(k)} = \frac{1}{k} \max_{Q_{X^k}} \min_{\substack{\widetilde{P}_{X^kY^k}\colon\, \widetilde{P}_{X^k}=Q_{X^k}, \widetilde{P}_{Y^k}=Q_{Y^k} \\ \mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)] \geq \mathbb{E}_Q[\log q^k(X^k,Y^k)]}} I_{\widetilde{P}}(X^k;Y^k) \tag{8.16}$$

$$\geq \frac{1}{k} \min_{\substack{\widetilde{P}_{X^kY^k}\colon\, \widetilde{P}_{X^k}=P_{X^k}, \widetilde{P}_{Y^k}=P_{Y^k} \\ \mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)] \geq \mathbb{E}_P[\log q^k(X^k,Y^k)]}} I_{\widetilde{P}}(X^k;Y^k) \tag{8.17}$$

$$= \frac{1}{k} H_P(X^k) - \frac{1}{k} \max_{\substack{\widetilde{P}_{X^kY^k}\colon\, \widetilde{P}_{X^k}=P_{X^k}, \widetilde{P}_{Y^k}=P_{Y^k}, \\ \mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)] \geq \mathbb{E}_P[\log q^k(X^k,Y^k)]}} H_{\widetilde{P}}(X^k|Y^k), \tag{8.18}$$

where (8.17) follows by bounding $\max_{Q_{X^k}}$ by the specific choice $Q_{X^k} = P_{X^k}$, and (8.18) follows by expanding the mutual information.

Since we are considering the case that all codewords in $\mathcal{C}_k$ are distinct, and P_{X^k} is uniform over the codebook, we obtain

$$\frac{1}{k} H_P(X^k) = \frac{1}{k} \log |\mathcal{C}_k|. \tag{8.19}$$

The second term in (8.18) requires more effort. Since the metric satisfies (8.13), we have $\mathbb{E}_P[\log q^k(X^k, Y^k)] = k \log q_{\max}$ and $q^k(x^k, y^k) \le q_{\max}^k$ for all (x^k, y^k), so the constraint $\mathbb{E}_{\widetilde{P}}[\log q^k(X^k, Y^k)] \ge \mathbb{E}_P[\log q^k(X^k, Y^k)]$ is equivalent to $\mathbb{P}_{\widetilde{P}}[q^k(X^k, Y^k) = q_{\max}^k] = 1$.

Let $\mathcal{A}_k$ denote the set of all $y^k \in \mathcal{Y}^k$ for which there is a unique maximizer of $\max_{x^k \in \mathcal{C}_k} q^k(x^k, y^k)$. If $q^k(X^k, Y^k) = q_{\max}^k$ with probability one, then given $Y^k = y^k$ with $y^k \in \mathcal{A}_k$, the sequence X^k can only equal this unique maximizer, implying that $H_{\widetilde{P}}(X^k | Y^k = y^k) = 0$. This observation allows us to simplify the second term in (8.18) to the following:

$$\frac{1}{k} \max_{\substack{\widetilde{P}_{X^k Y^k} \colon \widetilde{P}_{X^k} = P_{X^k}, \widetilde{P}_{Y^k} = P_{Y^k}, \\ \mathbb{P}_{\widetilde{P}}[q^k(X^k, Y^k) = q_{\max}^k] = 1}} H_{\widetilde{P}}(X^k | Y^k) \tag{8.20}$$

$$= \frac{1}{k} \max_{\substack{\widetilde{P}_{X^k Y^k} \colon \widetilde{P}_{X^k} = P_{X^k}, \widetilde{P}_{Y^k} = P_{Y^k}, \\ \mathbb{P}_{\widetilde{P}}[q^k(X^k, Y^k) = q_{\max}^k] = 1}} \sum_{y^k \in \mathcal{A}_k^c} P_{Y^k}(y^k) H_{\widetilde{P}}(X^k | Y^k = y^k) \tag{8.21}$$

$$\le \frac{1}{k} \sum_{y^k \in \mathcal{A}_k^c} P_{Y^k}(y^k) \cdot \log |\mathcal{C}_k|, \tag{8.22}$$

since X^k takes one of at most $|\mathcal{C}_k|$ values under P_{X^k}.

Substituting (8.19) and (8.22) into (8.18), we obtain

$$C_{\mathrm{LM}}^{(k)} \ge \frac{1}{k} \log |\mathcal{C}_k| \cdot \left(1 - \sum_{y^k \in \mathcal{A}_k^c} P_{Y^k}(y^k)\right). \tag{8.23}$$

Now, since we assume that ties are broken as errors (*cf.*, Section 1.1), any $y^k \in \mathcal{A}_k^c$ produces an error, and hence $\sum_{y^k \in \mathcal{A}_k^c} P_{Y^k}(y^k) \to 0$ due to our assumption $\epsilon_k \to 0$. Taking $k \to \infty$ on both sides of (8.23), we deduce that

$$\limsup_{k \to \infty} \frac{1}{k} \log |\mathcal{C}_k| \le C_{\mathrm{LM}}^{(\infty)}. \tag{8.24}$$

Having shown this to be true for codebooks with distinct codewords, it is straightforward to generalize to arbitrary codebooks. Specifically, since we assume that $\lim_{k\to\infty} \epsilon_k \to 0$, we have for k sufficiently large that at least half of the codewords of $\mathcal{C}_k$ must be unique. By applying the above argument to a smaller codebook with only these unique codewords, we deduce that (8.24) still holds, since removing half the codewords amounts to a negligible loss in the rate.

Since (8.24) holds for any sequence of codebooks $\mathcal{C}_k$ with vanishing error probability, it follows that $C_{\mathrm{M}} \leq C_{\mathrm{LM}}^{(\infty)}$, as desired. □

8.3.2 A Soft Converse for Rational Metrics

A partial result was given in [96, Thm. 3] stating conditions under which any length-n code attaining error probability $p_{\mathrm{e}} \ll \frac{1}{n}$ (i.e., $np_{\mathrm{e}} \to 0$) must have a rate no higher than $C_{\mathrm{LM}}^{(\infty)}$. We refer to this as a *soft converse*, since it does not directly preclude the possibility of higher-rate codes whose error probability tends to zero at a slower speed, such as $\frac{1}{\sqrt{n}}$ or $\frac{1}{\log n}$.

We proceed by stating the result formally. We say that a decoding metric is *rational* if each of its (log-)values can be written as

$$\log q(x,y) = \frac{j(x,y)}{j_{\mathrm{denom}}} \tag{8.25}$$

for some integers $j(x,y)$ and j_{denom}.[2] In addition, we assume that the metric can only equal zero when the channel transition probability is also zero:

$$q(x,y) = 0 \implies W(y|x) = 0. \tag{8.26}$$

This is a mild assumption, since, assuming without loss of generality that all inputs are used, the mismatch capacity is zero when (8.26) fails (see the discussion following Lemma 2.5).

We again let $\mathcal{C}_k$ be a codebook with block length k, and let ϵ_k be its associated average error probability when applied to the mismatched DMC (W, q). The following result is stated in [96, Thm. 3].

[2] The denominator j_{denom} being independent of (x,y) is without loss of generality, since one could always take the lowest common denominator.

Theorem 8.5 (Soft Converse for Rational Metrics)**.** For any mismatched DMC (W, q), if the decoding metric is rational (*cf.*, (8.25)) and satisfies (8.26), and $\mathcal{C}_k$ $(k = 1, 2, ...)$ is a sequence of length-k codebooks whose average error probability $\epsilon_k = p_{\mathrm{e}}(\mathcal{C}_k)$ satisfies $\lim_{k\to\infty} k \cdot \epsilon_k = 0$, then it holds that

$$\limsup_{k\to\infty} \frac{1}{k} \log |\mathcal{C}_k| \le C_{\mathrm{LM}}^{(\infty)}. \tag{8.27}$$

Proof. The proof in [96] was based on the dual expression (2.16) for the LM rate; here we provide an alternative proof based on the primal expression (2.15), also re-using some of the general steps given in the proof of Theorem 8.4. By the same argument used therein, it suffices to prove the result for codebooks $\mathcal{C}_k$ whose codewords are all distinct; we henceforth assume that this is the case.

Let $\eta_k > 0$ denote the minimal difference between distinct k-letter log-metric values with respect to the same length-k output sequence:

$$\eta_k = \min_{\substack{(x^k, \bar{x}^k, y^k) \in \mathcal{X}^k \times \mathcal{X}^k \times \mathcal{Y}^k : \\ q^k(x^k, y^k) > q^k(\bar{x}^k, y^k)}} \Big(\log q^k(x^k, y^k) - \log q^k(\bar{x}^k, y^k) \Big). \tag{8.28}$$

Note that if the minimization in (8.28) is empty, then we must have $q(x, y) = q(\bar{x}, y)$ for all $(x, \bar{x}, y)$. This case can be excluded, as it trivially gives a mismatch capacity of zero.

We observe that the assumptions of the theorem imply the following:

1. By (8.26), we can find a number $B \in (0, \infty)$ such that the metric is bounded as follows for all (x, y):

$$\begin{aligned} &\log q(x, y) \le B, \qquad \forall x, y, \\ &\log q(x, y) \ge -B \quad \forall x, y \ \text{ s.t. } \ W(y|x) > 0. \end{aligned} \tag{8.29}$$

   Indeed, the choice $B = \max_{x,y} |\log q(x, y)| \mathbb{1}\{W(y|x) > 0\}$ is finite and satisfies both of these conditions.

2. By the assumption of a rational metric, it is straightforward to show that η_k is bounded away from zero uniformly in k; see [96, Lemma 4] for a formal argument.[3]

[3] In [96] the log-metric is normalized by $\frac{1}{k}$, so in the notation therein the statement is that the decay rate of η_k to zero is no faster than $\frac{1}{k}$ (rather than being bounded away from zero).

Hence, to prove the theorem, it suffices to establish the following more general claim: For any mismatched DMC (W, q), if the metric is bounded according to (8.29), and $\lim_{k\to\infty} \frac{k\epsilon_k}{\eta_k} = 0$, then

$$\limsup_{k\to\infty} \frac{1}{k} \log |\mathcal{C}_k| \leq C_{\mathrm{LM}}^{(\infty)}. \tag{8.30}$$

We proceed with the proof of this claim.

As was done in the proof of Theorem 8.4, we let P_{X^k} be uniformly distributed over the codebook $\mathcal{C}_k$, and define $P_{X^kY^k} = P_{X^k} \times W^k$. Recall the chain of inequalities leading to (8.18), which is re-stated here for convenience:

$$C_{\mathrm{LM}}^{(k)} \geq \frac{1}{k} H_P(X^k) - \frac{1}{k} \max_{\substack{\widetilde{P}_{X^kY^k}\colon \widetilde{P}_{X^k}=P_{X^k}, \widetilde{P}_{Y^k}=P_{Y^k}, \\ \mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)] \geq \mathbb{E}_P[\log q^k(X^k,Y^k)]}} H_{\widetilde{P}}(X^k|Y^k). \tag{8.31}$$

We fix $\widetilde{P}_{X^kY^k}$ satisfying the constraints in (8.31), and seek to simplify the condition $\mathbb{E}_{\widetilde{P}}[\log q^k(X^k, Y^k)] \geq \mathbb{E}_P[\log q^k(X^k, Y^k)]$, which is equivalent to

$$\mathbb{E}[\log q^k(\overline{X}^k, Y^k) - \log q^k(X^k, Y^k)] \geq 0 \tag{8.32}$$

for any triplet $(X^k, Y^k, \overline{X}^k)$ with marginals $(X^k, Y^k) \sim P_{X^kY^k}$ and $(\overline{X}^k, Y^k) \sim \widetilde{P}_{X^kY^k}$. In addition to these random variables, we define

$$q_{\max}^k(Y^k) = \max_{x^k \in \mathcal{C}_k} q^k(x^k, Y^k), \tag{8.33}$$

and we let $\widehat{X}^k$ be the estimate of X^k produced when the mismatched decoder is applied to Y^k (i.e., $\widehat{X}^k$ is the codeword in $\mathcal{C}_k$ corresponding to the message estimate $\hat{m}$). Then, we simplify (8.32) by considering the following:

$$\begin{aligned}
&\mathbb{E}[\log q^k(\overline{X}^k, Y^k) - \log q^k(X^k, Y^k)] \\
&= \mathbb{E}\Big[\mathbb{1}\{\widehat{X}^k \neq X^k\}\big(\log q^k(\overline{X}^k, Y^k) - \log q^k(X^k, Y^k)\big)\Big] \\
&\quad + \mathbb{E}\Big[\mathbb{1}\{\widehat{X}^k = X^k\}\big(\log q^k(\overline{X}^k, Y^k) - \log q^k(X^k, Y^k)\big)\Big]
\end{aligned} \tag{8.34}$$

$$\leq 2kB \cdot \epsilon_k + \mathbb{E}\Big[\mathbb{1}\{\widehat{X}^k = X^k\}\big(\log q^k(\overline{X}^k, Y^k) - \log q^k(X^k, Y^k)\big)\Big] \tag{8.35}$$

$$= 2kB \cdot \epsilon_k + \mathbb{E}\Big[\mathbb{1}\{\widehat{X}^k = X^k\}\big(\log q^k(\overline{X}^k, Y^k) - \log q^k_{\max}(Y^k)\big)\Big] \tag{8.36}$$

$$= 2kB \cdot \epsilon_k + \mathbb{E}\Big[\mathbb{1}\{\widehat{X}^k = X^k, q^k(\overline{X}^k, Y^k) < q^k_{\max}(Y^k)\} \times \big(\log q^k(\overline{X}^k, Y^k) - \log q^k_{\max}(Y^k)\big)\Big] \tag{8.37}$$

$$\leq 2kB \cdot \epsilon_k - \eta_k \mathbb{P}[\widehat{X}^k = X^k, q^k(\overline{X}^k, Y^k) < q^k_{\max}(Y^k)] \tag{8.38}$$

$$\leq 2kB \cdot \epsilon_k - \eta_k\big((1 - \epsilon_k) + \mathbb{P}[q^k(\overline{X}^k, Y^k) < q^k_{\max}(Y^k)] - 1\big) \tag{8.39}$$

$$= 2kB \cdot \epsilon_k + \eta_k(\epsilon_k - \mathbb{P}[q^k(\overline{X}^k, Y^k) < q^k_{\max}(Y^k)]), \tag{8.40}$$

where:

- (8.35) follows from the upper bound $\log q^k(\overline{X}^k, Y^k) - \log q^k(X^k, Y^k) \leq 2kB$, which holds due to (8.29);
- (8.36) holds since whenever $\widehat{X}^k = X^k$, one has $q^k(X^k, Y^k) = q^k_{\max}(Y^k)$ by the definition of $\widehat{X}^k$ as a codeword attaining the highest metric;
- (8.37) follows since the difference in log-metric values is trivially zero when $q^k(\overline{X}^k, Y^k) = q^k_{\max}(Y^k)$;
- (8.38) holds since the condition $q^k(\overline{X}^k, Y^k) < q^k_{\max}(Y^k)$ combined with the definition of η_k yields $\log q^k(\overline{X}^k, Y^k) - \log q^k(X^k, Y^k) \leq -\eta_k$;
- (8.39) follows from the fact that $\mathbb{P}[\mathcal{E}_1 \cap \mathcal{E}_2] \geq \mathbb{P}[\mathcal{E}_1] + \mathbb{P}[\mathcal{E}_2] - 1$ for any two events $\mathcal{E}_1$ and $\mathcal{E}_2$, and since $\mathbb{P}[\widehat{X}^k = X^k] = 1 - \epsilon_k$ under our assumption that the codewords of $\mathcal{C}_k$ are distinct.

Re-arranging (8.40), we deduce that if (8.32) holds, then

$$\mathbb{P}[q^k(\overline{X}^k, Y^k) < q^k_{\max}(Y^k)] \leq 2kB \cdot \frac{\epsilon_k}{\eta_k} + \epsilon_k, \tag{8.41}$$

or equivalently, taking one minus both sides,

$$\mathbb{P}[q^k(\overline{X}^k, Y^k) = q^k_{\max}(Y^k)] \geq 1 - \left(\frac{2kB}{\eta_k} + 1\right)\epsilon_k. \tag{8.42}$$

Letting $\mathcal{A}_k$ denote the set of y^k sequences for which the maximizer of $\max_{x^k \in \mathcal{C}_k} q^k(x^k, y^k)$ is unique, we can upper bound the left-hand side of (8.42) as follows:

$$\begin{aligned} &\mathbb{P}[q^k(\overline{X}^k, Y^k) = q^k_{\max}(Y^k)] && (8.43)\\ &\quad \le \mathbb{P}[q^k(\overline{X}^k, Y^k) = q^k_{\max}(Y^k), Y^k \in \mathcal{A}_k] + \mathbb{P}[Y^k \in \mathcal{A}_k^c] && (8.44)\\ &\quad \le \mathbb{P}[q^k(\overline{X}^k, Y^k) = q^k_{\max}(Y^k), Y^k \in \mathcal{A}_k] + \epsilon_k, && (8.45) \end{aligned}$$

where (8.46) follows since the event $\mathcal{A}_k^c$ implies a decoding error, due to the fact that ties are decoded as errors (see Section 1.1). Hence, combining (8.42) and (8.45), we have

$$\mathbb{P}[q^k(\overline{X}^k, Y^k) = q^k_{\max}(Y^k), Y^k \in \mathcal{A}_k] \ge 1 - 2\left(\frac{kB}{\eta_k} + 1\right)\epsilon_k. \quad (8.46)$$

Recall that the preceding analysis (and hence (8.46)) holds whenever $(X^k, Y^k) \sim P_{X^kY^k}$ and $(\overline{X}^k, Y^k) \sim \widetilde{P}_{X^kY^k}$ with $\widetilde{P}_{X^kY^k}$ satisfying the constraints in (8.31). Hence, reverting back from the notation $(X^k, \overline{X}^k, Y^k)$ to the notation $\mathbb{P}_{\widetilde{P}}[\cdot]$, we can write (8.46) as

$$\mathbb{P}_{\widetilde{P}}[q^k(X^k, Y^k) = q^k_{\max}(Y^k), Y^k \in \mathcal{A}_k] \ge 1 - 2\left(\frac{kB}{\eta_k} + 1\right)\epsilon_k. \quad (8.47)$$

Now, similarly to the derivation of Fano's inequality, we introduce the binary random variable

$$S = \mathbb{1}\{q^k(X^k, Y^k) = q^k_{\max}(Y^k), Y^k \in \mathcal{A}_k\}. \quad (8.48)$$

By the definition of $\mathcal{A}_k$ following (8.42), and the definition of $q^k_{\max}$ in (8.33), we see that when $S = 1$ it must be the case that X^k is the unique maximizer of the k-letter decoding metric. This means that $H_{\widetilde{P}}(X^k|Y^k, S = 1) = 0$, and hence

$$\begin{aligned} H_{\widetilde{P}}(X^k|Y^k) &\le H_{\widetilde{P}}(X^k, S|Y^k) && (8.49)\\ &= H_{\widetilde{P}}(S|Y^k) + H_{\widetilde{P}}(X^k|Y^k, S) && (8.50)\\ &\le 1 + \mathbb{P}_{\widetilde{P}}[S = 0] H_{\widetilde{P}}(X^k|Y^k, S = 0) && (8.51)\\ &\le 1 + 2\left(\frac{kB}{\eta_k} + 1\right)\epsilon_k \log|\mathcal{C}_k|, && (8.52) \end{aligned}$$

where (8.50) follows from the chain rule, (8.51) uses the fact that $H_{\widetilde{P}}(X^k|Y^k, S = 1) = 0$, and (8.52) follows by upper bounding

$\mathbb{P}_{\widetilde{P}}[S=0] = 1 - \mathbb{P}_{\widetilde{P}}[S=1]$ using (8.47), and by trivially upper bounding the conditional entropy by $\log|\mathcal{C}_k|$.

We are now ready to simplify the expression given in (8.31). As we already established in (8.19), the first term equals $\frac{1}{k}\log|\mathcal{C}_k|$ under our assumption that the codewords of $\mathcal{C}_k$ are distinct. As for the second term, we have

$$\frac{1}{k} \max_{\substack{\widetilde{P}_{X^k,Y^k}\colon\, \widetilde{P}_{X^k}=P_{X^k},\widetilde{P}_{Y^k}=P_{Y^k},\\ \mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)]\geq\mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)]}} H_{\widetilde{P}}(X^k|Y^k) \tag{8.53}$$

$$\leq \frac{1}{k} \max_{\substack{\widetilde{P}_{X^k,Y^k}\colon\, \widetilde{P}_{X^k}=P_{X^k},\widetilde{P}_{Y^k}=P_{Y^k},\\ \mathbb{P}_{\widetilde{P}}[q^k(X^k,Y^k)=q^k_{\max}(Y^k),Y^k\in\mathcal{A}_k]\geq 1-2\left(\frac{kB}{\eta_k}+1\right)\epsilon_k}} H_{\widetilde{P}}(X^k|Y^k) \tag{8.54}$$

$$\leq \frac{1}{k}\left(1+2\left(\frac{kB}{\eta_k}+1\right)\epsilon_k \cdot \log|\mathcal{C}_k|\right), \tag{8.55}$$

where (8.54) follows from the fact established above that the constraint $\mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)] \geq \mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)]$ implies (8.47), and (8.55) follows from (8.52).

Substituting (8.19) and (8.55) into (8.31), we find that

$$C_{\mathrm{LM}}^{(k)} \geq \left(1-2\left(\frac{kB}{\eta_k}+1\right)\epsilon_k\right)\cdot\frac{1}{k}\log|\mathcal{C}_k| - \frac{1}{k}. \tag{8.56}$$

Taking the limit as k tends to infinity, and recalling the assumption $\lim_{k\to\infty}\frac{k\epsilon_k}{\eta_k}=0$ stated above (8.30), we obtain (8.27) as desired. □

By the results of Section 7 (applied to the product channel), we know that at any rate below $C_{\mathrm{LM}}^{(k)}$, there exist sequences of codes whose error probability vanishes exponentially fast. Combined with Theorem 8.5, we deduce that for rational metrics, $C_{\mathrm{LM}}^{(\infty)}$ is the supremum of all rates for which such exponential decay can be attained.

8.4 Other Decoding Rules

Throughout the monograph, we have focused on the standard maximum-metric decoder defined in (1.1). In this subsection, we consider two variations of this decoder in which $C_{\mathrm{LM}}^{(\infty)}$ can be deduced as an upper

bound on the mismatch capacity: maximum-metric decoding with a margin, and threshold decoding.

8.4.1 Maximum-Metric Decoding with a Margin

We consider mismatched decoding with a δ-*margin*, meaning that the decoder selects a message whose codeword yields a normalized log-metric exceeding that of all other codewords by at least δ, i.e., the decoder chooses message $\hat{m}$ if and only if the codeword $\boldsymbol{x}^{(\hat{m})}$ is the only codeword such that

$$\frac{1}{n}\log q^n(\boldsymbol{x}^{(\hat{m})}, \boldsymbol{y}) \geq \frac{1}{n}\log q^n(\boldsymbol{x}^{(j)}, \boldsymbol{y}) + \delta, \quad \forall j \neq \hat{m}. \tag{8.57}$$

If no such codeword exists, then an error is declared. This decoder is clearly no better than the standard maximum-metric decoder, since whenever $\hat{m}$ is selected it must be the case that $\boldsymbol{x}^{(\hat{m})}$ is the maximum-metric codeword.

For this setup, the following was proved in [96].

Theorem 8.6 (Mismatch Capacity for Maximum-Metric Decoding with a Margin)**.** For any mismatched DMC (W, q), under maximum-metric decoding with a margin according to (8.57), we have the following:

(i) For any given $\epsilon > 0$, there exists $\delta > 0$ such that the rate $C_{\text{LM}}^{(\infty)} - \epsilon$ is achievable when the margin parameter is δ.

(ii) For any given $\delta > 0$ and $\epsilon > 0$, there does not exist any sequence of codes having rate $R > C_{\text{LM}}^{(\infty)} + \epsilon$ such that the error probability vanishes when the margin parameter is δ.

The first of these claims (i.e., achievability) is proved via a straightforward extension of that of the analogous claim for maximum-metric decoding, with the additional step of showing that the function

$$C_{\text{LM}}^{\text{marg}}(\delta) = \max_{Q_X} \min_{\substack{\widetilde{P}_{XY}:\ \widetilde{P}_X = Q_X, \widetilde{P}_Y = Q_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)] - \delta}} I_{\widetilde{P}}(X;Y) \tag{8.58}$$

(which equals the LM rate C_{LM} when $\delta = 0$) is continuous as $\delta \to 0$ from above. The second claim above (i.e., the converse result) is proved using similar ideas to the proof of Theorem 8.5; the details can be found in [96, Thm. 1].

8.4.2 Mismatched Threshold Decoding

Another important decoding rule that is related to the mismatched maximum-metric decoder is the mismatched threshold decoder, defined as follows: For a fixed threshold $\tau > 0$, the decoded message $\hat{m}$ is the unique message whose decoding metric is above the threshold, i.e., the decoder outputs $\hat{m}$ if and only if

$$\frac{1}{n}\log q^n(\boldsymbol{x}^{(\hat{m})}, \boldsymbol{y}) \geq \tau \quad \text{and} \quad \frac{1}{n}\log q^n(\boldsymbol{x}^{(j)}, \boldsymbol{y}) < \tau, \ \forall j \neq \hat{m}. \tag{8.59}$$

If no such codeword exists, then an error is declared. Similarly to the margin decoder introduced above, this decoder is no better than the standard maximum-metric decoder, since whenever the decoder declares $\hat{m}$ it must be the case that $\hat{m}$ yields the highest metric. In principle, the threshold τ could vary with n, and with the channel output $\boldsymbol{y}$, but the result given below has only been proved in the fixed-threshold setting.

For this setup, we have the following result [96], which is analogous to Theorem 8.6, but has additional technical assumptions in the converse part.

Theorem 8.7 (Mismatch Capacity for Threshold Decoding)**.** For any mismatched DMC (W, q), under threshold decoding according to (8.59), we have the following:

(i) For any given $\epsilon > 0$, there exists $\tau > 0$ such that the rate $C_{\text{LM}}^{(\infty)} - \epsilon$ is achievable when the threshold parameter is τ.

(ii) Suppose that there exists at least one symbol x^* such that

$$\max_{y,y' \in \mathcal{Y}} \{W(y|x^*)q(x^*, y) - W(y'|x^*)q(x^*, y')\} \neq 0. \tag{8.60}$$

Then, for any fixed threshold $\tau > 0$ and constant $\epsilon > 0$, and any sequence of codebooks (indexed by the block length n) having codewords of the same type $Q_{X,n}$ converging to a limiting distribution Q_X sufficiently fast so that $\lim_{n\to\infty} \sqrt{n} \max_x |Q_{X,n}(x) - Q_X(x)| = 0$, if $R > C_{\text{LM}}^{(\infty)} + \epsilon$, then the error probability is bounded away from zero as $n \to \infty$.

The first claim (i.e., achievability) is implicit in the derivation of the LM rate in [39] that is based on upper bounding the maximum-metric

decoding achievable rate in terms of that of threshold decoding. The idea is that by employing constant-composition random coding and setting $\tau = \mathbb{E}_P[\log q(X,Y)] - \epsilon$ for some $\epsilon > 0$, the resulting achievable rate is

$$C_{\mathrm{LM}}^{\mathrm{thresh}}(\tau) = \max_{Q_X} \min_{\substack{\widetilde{P}_{XY}\colon\ \widetilde{P}_X = Q_X, \widetilde{P}_Y = Q_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \tau}} I_{\widetilde{P}}(X;Y), \tag{8.61}$$

which can be made arbitrarily close to the LM rate C_{LM} for small enough $\epsilon > 0$.

For the second claim (i.e., the converse) we outline the main ideas of the proof as follows. First, it is shown that for successful threshold decoding, the threshold must be strictly less than $\mathbb{E}_{Q_X \times W}[\log q(X,Y)]$, where Q_X is the limiting distribution in the theorem statement. Intuitively, this is because the normalized log-metric $\frac{1}{n} \log q^n(\boldsymbol{X}, \boldsymbol{Y})$ for the transmitted codeword $\boldsymbol{X}$ concentrates around this value. This is formalized using the central limit theorem (CLT), and making use of the assumptions in the theorem: (i) the assumption (8.60) ensures that $\frac{1}{n} \log q^n(\boldsymbol{X}, \boldsymbol{Y})$ has positive variance;[4] (ii) the assumption $\lim_{n\to\infty} \sqrt{n} \max_x |Q_{X,n}(x) - Q_X(x)| = 0$ is utilized since the CLT concerns deviations on the order of $\frac{1}{\sqrt{n}}$ from the mean. Since we consider the case that τ does not depend on n, we deduce that $\tau = \mathbb{E}_{Q \times W}[\log q(X,Y)] - \epsilon'$ for some $\epsilon' > 0$.

The second idea is to relate the performance of threshold decoding to that of maximum-metric decoding with a margin (*cf.*, Section 8.4.1). Specifically, letting $p_{\mathrm{e,thresh}}^{(\tau)}$ and $p_{\mathrm{e,marg}}^{(\delta')}$ be the corresponding error probabilities, it can be shown that $p_{\mathrm{e,marg}}^{(\delta')} \leq p_{\mathrm{e,thresh}}^{(\tau)} + \mathbb{P}[\frac{1}{n} \log q^n(\boldsymbol{X}, \boldsymbol{Y}) \leq \tau + \delta']$. By the above choice of τ and the law of large numbers, the second term on the right-hand side tends to zero under the choice $\delta' = \epsilon'/2$, and the second claim of Theorem 8.7 then follows from the analogous claim for maximum-metric decoding with a margin. The details can be found in [96, Thm. 2].

[4] In fact, if the assumption (8.60) does not hold, then one can prove that $C_{\mathrm{LM}}^{(\infty)}$ upper bounds the mismatch capacity by following the approach used for erasures-only metrics, *cf.*, (8.13).

8.5 Further Multi-Letter Upper Bounds

The quantity $C_{\mathrm{LM}}^{(\infty)}$ (*cf.*, (8.9)) is multi-letter, in the sense that it is not expressed as the optimization of information measures on $\mathcal{X}$ and $\mathcal{Y}$, but rather, on $\mathcal{X}^k$ and $\mathcal{Y}^k$ (with $k \to \infty$). Further multi-letter upper bounds on C_{M} were derived in [91, 96], some in the spirit of $C_{\mathrm{LM}}^{(\infty)}$, and others building on information-spectrum methods [45, 104]. In this subsection, we overview some of the results attained.

Before proceeding, we briefly discuss the merit in deriving multi-letter upper bounds on C_{M}, which may appear to be of limited value due to being non-computable. Despite this limitation, such bounds can serve to deduce certain properties of the mismatch capacity C_{M} and related quantities. For instance, by a suitable comparison of multi-letter expressions, a precise characterization was given in [94] on how much the mismatch capacity increases under certain list decoding recovery criteria. Multi-letter bounds could also potentially serve as starting points towards further results, such as settling the Csiszár and Narayan conjecture, attaining new single-letter bounds, and so on. In fact, such benefits have already been observed in other topics in information theory, e.g., a multi-letter characterization of the interference channel capacity region [3] was used to derive single-letter outer bounds, and to characterize certain extreme points [66].

We also briefly mention that multi-letter *achievability* bounds have been established, including a multi-letter counterpart to the LM rate [39] and lower bounds on C_{M} with a similar flavor to the upper bounds summarized below [96]. These achievability results are omitted from our discussion, as the focus of this section is on converse results.

8.5.1 A Max-Min Multi-Letter Upper Bound

Substituting the LM rate expression (8.1) into (8.9), we find that $C_{\mathrm{LM}}^{(\infty)}$ can be written as

$$C_{\mathrm{LM}}^{(\infty)} = \lim_{k\to\infty} \max_{Q_{X^k}} \min_{\substack{\widetilde{P}_{X^kY^k}:\ \widetilde{P}_{X^k}=Q_{X^k},\widetilde{P}_{Y^k}=P_{Y^k},\\ \mathbb{E}_{\widetilde{P}}[\log q^k(X^k,Y^k)]\geq\mathbb{E}_P[\log q^k(X^k,Y^k)]}} \frac{1}{k} I_{\widetilde{P}}(X^k;Y^k), \tag{8.62}$$

where $P_{X^kY^k} = Q_{X^k} \times W^k$. The following upper bound on C_{M} has the same form as (8.62), but differs in the set over which the inner minimization is performed.

Theorem 8.8 (Multi-Letter Upper Bound)**.** For any mismatched DMC (W, q), we have

$$C_{\mathrm{M}} \leq \liminf_{k\to\infty} \max_{Q_{X^k}} \min_{\substack{\widetilde{P}_{X^kY^k}:\ \widetilde{P}_{X^k}=Q_{X^k},\\ P_{Y^k,q^k(X^k,Y^k)}=\widetilde{P}_{Y^k,q^k(X^k,Y^k)}}} \frac{1}{k} I_{\widetilde{P}}(X^k;Y^k), \tag{8.63}$$

where $\widetilde{P}_{Y^k,q^k(X^k,Y^k)}$ is the joint distribution of Y^k and the metric value $q^k(X^k, Y^k)$ at the output of the auxiliary channel $\widetilde{P}_{Y^k|X^k}$, and $P_{Y^k,q^k(X^k,Y^k)}$ is the corresponding joint distribution for the true channel W^k.

The idea of the proof is the following [96]: If two (multi-letter) channels lead to the same joint distribution of Y^k and $q^k(X^k, Y^k)$, then they have the same average probability of error under metric q. For any such channel, at most $I_{\widetilde{P}}(X^k;Y^k)$ bits of information can be reliably transmitted, and the result follows by minimizing over all such channels.

8.5.2 Information-Spectrum Formula for the Mismatch Capacity

Next, we turn to an *exact* formula for the mismatch capacity derived in [91] using the the information-spectrum method. While the result in [91] applies to general channels and metrics with memory, we restrict our attention to the case of mismatched DMCs here. For a sequence of real-valued random variables $\{X_i\}_{i=1}^{\infty}$, the limit inferior in probability is defined as follows

$$\mathrm{p}\text{-}\liminf_{n\to\infty} X_n = \sup\Big\{t:\ \lim_{n\to\infty} \mathbb{P}[X_n < t] = 0\Big\}. \tag{8.64}$$

In the following, we consider an n-letter input distribution $Q_{\boldsymbol{X}}$, and define the pairwise error probability random variable

$$\Phi_n(\boldsymbol{X}, \boldsymbol{Y}) = \mathbb{P}[q^n(\overline{\boldsymbol{X}}, \boldsymbol{Y}) \geq q^n(\boldsymbol{X}, \boldsymbol{Y}) | \boldsymbol{X}, \boldsymbol{Y}], \tag{8.65}$$

where $(\boldsymbol{X}, \boldsymbol{Y}, \overline{\boldsymbol{X}}) \sim Q_{\boldsymbol{X}}(\boldsymbol{x}) W^n(\boldsymbol{y}|\boldsymbol{x}) Q_{\boldsymbol{X}}(\bar{\boldsymbol{x}})$. This is the conditional probability of a codeword $\overline{\boldsymbol{X}}$ independently drawn from $Q_{\boldsymbol{X}}$ yielding a

higher decoding metric than $q^n(\boldsymbol{X}, \boldsymbol{Y})$, and this also appeared in the derivation of the GMI and LM rate (see Section 2.6).

Theorem 8.9 (Multi-Letter Expression for the Mismatch Capacity)**.** For any mismatched DMC (W, q), the mismatch capacity is given by

$$C_{\mathrm{M}} = \sup_{\boldsymbol{Q}} \ \mathrm{p}\text{-}\liminf_{n\to\infty} -\frac{1}{n} \log \Phi_n(\boldsymbol{X}, \boldsymbol{Y}), \tag{8.66}$$

where $(\boldsymbol{X}, \boldsymbol{Y}) \sim Q_{\boldsymbol{X}} \times W^n$, Φ_n is defined in (8.65), and $\boldsymbol{Q}$ denotes the infinite sequence of multi-letter input distributions $(Q_{X^1}, Q_{X^2}, Q_{X^3}, \dots)$.

The proof is based on the following observation, providing an exact expression for the error probability for a given length-n codebook $\mathcal{C}_n$ in terms of Φ_n. Let $\boldsymbol{X}$ be a random variable equiprobable over $\mathcal{C}_n$, and let $\boldsymbol{Y}$ be the output of W^n with $\boldsymbol{X}$ as the input. Then, we have

$$p_{\mathrm{e}}(\mathcal{C}_n) = \mathbb{P}\left[-\frac{1}{n} \log \Phi_n(\boldsymbol{X}, \boldsymbol{Y}) < \frac{1}{n} \log |\mathcal{C}_n|\right], \tag{8.67}$$

where Φ_n is computed with respect to $Q_{\boldsymbol{X}}$, the equiprobable distribution on $\mathcal{C}_n$. Theorem 8.9 can then be deduced using the definition of lim inf in probability. The details can be found in [91].

The exact expression of Theorem 8.9 can be upper bounded by the same expression with the limit inferior in probability replaced by the limit inferior of the series of expectations (see [91, Cor. 2]), yielding

$$C_{\mathrm{M}} \le \sup_{\boldsymbol{Q}} \ \liminf_{n\to\infty} -\frac{1}{n} \mathbb{E}\big[\log \Phi_n(\boldsymbol{X}, \boldsymbol{Y})\big]. \tag{8.68}$$

Moreover, this bound was shown be tight for finite-input channels whenever the strong converse property holds for mismatched decoding (see [91, Lemma 7]), though understanding when the latter property holds remains an open problem.

The bound (8.68) was also shown to be tight in the case of the erasures-only metric $q(x, y) = \mathbb{1}\{W(y|x) > 0\}$, in which case we have

$$\Phi_n(\boldsymbol{X}, \boldsymbol{Y}) = \mathbb{P}\big[W^n(\boldsymbol{Y}|\overline{\boldsymbol{X}}) > 0 | \boldsymbol{Y}\big]. \tag{8.69}$$

Hence, in this case, the multi-letter capacity formula becomes [91, Proposition 1]

$$C_{\mathrm{M}} = \sup_{\boldsymbol{Q}} \ \liminf_{n\to\infty} \ -\frac{1}{n} \mathbb{E}\Big[\log \mathbb{P}\big[W^n(\boldsymbol{Y}|\overline{\boldsymbol{X}}) > 0 | \boldsymbol{Y}\big]\Big]. \tag{8.70}$$

8.6 A Single-Letter Upper Bound

The task of attaining single-letter upper bounds on the mismatch capacity for general mismatched DMCs has proved to be notoriously difficult since the introduction of the mismatched decoding problem. In this subsection, we present a recent result providing a single-letter upper bound on the mismatch capacity that can be strictly smaller than the matched capacity [11] (see also [9, 10]).

The high-level idea of the upper bound is to relate the error probability of the mismatched DMC (W, q) to the error probability associated with an auxiliary channel $P_{\overline{Y}|X}$ under optimal maximum-likelihood (ML) decoding. Then, a standard mutual information based upper bound on the capacity of $P_{\overline{Y}|X}$ translates to a non-trivial upper bound on the mismatch capacity of (W, q).

In order to state the upper bound, we introduce some definitions. For any two outputs $y, \overline{y} \in \mathcal{Y}$, we define the following set of inputs maximizing the difference of log-metric values:

$$\mathcal{X}_q^*(y, \overline{y}) = \Big\{ x \in \mathcal{X} \colon x \in \arg\max_{x' \in \mathcal{X}} \big(\log q(x', \overline{y}) - \log q(x', y) \big) \Big\}. \quad (8.71)$$

In addition, we define

$$\mathcal{M}_{\max}(q) \triangleq \Big\{ P_{Y\overline{Y}|X} \in \mathcal{P}(\mathcal{Y}^2|\mathcal{X}) \colon P_{Y\overline{Y}|X}(y, \overline{y}|x) = 0, \ \forall x \notin \mathcal{X}_q^*(y, \overline{y}) \Big\}, \quad (8.72)$$

which is the set of conditional joint distributions such that x maximizes the log-metric difference in (8.71) with probability one. We henceforth refer to any conditional joint distribution in $\mathcal{M}_{\max}(q)$ as being *maximal*.

Theorem 8.10 (Single Letter Upper Bound on the Mismatch Capacity)**.** For any mismatched DMC (W, q), we have

$$C_{\text{M}}(W, q) \leq \overline{R}(W, q), \quad (8.73)$$

where

$$\overline{R}(W, q) \triangleq \max_{Q_X \in \mathcal{P}(\mathcal{X})} \min_{\substack{P_{Y\overline{Y}|X} \in \mathcal{M}_{\max}(q): \\ P_{Y|X} = W}} I_P(X; \overline{Y}), \quad (8.74)$$

where $I_P(X; \overline{Y})$ is computed with respect to $Q_X \times P_{\overline{Y}|X}$.

Observe that $\overline{R}(W, q)$ has a similar form to the input-optimized LM rate C_{LM} in its primal form (*cf.*, (8.1)–(8.2)), with a maximization over the input distribution and a minimization over the (joint) auxiliary channel. The defining feature here is the constraint that $P_{Y\overline{Y}|X}$ is maximal.

Before outlining the proof of Theorem 8.10, we present some examples and properties of $\overline{R}(W, q)$.

Examples. We return to the binary-input ternary-output example of Section 6.3.4, in which $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{0, 1, 2\}$, and the channel and metric are described by the entries of the $|\mathcal{X}| \times |\mathcal{Y}|$ matrices

$$\boldsymbol{W} = \begin{bmatrix} 0.97 & 0.03 & 0 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \quad \boldsymbol{q} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0.5 & 1.36 \end{bmatrix}. \tag{8.75}$$

We saw in Section 6.3.4 that the 2-letter superposition coding rate $C_{\text{SC}}^{(2)} = 0.19908$ bits/use is strictly higher than the LM rate $0.19746 \leq C_{\text{LM}} \leq 0.19751$, implying that $C_{\text{M}} > C_{\text{LM}}$ [76].

Since (8.74) minimizes over $P_{Y\overline{Y}|X}$, we can substitute any specific choice therein and still have a valid upper bound on C_{M}. We consider the choice in Table 8.1 (with all triplets not shown therein having probability zero), which is easily verified to be maximal and to satisfy $P_{Y|X} = W$. Marginalizing out the Y variable, we find that

$$\left[P_{\overline{Y}|X}(\overline{y}|x)\right]_{x\in\mathcal{X},\overline{y}\in\mathcal{Y}} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}. \tag{8.76}$$

A numerical calculation with an efficient iterative algorithm (discussed below) establishes that the joint conditional distribution in Table 8.1 attains the minimum in (8.74), and gives the following (to five decimal places):

$$\overline{R}(W, q) = 0.61823 \text{ bits/use}. \tag{8.77}$$

On the other hand, a basic numerical calculation establishes that W has matched capacity

$$C(W) = 0.71329 \text{ bits/use}. \tag{8.78}$$

Table 8.1: Non-zero entries of $P_{Y\overline{Y}|X}(y,\overline{y}|x)$ for the example in (8.75)

$(x,y,\overline{y})$	$P_{Y\overline{Y}\|X}(y,\overline{y}\|x)$	$(x,y,\overline{y})$	$P_{Y\overline{Y}\|X}(y,\overline{y}\|x)$
$(0,0,0)$	0.5	$(1,0,0)$	0.1
$(0,0,1)$	0.47	$(1,1,1)$	0.1
$(0,1,1)$	0.03	$(1,2,2)$	0.8

Hence, the upper bound $\overline{R}(W,q)$ is strictly smaller than the matched capacity, albeit still far from the best known achievable rate (i.e., lower bound on the mismatch capacity) in this example.

In addition to the above example, the following observations were made in [11]:

- If we change $q(2,2)$ from 0.5 to 1 in (8.75), then the choice of $P_{Y\overline{Y}|X}$ in Table 8.1 remains the optimal solution of the iterative algorithm, meaning that $\overline{R} = 0.61823$ bits/use. In this case, however, a numerical evaluation of the LM rate also gives $C_{\text{LM}} = 0.61823$ bits/use, and hence, the upper and lower bounds coincide (at least numerically).

- In the binary-input binary-output case $|\mathcal{X}| = |\mathcal{Y}| = 2$, the upper bound $\overline{R}(W,q)$ always equals the mismatch capacity, as characterized in Lemma 8.2.

- A sufficient condition can be given for a binary-input mismatched DMC to have $C_{\text{M}} < C$, where C is the matched capacity. To state this condition, we start with the following definition: Two sequences $\{\alpha_k\}_{k=1}^K$ and $\{\beta_k\}_{k=1}^K$ are said to have the same order if, for all $k_1, k_2 \in \{1,\ldots,K\}$, it holds that

$$\alpha_{k_1} \geq \alpha_{k_2} \Rightarrow \beta_{k_1} \geq \beta_{k_2}. \tag{8.79}$$

  Then, if $\max_{x,y} W(y|x) > 0$ and the sequences $\big\{\log W(y_k|1) - \log W(y_k|2)\big\}_{k=1}^{|\mathcal{Y}|}$ and $\big\{\log q(1,y_k) - \log q(2,y_k)\big\}_{k=1}^{|\mathcal{Y}|}$ do not have the same order, then $C_{\text{M}} < C$.

- As for a non-binary example, consider the following ternary-input quaternary-output channel and the corresponding erasures-only

metric (Section 1.2.7):

$$\boldsymbol{W} = \begin{bmatrix} 0.25 & 0 & 0.05 & 0.7 \\ 0.3 & 0.55 & 0 & 0.15 \\ 0.05 & 0.5 & 0.45 & 0 \end{bmatrix}, \quad \boldsymbol{q} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}. \tag{8.80}$$

The matched capacity of W is $C = 0.78537$ bits/use, and the above-mentioned iterative algorithm gives $\overline{R} = 0.62318$ bits/use. The LM rate computed by an exhaustive search over the input distributions is approximately $C_{\text{LM}} = 0.42922$ bits/use.

Properties of $\overline{R}(W, q)$. Here we discuss some convexity properties of $\overline{R}(W, q)$ and their implications for its computation. We first present a lemma stating that the minimization in (8.74) is performed over a convex set. The proof is straightforward, and can be found in [11].

Lemma 8.11 (Convexity of the Constraint Set)**.** For any mismatched DMC (W, q), the set of conditional joint distributions $P_{Y\overline{Y}|X} \in \mathcal{M}_{\max}(q)$ such that $P_{Y|X} = W$ is convex.

Since the mutual information is concave in Q_X and convex in $P_{\overline{Y}|X}$, Lemma 8.11 implies that the optimization problem in (8.74) is a convex-concave saddlepoint problem, and can be simplified as follows:

$$\overline{R}(W, q) = \max_{Q_X} \min_{\substack{P_{Y\overline{Y}|X} \in \mathcal{M}_{\max}(q): \\ P_{Y|X} = W}} I_P(X; \overline{Y}) \tag{8.81}$$

$$= \min_{\substack{P_{Y\overline{Y}|X} \in \mathcal{M}_{\max}(q): \\ P_{Y|X} = W}} \max_{Q_X} I_P(X; \overline{Y}) \tag{8.82}$$

$$= \min_{\substack{P_{Y\overline{Y}|X} \in \mathcal{M}_{\max}(q): \\ P_{Y|X} = W}} C(P_{\overline{Y}|X}). \tag{8.83}$$

Two further implications of the convex-concave structure of (8.74) are as follows (see [11] for details): (i) One can establish an efficient iterative algorithm to compute $\overline{R}(W, q)$; (ii) Multi-letter extensions of the bound do not yield any improvement, i.e., we have for any positive integer k that

$$\frac{1}{k}\overline{R}^{(k)}(W, q) = \overline{R}(W, q), \tag{8.84}$$

where $\overline{R}^{(k)}(W,q)$ denotes the multi-letter extension of $\overline{R}(W,q)$. This is shown by first deriving the KKT conditions for both the single-letter and multi-letter problems, and then showing that the k-fold product of any optimal distribution for the single-letter problem also satisfies the KKT conditions for the multi-letter problem.

Outline of the Derivation of $\overline{R}(W,q)$. We provide a brief overview of the proof of Theorem 8.10, and refer the reader to [11] for the details. As we mentioned previously, the key idea is to lower bound the error probability of a codebook $\mathcal{C}_n$ applied to the mismatched DMC (W,q) by that of the same codebook over a different channel $P_{\overline{Y}|X}$ with optimal (maximum-likelihood) decoding.

We first state some standard assumptions on the codebook $\mathcal{C}_n = \{\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(M)}\}$ that incur no loss of generality. Since any code contains a constant-composition sub-code with the same asymptotic rate (see Section 2.2), we may assume that $\mathcal{C}_n$ is a constant-composition codebook, say with type $Q_{X,n}$. In addition, it suffices to consider the maximal error probability $p_{\mathrm{e,max}}$ (with respect to $m=1,\ldots,M$) instead of the average error probability p_{e}, since these two notions lead to the same mismatch capacity (see Remark 1.2).

We introduce the following counterpart of (8.72):

$$\mathcal{M}_{\max,n}(q) \triangleq \Big\{ P_{Y\overline{Y}|X,n} \in \mathcal{P}_n(\mathcal{Y}^2|\mathcal{X}) : P_{Y\overline{Y}|X,n}(y,y'|x) = 0,\ \forall x \notin \mathcal{X}_q^*(y,y') \Big\}, \tag{8.85}$$

which has the same constraints as $\mathcal{M}_{\max}(q)$, but has the further constraint of being a conditional joint type. Then, for a given codeword $\boldsymbol{x} \in \mathcal{T}^n(Q_{X,n})$ and a maximal conditional joint type $P_{Y\overline{Y}|X,n} \in \mathcal{M}_{\max,n}(q)$, consider the bipartite graph

$$\mathcal{G}_{\boldsymbol{x}}(P_{Y\overline{Y}|X,n}) = \{\mathcal{T}_{\boldsymbol{x}}^n(P_{Y|X,n}), \mathcal{T}_{\boldsymbol{x}}^n(P_{\overline{Y}|X,n}), \mathcal{E}_{\boldsymbol{x}}\} \tag{8.86}$$

with left-node set $\mathcal{T}_{\boldsymbol{x}}^n(P_{Y|X,n})$, right-node set $\mathcal{T}_{\boldsymbol{x}}^n(P_{\overline{Y}|X,n})$, and edge set

$$\mathcal{E}_{\boldsymbol{x}} = \{(\boldsymbol{y},\overline{\boldsymbol{y}}) | \widehat{P}_{\boldsymbol{y}\overline{\boldsymbol{y}}|\boldsymbol{x}} = P_{Y\overline{Y}|X,n}\}, \tag{8.87}$$

with $\widehat{P}_{\boldsymbol{y}\overline{\boldsymbol{y}}|\boldsymbol{x}}$ denoting the joint conditional type induced by $(\boldsymbol{x},\boldsymbol{y},\overline{\boldsymbol{y}})$.

Thus, the graph connects $\boldsymbol{y}$, the output of the channel $P_{Y|X,n}$, to $\overline{\boldsymbol{y}}$, the output of the channel $P_{\overline{Y}|X,n}$, whenever their joint conditional type $\widehat{P}_{\boldsymbol{y}\overline{\boldsymbol{y}}|\boldsymbol{x}}$ is equal to $P_{Y\overline{Y}|X,n}$. It can be shown that the graph $\mathcal{G}_{\boldsymbol{x}}(P_{Y\overline{Y}|X,n})$ is regular, i.e., all left-nodes have the same degree, and all right-nodes have the same degree.

With the preceding definitions in place, the key step in the analysis is as follows: Show that for two constant-composition codewords $\boldsymbol{x}, \overline{\boldsymbol{x}} \in \mathcal{C}_n$ and an output sequence $\boldsymbol{y}$, if $P_{Y\overline{Y}|X,n}$ is a maximal distribution whose marginal distribution satisfies $P_{Y|X,n} = \widehat{P}_{\boldsymbol{y}|\boldsymbol{x}}$, then there exists some $\overline{\boldsymbol{y}}$ connected to $\boldsymbol{y}$ in $\mathcal{G}_{\boldsymbol{x}}(P_{Y\overline{Y}|X,n})$ for which we have

$$\widehat{P}_{\overline{\boldsymbol{y}}|\boldsymbol{x}} = \widehat{P}_{\overline{\boldsymbol{y}}|\overline{\boldsymbol{x}}}. \tag{8.88}$$

Here, as usual, $\boldsymbol{x}$ represents the transmitted codeword, and $\overline{\boldsymbol{x}}$ represents some non-transmitted codeword. Then, the following crucial lemma shows that (8.88) implies $q^n(\boldsymbol{x}, \boldsymbol{y}) \leq q^n(\overline{\boldsymbol{x}}, \boldsymbol{y})$.

Lemma 8.12 (Connection Between Conditional Type Error and Decoding Error)**.** Consider a conditional type $Q_{X,n} \in \mathcal{P}_n(\mathcal{X})$ and sequences $\boldsymbol{x}, \overline{\boldsymbol{x}} \in \mathcal{T}^n(Q_{X,n})$, and let $P_{Y\overline{Y}|X,n} \in \mathcal{M}_{\max,n}(q)$ be a maximal joint conditional type satisfying $P_{Y|X,n} = \widehat{P}_{\boldsymbol{y}|\boldsymbol{x}}$. If there exists some $\overline{\boldsymbol{y}} \in \mathcal{T}^n_{\boldsymbol{x}}(P_{\overline{Y}|X,n}) \cap \mathcal{T}^n_{\overline{\boldsymbol{x}}}(P_{\overline{Y}|X,n})$ connected to $\boldsymbol{y} \in \mathcal{T}^n_{\boldsymbol{x}}(P_{Y|X,n})$ in $\mathcal{G}_{\boldsymbol{x}}(P_{Y\overline{Y}|X,n})$, then

$$q^n(\boldsymbol{x}, \boldsymbol{y}) \leq q^n(\overline{\boldsymbol{x}}, \boldsymbol{y}). \tag{8.89}$$

Proof. First consider the following standard marginalization property for any triplet $(\boldsymbol{x}', \boldsymbol{y}, \overline{\boldsymbol{y}})$:

$$\widehat{P}_{\boldsymbol{y}\overline{\boldsymbol{y}}}(y, \overline{y}) = \sum_x \widehat{P}_{\boldsymbol{x}'\boldsymbol{y}\overline{\boldsymbol{y}}}(x, y, \overline{y}). \tag{8.90}$$

Using this property once with $\boldsymbol{x}' = \boldsymbol{x}$ and again with $\boldsymbol{x}' = \overline{\boldsymbol{x}}$, it follows that

$$\sum_x \widehat{P}_{\boldsymbol{x}\boldsymbol{y}\overline{\boldsymbol{y}}}(x, y, \overline{y}) = \sum_x \widehat{P}_{\overline{\boldsymbol{x}}\boldsymbol{y}\overline{\boldsymbol{y}}}(x, y, \overline{y}). \tag{8.91}$$

We now consider the following difference of log-metric values:

$$
\begin{aligned}
&\log q^n(\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}) - \log q^n(\overline{\boldsymbol{x}}, \boldsymbol{y}) \\
&= n \sum_{x,y,\overline{y}} \widehat{P}_{\overline{\boldsymbol{x}}\boldsymbol{y}\overline{\boldsymbol{y}}}(x, y, \overline{y}) \big(\log q(x, \overline{y}) - \log q(x, y) \big) && (8.92) \\
&\le n \sum_{y,\overline{y}} \Big(\sum_{x} \widehat{P}_{\overline{\boldsymbol{x}}\boldsymbol{y}\overline{\boldsymbol{y}}}(x, y, \overline{y}) \Big) \max_{x \in \mathcal{X}} \big(\log q(x, \overline{y}) - \log q(x, y) \big) && (8.93) \\
&= n \sum_{y,\overline{y}} \Big(\sum_{x} \widehat{P}_{\boldsymbol{x}\boldsymbol{y}\overline{\boldsymbol{y}}}(x, y, \overline{y}) \Big) \max_{x \in \mathcal{X}} \big(\log q(x, \overline{y}) - \log q(x, y) \big) && (8.94) \\
&= n \sum_{y,\overline{y}} \sum_{x} \widehat{P}_{\boldsymbol{x}\boldsymbol{y}\overline{\boldsymbol{y}}}(x, y, \overline{y}) \big(\log q(x, \overline{y}) - \log q(x, y) \big) && (8.95) \\
&= \log q^n(\boldsymbol{x}, \overline{\boldsymbol{y}}) - \log q^n(\boldsymbol{x}, \boldsymbol{y}) && (8.96)
\end{aligned}
$$

where (8.92) uses $\log q^n(\boldsymbol{x}, \boldsymbol{y}) = n \sum_{x,y} \widehat{P}_{\boldsymbol{x}\boldsymbol{y}}(x, y) \log q(x, y)$, (8.93) follows by upper bounding $\log q(x, \overline{y}) - \log q(x, y)$ by its maximum over the inputs, (8.94) follows from (8.91), (8.95) follows from the maximality of $P_{Y\overline{Y}|X,n}$ (see (8.71)) and the graph construction $\mathcal{G}_{\boldsymbol{x}}(P_{Y\overline{Y}|X,n})$ (see (8.86)–(8.87)), and (8.96) follows similarly to (8.92).

Now, the assumption $\overline{\boldsymbol{y}} \in \mathcal{T}^n_{\boldsymbol{x}}(P_{\overline{Y}|X,n}) \cap \mathcal{T}^n_{\overline{\boldsymbol{x}}}(P_{\overline{Y}|X,n})$ implies that $\log q^n(\boldsymbol{x}, \overline{\boldsymbol{y}}) = \log q^n(\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}})$, and combining this with (8.96), we get the desired result $\log q^n(\boldsymbol{x}, \boldsymbol{y}) \le \log q^n(\overline{\boldsymbol{x}}, \boldsymbol{y})$. □

The event $\widehat{P}_{\overline{\boldsymbol{y}}|\boldsymbol{x}} = \widehat{P}_{\overline{\boldsymbol{y}}|\overline{\boldsymbol{x}}}$ is referred to as a *conditional type error*, corresponding to the correct codeword and some incorrect codeword inducing the same conditional type with $\overline{\boldsymbol{y}}$. Lemma 8.12 implies that whenever there is a conditional type error in the channel $P_{\overline{Y}|X,n}$, there also is a mismatched decoding error in the channel $P_{Y|X,n}$, which in turn should be close to the original channel W by a typicality argument (recall that $P_{Y|X,n} = \widehat{P}_{\boldsymbol{y}|\boldsymbol{x}}$).

As hinted above, it can be shown that for rates R above $I_P(X; \overline{Y})$, there exists a message that yields a conditional type error in the channel $P_{\overline{Y}|X,n}$, and hence a mismatched decoding error in the channel $P_{Y|X,n}$. Since this holds for every maximal $P_{Y\overline{Y}|X,n} \in \mathcal{M}_{\max,n}(q)$ for which $P_{Y|X,n}$ is close to W, we can choose such a $P_{Y\overline{Y}|X,n}$ with the minimal value of $I_P(X; \overline{Y})$, in accordance with the minimization present in (8.74). Then, the remainder of the proof revolves around the technicalities related to passing from types to general distributions.

9

Overview of Other Topics

In this monograph, we have focused our attention on what we view as the most fundamental topics in mismatched decoding. Naturally, the material that we have presented is far from exhaustive. In this section, we provide a brief overview of several other topics in the literature that we did not cover, pointing the reader to the relevant references.

9.1 Alternative Achievable Rates and Capacity Notions

Generalized Cutoff Rate. For the standard single-user mismatched decoding problem (*cf.*, Section 2), several early works focused on an achievable rate known as the generalized cutoff rate [29, 49, 68, 75], defined as follows in analogy with the GMI in (2.10):

$$R_{\mathrm{GCR}}(Q_X) = \sup_{s \geq 0} \log \sum_{x,y} Q_X(x) W(y|x) \frac{q(x,y)^s}{\sum_{\overline{x}} Q_X(\overline{x}) q(\overline{x},y)^s}. \tag{9.1}$$

This can be thought of as the rate obtained by naively upper bounding the random-coding error probability under the i.i.d. ensemble by $(M-1)\mathbb{P}[q^n(\overline{\boldsymbol{X}},\boldsymbol{Y}) \geq q^n(\boldsymbol{X},\boldsymbol{Y})]$, where $(\boldsymbol{X},\boldsymbol{Y},\overline{\boldsymbol{X}}) \sim Q_X^n(\boldsymbol{x})W^n(\boldsymbol{y}|\boldsymbol{x})Q_X^n(\overline{\boldsymbol{x}})$. By applying Jensen's inequality in (2.10) we

find that $R_{\mathrm{GCR}}(Q_X) \leq I_{\mathrm{GMI}}(Q_X)$, and it was shown in [53, Sec. III] that the gap between the two can be arbitrarily large.

Alternative Notions of Mismatch Capacity. In [65], it was shown that seemingly innocuous adjustments to the definition of the mismatch capacity can drastically affect what rates are achievable. In particular, variations based on bit-error probability (instead of block error probability), randomized decoding, and encoders with added flexibility were all shown to potentially increase the mismatch capacity. As an example, we elaborate on the last of these: If the true channel is binary and noiseless, but the decoder adopts a decoding metric corresponding to the channel that deterministically flips its input, then the classical mismatch capacity is zero. However, if the encoder is allowed to pre-process the selected codeword before passing it through the channel, it can simply flip each bit prior to transmission, so that a rate of 1 bit/use is trivially attained.

9.2 Continuous and Fading Channels

Inter-Symbol Interference. The vast majority of the mismatched decoding literature focuses on memoryless channels, as these are the most convenient to analyze mathematically. However, an ongoing line of works has studied the important class of inter-symbol interference noise channels under mismatched decoding [1, 47, 70, 72] (see also [71] for more general linear channel models). In particular, ensemble-tight achievable rates were given in [47] for an autoregressive random-coding ensemble and a counterpart with auxiliary costs, generalizing the rates obtained in [1] for the i.i.d. and shell ensembles.

Block Fading Channels. The block fading channel model is a widely-adopted model in theoretical studies of wireless communication, in which a multiplicative noise variable stays constant within fixed-size blocks, but then independently changes between blocks. In [12, 13], various performance measures for block fading channels were studied under mismatched nearest-neighbor decoding resulting from imperfect channel state information. In particular, a notion of generalized outage

probability was studied, characterizing the probability that the GMI is below a certain target rate.

Output Processing. In [109], a variation of the mismatched decoding problem with additive noise channels was considered, in which a quantizer is applied to the output Y before being passed to the decoder. Adopting the GMI as the performance measure, the authors studied the problem of optimizing the quantizer subject to suitable constraints in order to maximize the GMI. In addition, gains in the GMI resulting from faster-than-Nyquist sampling of the output were characterized.

Rescaled Codewords at the Transmitter. In [32], it was shown that a simple re-scaling operation at the transmitter can increase the error exponent of i.i.d. random codes for a Gaussian channel with a given discrete constellation at the input. Specifically, the transmitter simply scales the codeword so that the power constraint is met with equality, whereas the decoder still uses the original codebook and is therefore mismatched. This technique provides a simple procedure for reaping some (but not all) of the gains that constant-composition codes provide over i.i.d. codes.

9.3 Refined Asymptotics

Fixed-Error Asymptotics and Moderate Deviations. In Section 7, we focused on error exponents, which characterize the exponential speed of decay of the error probability at fixed rates. In [79], building on recent studies in the matched setting [69, 100], two other notions of refined asymptotics were studied: (i) fixed-error asymptotics, in which one fixes a target error probability and studies the speed of convergence to the "first-order" achievable rate; (ii) moderate deviations asymptotics, in which the rate approaches its asymptotic limit and the error probability approaches zero simultaneously. In addition, a saddlepoint approximation was introduced (see also [60, 61]) that unifies the asymptotics for both fixed and varying rates under i.i.d. random coding.

Refined Asymptotics for Gaussian Sources and Channels. In [84], the fixed-error asymptotics were studied for additive non-Gaussian channels (see Section 3.8.2), providing a refined characterization of Lapidoth's study of the achievable rates [53]. The rate-distortion counterpart (see Section 4.8) was studied in [111], and [112] unified these two lines of works by considering joint-source channel coding with Gaussian coding techniques applied to general sources and channels.

9.4 Modified Channel Coding Settings

Capacity Per Unit Cost. In the presence of a system cost constraint (*cf.*, Section 3.2), as an alternative to considering the capacity with a given cost constraint, one can seek to obtain the highest possible capacity per unit cost for a fixed cost function without a hard constraint. This problem was studied in the mismatched decoding setting in [39]. It was shown that the mismatch capacity per unit cost is achieved in the low-cost limit when a zero-cost input exists, but somewhat surprisingly, the counterpart to this result no longer holds in general when the mismatch capacity is replaced by the LM rate.

Mismatched Decoding with Feedback. In [55], the problem of channel coding with feedback and mismatched decoding was introduced. It was shown that feedback can strictly increase the mismatch capacity, and sufficient conditions were given under which the mismatch capacity equals the matched capacity. A related study on the zero-undetected error capacity with feedback was given in [20], where it was shown that the maximal achievable rate is always equal to either zero or the matched capacity.

Channels with a State. In [33], various random-coding based achievable rates were given for channels with a state, taking the form $W(y|x,s)$ when the state takes the value s. In analogy with the matched setting [30, Ch. 7], the achievable rates vary according to whether the state is known at the encoder and/or decoder, and whether it is known causally or non-causally. This setup is also directly related to fading channels,

but the focus in [33] is on discrete memoryless channels, leading to different analysis techniques.

Identification via Channels. The problem of identification via channels seeks to allow a decoder to reliably determine whether or not a specific pre-specified message was transmitted, which is a weaker requirement than determining an arbitrary message. In [95], this problem was studied in the presence of mismatched decoding. Generalizing a classical result [4], it was shown that the identification problem can be performed reliably whenever the number M of messages satisfies $\lim_{n\to\infty} \log\log M < I_{\mathrm{LM}}(Q_X)$, i.e., a double-exponentially large number of messages.

9.5 Other Decoders

Stochastic Likelihood Decoding. In the matched setting, an alternative to maximum-likelihood decoding is stochastic likelihood decoding [108], in which each codeword is selected with probability proportional to its likelihood. In [80], ensemble-tight rates and error exponents were studied for a mismatched variant of this decoder in which the likelihood $W^n(\boldsymbol{y}|\boldsymbol{x})$ is replaced by a single-letter decoding metric $q^n(\boldsymbol{x},\boldsymbol{y})$. These results were further generalized to certain multi-letter metrics in [64], where it was additionally shown that this decoder provides a novel approach to deriving expurgated error exponents.

Threshold Decoding. Another well-known alternative to the maximum-likelihood decoder is the threshold decoder: Search for a unique threshold whose likelihood exceeds some threshold, or declare an error if no such codeword exists. This decoding rule was studied in [26, 60, 91] in distinct directions: [26] posed the question of whether mismatched threshold-decoding and maximum-metric decoding yield the same mismatch capacity (which remains unsolved in general), [60] developed a saddlepoint approximation of the random-coding error probability, and [91] established a multi-letter expression for the mismatch capacity, as we outlined in Section 8.4.2.

Successive Decoding. In Section 5, we considered maximum-metric decoding for the multiple-access channel (MAC). In [82], a mismatched multi-letter successive decoding rule was instead adopted, with the two decoding steps taking the form

$$\hat{m}_1 = \arg\max_i \sum_j q^n(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y}), \tag{9.2}$$

$$\hat{m}_2 = \arg\max_j q^n(\boldsymbol{x}_1^{(\hat{m}_1)}, \boldsymbol{x}_2^{(j)}, \boldsymbol{y}). \tag{9.3}$$

The first step is motivated by the fact that substituting $q = W$ in the first step yields the optimal rule for estimating m_1 alone. It was demonstrated that for a given random-coding ensemble, there exist rate pairs that can be attained via successive decoding but not maximum-metric decoding, and vice versa.

List Decoding. The list decoding criterion allows the decoder to output a list of messages of a certain size, with the decoding considered successful if the true message is included in the list. A mismatched variant of list decoding was studied in [94], with an emphasis on multi-letter formulas analogous to those given in Section 8. Among other things, it was shown that when the list size grows exponentially as $e^{n\Theta}$, the resulting increase in the mismatch capacity is exactly Θ. In particular, if the list size grows sub-exponentially fast, then the mismatch capacity is the same as that of regular decoding.

Universal Decoding. One of the earliest works in the mismatched decoding literature [99] sought to adopt a single decoding metric that simultaneously works well for all channels in some class. In a more recent work [63], a complementary problem is studied, in which some *class of decoding metrics* is fixed, and the goal is to construct a channel-independent decoder that is guaranteed to work nearly as well as the best decoding metric in the class. More specifically, for a given random coding ensemble, it is shown that the proposed universal decoder has an error probability within a sub-exponential factor of the best fixed decoder in the class.

9.6 Mismatch in Practical Codes

Huffman Coding. Perhaps the most well-known example of mismatch in information theory is that for Huffman coding in lossless symbol-wise source coding [22, Sec. 5.6]: If the true source distribution is P_X but Huffman coding is performed according to some Q_X, then the resulting average code length L_{avg} satisfies $H(P_X) + D(P_X\|Q_X) \leq L_{\mathrm{avg}} < H(P_X) + D(P_X\|Q_X) + 1$. When $P_X = Q_X$, this reduces to the well-known bound $H(P_X) \leq L_{\mathrm{avg}} < H(P_X) + 1$.

Viterbi Decoding. The Viterbi algorithm [35] is a classical technique for the maximum-likelihood decoding of convolution codes, and related decoding tasks involving Markov models. In [59], a mismatched variant of the Viterbi algorithm was analyzed for convolutional coding over a symmetric channel, using a symmetric decoding metric that may differ from the true likelihood function. The main result is a lower bound on the bit-error probability; in broad scenarios, this bound is asymptotically tight in the limit of a high signal-to-noise ratio.

Polar Codes. In a recent line of works [5–7], mismatched decoding is analyzed in the context of polar codes over binary-input discrete memoryless channels. The achievable rate derived resembles the dual form of the GMI (see (2.10)), but with two notable modifications: (i) The input distribution is always equiprobable, i.e., $Q_X(0) = Q_X(1) = \frac{1}{2}$; (ii) The optimization over $s \geq 0$ is replaced by the specific value $s = 1$ coinciding with Fisher's rate [34]. This rate can match the GMI in certain special cases, but in general may be significantly lower, or even negative.

10

Conclusion

In this monograph, we have surveyed a wide range of classical results and recent developments on the problem of point-to-point communication with mismatched decoding, as well as the analogous multiple-access communication problem, and that of mismatched encoding in rate-distortion theory. In the following, we provide a brief summary of each section, and list several open problems of interest.

In Section 2, we studied the rates achievable by random coding with independent codewords for mismatched DMCs, with an emphasis on the i.i.d. ensemble and the constant-composition ensemble. We introduced a number of properties of the resulting achievable rates, referred to as the GMI and the LM rate. In Section 3, we presented the generalization of these rates to continuous-alphabet channels, and introduced the cost-constrained random-coding ensemble as a means for proving the achievability of the LM rate. In addition, we evaluated these rates on a variety of additive noise channels, including the prominent examples of: (i) Gaussian codes for additive non-Gaussian noise channels; and (ii) fading channels with mismatched channel state information at the decoder.

In the case of discrete alphabets, ensemble tightness results are well-established for both the GMI and the LM rate for their respective ensembles, whereas for continuous-alphabet channels, only the ensemble tightness of the GMI is known (see Section 3.7). This raises the following open problem, for which we expect the answer to be affirmative.

Open Problem 1. Is the LM rate tight with respect to the ensemble average for the cost-constrained ensemble for continuous-alphabet memoryless channels with mismatched decoding?

The following question is relatively open-ended, and serves to highlight the fact that the GMI has been the focus of many specific continuous channels, with the potential gains of the LM rate remaining elusive other than the simple improvement given in Theorem 3.9.

Open Problem 2. To what extent can the LM rate improve over the GMI for additive noise channels and/or fading channels with mismatched decoding?

In Section 4, we turned to the problem of rate distortion theory with mismatched encoding, and studied an achievable rate based on random coding that can be viewed as a counterpart to the LM rate for channel coding. Despite this analogy, there are several concepts and techniques that have remained unexplored in this context despite being studied extensively in channel coding. We summarize three notable examples as follows.

Open Problem 3. Does there exist an equivalent "dual" expression for the achievable mismatched distortion-rate function derived in Section 4.3 based on constant-composition random coding? If so, can it also be attained via a suitably-designed cost-constrained random-coding ensemble?

Open Problem 4. For the rate-distortion problem with mismatched encoding, what rates can be achieved for continuous sources?

Open Problem 5. For the rate-distortion problem with mismatched encoding, what rates can be achieved via multi-user coding techniques, and to what extent do these rates improve over the rates attained using independent codewords?

Regarding Open Problems 3 and 4, we note that certain dual expressions and continuous-alphabet extensions were indeed stated in Section 4.7, but only for the case of a single (known) distortion measure and a mismatched random coding distribution, as opposed to the setup of Section 4.3 in which the encoder adopts a mismatched distortion measure.

In Section 5, we introduced the mismatched multiple-access channel, and argued that it is not only of independent interest, but provides valuable insight and results concerning point-to-point communication. In Section 6, we took this idea further by studying multi-user coding techniques for point-to-point communication in more detail, with an emphasis on superposition coding and its refined variant.

Despite having established several analogies between the achievable rates of these sections and the LM rate, there remain several ways in which our understanding of the former is comparatively lacking. For instance, while Lemma 2.7 provides precise conditions under which the LM rate equals the matched capacity, no analog has been established for the rates of multi-user coding techniques.

Open Problem 6. Under what conditions does the achievable rate region of Section 5.3 recover the matched capacity region? Similarly, under what conditions do the superposition coding rates of Sections 6.3 and 6.4 attain the matched capacity?

In addition, while refined superposition coding is known to attain a rate at least as high as standard superposition coding, and the gap can be strict for a given input distribution, little is known regarding how these rates compare when the input distribution is optimized.

Open Problem 7. Do there exist mismatched DMCs for which the refined superposition coding rate exceeds that of standard superposition coding even after the optimization of the auxiliary alphabet and the input distribution?

In addition, there is no apparent reason to believe that the refined superposition coding rate equals the mismatch capacity general, leading us to ask the following.

Open Problem 8. Do there exist multi-user random coding techniques, or alternative coding techniques, that attain rates higher than that of refined superposition coding?

In Section 7, we surveyed the notion of error exponents for mismatched memoryless channels, and saw that their study can lead to various insights not seen in the study of achievable rates alone. In fact, it was observed that the random-coding ensembles attaining higher rates are often coupled with smaller exponents at low rates, raising the following open problem.

Open Problem 9. Do there exist random coding techniques that are suited to achieving both high rates and high error exponents?

In addition, the following basic question proposed by Csiszár and Narayan [26] still remains open.

Open Problem 10. Is it always possible to attain exponentially decaying error probability for all rates below the mismatch capacity?

In Section 8, we turned to upper bounds on the mismatch capacity, i.e., converse results, for which there have been comparatively few results compared to lower bounds. We paid particular attention to the multi-letter achievable rate $C_{\mathrm{LM}}^{(\infty)}$, obtained by considering the k-letter improvement of the LM rate in the limit as $k \to \infty$. We extensively discussed the Csiszár–Narayan conjecture, which remains unresolved in general, and is stated in the following.

Open Problem 11. Does the multi-letter achievable rate $C_{\mathrm{LM}}^{(\infty)}$ equal the mismatch capacity C_{M} for arbitrary mismatched DMCs?

As we mentioned in Section 6.3.4, the early converse result of Balakirsky [14] claiming that the tightness of the LM rate was refuted via a numerical counter-example in [76], raising the following question.

Open Problem 12. Is there a purely analytical proof that the LM rate with an optimized input distribution can be strictly less than the mismatch capacity for binary-input mismatched DMCs?

The following question asks whether the strong converse holds; this was originally posed by Csiszár and Narayan [26], and still remains unsolved.

Open Problem 13. Does the smallest possible (maximal or average) error probability tend to one for all rates strictly above the mismatch capacity? If so, does it approach one exponentially fast?

The second part of this open problem serves as a counterpart to Open Problem 10, as it concerns rates above the mismatch capacity rather than rates below the mismatch capacity. In fact, Open Problem 10 is also of direct interest in the development of converse results: An affirmative answer would mean that the soft converse of Section 8.3.2 immediately implies the corresponding regular converse.

Of course, given the apparent considerable difficulty in establishing single-letter upper bounds [9, 14, 76], the following open-ended question remains particularly important.

Open Problem 14. What additional single-letter upper bounds on the mismatch capacity can be established, and how do they compare with the best known achievable rates?

In Section 9, we highlighted several additional topics that we did not survey in detail, and many of these are still in their early stages and warrant further study. To name just two examples, we expect many further developments on channels and/or decoding metrics with memory to be possible, and we believe that additional developments in the theory of practical codes with mismatched decoding would be of considerable interest, complementing the purely information-theoretic results surveyed this monograph.

Finally, there remain both classical notions in information theory (e.g., variable-length codes) and modern applications of information theory (e.g., machine learning) for which the mismatched decoding viewpoint appears to have been largely or completely unexplored, thereby posing potentially exciting avenues for future research.

Appendix

A

Appendix

A.1 Continuity Arguments

In this appendix, we present some of the technical continuity arguments required for the primal-domain analysis of the GMI and LM rate in Section 2. We also briefly comment on the analogous arguments for the multi-user random coding techniques of Sections 5 and 6.

A.1.1 Continuity of the GMI and LM Rate (Lemma 2.5)

Here we prove Lemma 2.5, which states that the GMI and LM rate are continuous in the pair (Q_X, W) within the space of pairs satisfying $Q_X(x) > 0 \cap q(x,y) = 0 \implies W(y|x) = 0$. This assumption ensures that the distribution $P_{XY} = Q_X \times W$ yields $\mathbb{E}_P[\log q(X,Y)] > -\infty$.

We present the proof for the LM rate (taken from [26]); similar steps hold for the GMI. Fix a pair (Q_X, W), and let $(Q_{X,n}, W_n)$ be a sequence converging to (Q_X, W). To simplify the notation, we subsequently represent these pairs in terms of their joint distributions, denoted by P_{XY} and $P_{XY,n}$ respectively. The corresponding LM rates are written as $I_{\text{LM}}(P_{XY})$ and $I_{\text{LM}}(P_{XY,n})$, and we let $\widetilde{P}^*_{XY}$ and $\widetilde{P}^*_{XY,n}$ denote corresponding joint distributions achieving the minimum in (2.15).

Moreover, the constraint sets corresponding to (2.15) are written as $\mathcal{S}'(P_{XY})$ and $\mathcal{S}'(P_{XY,n})$.

We establish continuity in two steps, first showing the lower bound $\liminf_{n\to\infty} I_{\text{LM}}(P_{XY,n}) \geq I_{\text{LM}}(P_{XY})$ and then the upper bound $\limsup_{n\to\infty} I_{\text{LM}}(P_{XY,n}) \leq I_{\text{LM}}(P_{XY})$. For the lower bound, note that since the probability simplex is compact, any infinite subsequence of $\widetilde{P}^*_{XY,n}$ must have a further subsequence converging to some $\widetilde{P}^*_{XY,\infty}$. Since $P_{XY,n} \to P_{XY}$ and the constraint functions defining $\mathcal{S}'(P_{XY})$ are continuous in P_{XY}, the fact that $\widetilde{P}^*_{XY,n} \in \mathcal{S}'(P_{XY,n})$ implies that $\widetilde{P}^*_{XY,\infty} \in \mathcal{S}'(P_{XY})$. That is, $\widetilde{P}^*_{XY,\infty}$ satisfies the constraints in the minimization problem of the LM rate, and hence $I_{\text{LM}}(P_{XY}) \leq I_{\widetilde{P}^*_{XY,\infty}}(X;Y)$. Since this argument holds with $\widetilde{P}^*_{XY,\infty}$ corresponding to any arbitrary subsequence of $P_{XY,n}$, it follows that

$$\liminf_{n\to\infty} I_{\text{LM}}(P_{XY,n}) \geq I_{\text{LM}}(P_{XY}). \tag{A.1}$$

To establish the matching upper bound, we will show that there exists a sequence $\widetilde{P}_{XY,n} \in \mathcal{S}'(P_{XY,n})$ converging to $\widetilde{P}^*_{XY}$. Once this is done, the continuity of mutual information implies

$$I_{\text{LM}}(P_{XY,n}) \leq I_{\widetilde{P}^*_{XY,n}}(X;Y) \to I_{\widetilde{P}^*_{XY}}(X;Y) = I_{\text{LM}}(P_{XY}), \tag{A.2}$$

and taking $\limsup_{n\to\infty}$ on both sides completes the proof.

The above-mentioned sequence $\widetilde{P}_{XY,n}$ is selected as follows for some vanishing sequence of values $\epsilon_n > 0$:

$$\widetilde{P}_{XY,n}(x,y) = P_{XY,n}(x,y) + (1-\epsilon_n)(\widetilde{P}^*_{XY}(x,y) - P_{XY}(x,y)). \tag{A.3}$$

By the assumption that $P_{XY,n}$ converges to P_{XY}, we have that $P_{XY,n}(x,y) \geq (1-\epsilon_n)P_{XY}(x,y)$ for sufficiently large n as long as ϵ_n is chosen to decay to zero sufficiently slowly. Hence, the right-hand side of (A.3) is non-negative. In addition, since $P_{XY,n}$, $\widetilde{P}^*_{XY}$, and P_{XY} all sum to one, we can sum both sides of (A.3) to obtain $\sum_{x,y} \widetilde{P}_{XY,n}(x,y) = 1$, meaning that $\widetilde{P}_{XY,n}$ is a valid joint distribution.

Recall that the two desired properties of $P_{XY,n}$ are that $\widetilde{P}_{XY,n}(x,y) \to \widetilde{P}^*_{XY}(x,y)$, and that $\widetilde{P}_{XY,n} \in \mathcal{S}'(P_{XY,n})$. The former of these follows immediately by substituting the assumptions $P_{XY,n}(x,y) \to P_{XY}(x,y)$ and $\epsilon_n \to 0$ into (A.3), so it only remains

to check that $\widetilde{P}_{XY,n} \in \mathcal{S}'(P_{XY,n})$, i.e., the constraints defining the LM rate are satisfied under the input distribution and channel corresponding to $P_{XY,n}$.

To see that the marginal constraints $\widetilde{P}_{X,n} \in P_{X,n}$ and $\widetilde{P}_{Y,n} \in P_{Y,n}$ (*cf.*, (2.15)) are satisfied, observe that if we sum both sides of (A.3) over either x or y, the second term cancels to zero due to the fact that $\widetilde{P}^*_{XY} \in \mathcal{S}'(P_{XY})$ and hence $\widetilde{P}^*_X = P_X$ and $\widetilde{P}^*_Y = P_Y$. Similarly, for the metric constraint, we multiply both sides of (A.3) by $\log q(x,y)$ and sum over x, y. The fact that $\widetilde{P}^*_{XY} \in \mathcal{S}'(P_{XY})$ implies that $\mathbb{E}_{\widetilde{P}^*}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]$, meaning that the second term on the right-hand side of (A.3) contributes a non-negative amount, and we are left with $\mathbb{E}_{\widetilde{P}_{XY,n}}[\log q(X,Y)] \geq \mathbb{E}_{P_{XY,n}}[\log q(X,Y)]$ as required.

A.1.2 Technical Ensemble Tightness Step (Lemma 2.14)

Here we prove the technical result stated in Lemma 2.14, in which the minimization over joint types is upper bounded in terms of a matching minimization over all joint distributions. To our knowledge, the proof that we present here is the first one to hold for arbitrary mismatched DMCs. See [58, 65] for simpler proofs under the assumption that there exists some $\widetilde{P}_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ such that $\widetilde{P}_X = P_X$, $\widetilde{P}_Y = P_Y$, and $\mathbb{E}_{\widetilde{P}}[\log q(X,Y)] > \mathbb{E}_P[\log q(X,Y)]$, i.e., strict inequality in the metric constraint. This is true for most mismatched DMCs, but can sometimes fail, e.g., for the erasures-only metric introduced in Section 1.2.7.

For convenience, we write the left-hand side of (2.115) as

$$I_{\mathrm{LM},n}(P_{XY}) = \min_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_X = P_X,\, \widetilde{P}_Y = P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y). \tag{A.4}$$

We need to show that, for any joint type $P_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$ and arbitrarily small $\delta > 0$, it holds that

$$I_{\mathrm{LM},n}(P_{XY}) \leq I_{\mathrm{LM}}(P_{XY}) + \delta, \tag{A.5}$$

where

$$I_{\mathrm{LM}}(P_{XY}) = \min_{\substack{\widetilde{P}_{XY} \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_X = P_X,\, \widetilde{P}_Y = P_Y,\\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} I_{\widetilde{P}}(X;Y). \tag{A.6}$$

Fix $\widetilde{P}_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ (not necessarily a joint type) satisfying the constraints in (A.6). Since P_{XY} is a joint type by assumption, the marginal constraints in (A.6) ensure that $\widetilde{P}_X$ and $\widetilde{P}_Y$ are types. We use $\widetilde{P}_{XY}$ to construct a joint type $\widetilde{P}_{XY,n}$ as follows:

- Initialize $\widetilde{P}'_{Y|X,n}(y|x) = \frac{1}{nP_X(x)} \lfloor nP_X(x)\widetilde{P}_{Y|X}(y|x) \rfloor$; note that because we are rounding down, this in itself may not be a valid conditional distribution. The constraint $\widetilde{P}_X = P_X$ leads to $\sum_x \widetilde{P}_X(x)\widetilde{P}'_{Y|X,n}(y|x) \le P_Y(y)$ for all y, and $\sum_y \widetilde{P}'_{Y|X,n}(y|x) \le 1$ for all x, but some of these inequalities might be strict.

- Since the marginals $\widetilde{P}_X$ and $\widetilde{P}_Y$ are types, we can think of them in terms of length-n sequences $\boldsymbol{x}$ and $\boldsymbol{y}$, which we seek to suitably arrange to form a joint type $\widetilde{P}_{XY,n}$. The above-defined conditional distribution $\widetilde{P}'_{Y|X,n}$ specifies most of the pairs (x_i, y_i) (for $i = 1, \ldots, n$) corresponding to a joint type, but we are still left with some entries of $\boldsymbol{x}$ and some entries of $\boldsymbol{y}$ that have not yet been assigned. Specifically, for each (x, y) pair, $nP_X(x)\widetilde{P}'_{Y|X,n}(y|x) = \lfloor nP_X(x)\widetilde{P}_{Y|X}(y|x) \rfloor$ pairs have been assigned to (x, y), but the rest remain unassigned.

- We let N_x (respectively, N_y) be the number of unassigned entries of $\boldsymbol{x}$ with value x (respectively, of $\boldsymbol{y}$ with value y), and match these up in a way that maximizes the metric:

$$\begin{aligned} \underset{\{N_{xy}\}_{x\in\mathcal{X},y\in\mathcal{Y}}}{\text{maximize}} \quad & \sum_{x,y} N_{xy} \log q(x,y) \\ \text{subject to} \quad & \sum_x N_{xy} = N_y, \ \ \forall y, \\ & \sum_y N_{xy} = N_x, \ \ \forall x, \\ & N_{xy} \in \mathbb{Z}_{\ge 0}, \ \ \forall x, y, \end{aligned} \tag{A.7}$$

  where N_{xy} represents the number of additional occurrences of the pair (x, y) formed in this matching-up step, and $\mathbb{Z}_{\ge 0}$ is the set of non-negative integers. This is an *integer programming* (IP) problem.

Noting that we established that $\lfloor nP_X(x)\widetilde{P}_{Y|X}(y|x)\rfloor$ pairs are assigned in the first step above for each (x, y), we can add these counts together to get $\sum_{x,y}\lfloor nP_X(x)\widetilde{P}_{Y|X}(y|x)\rfloor \geq \sum_{x,y}\left(nP_X(x)\widetilde{P}_{Y|X}(y|x) - 1\right) = n - |\mathcal{X}|\cdot|\mathcal{Y}|$, so the number of unassigned pairs remaining is at most $|\mathcal{X}|\cdot|\mathcal{Y}|$. Thus, $\widetilde{P}_{XY,n}$ constructed above is $\frac{|\mathcal{X}|\cdot|\mathcal{Y}|}{n}$-close to $\widetilde{P}_{XY}$ in terms of the ℓ_1-norm (i.e., sum of absolute differences), and hence

$$I_{\widetilde{P}_{XY,n}}(X;Y) \leq I_{\widetilde{P}_{XY}}(X;Y) + \delta \tag{A.8}$$

for any $\delta > 0$ and sufficiently large n by the continuity of mutual information. Since the above findings hold for an arbitrary choice of $\widetilde{P}_{XY}$ satisfying the constraints in (A.6), establishing (A.5) now only requires showing that our constructed $\widetilde{P}_{XY,n}$ is feasible in (A.4). The marginal constraints are already satisfied by construction, so it only remains to check the inequality constraint containing the decoding metric.

If we relax the integer constraint $N_{xy} \in \mathbb{Z}_+$ to $N_{xy} \in \mathbb{R}_+$ (i.e., the non-negative reals) in (A.7), we get a *linear program* (LP). With this modification, the assignments $N_{xy} = nP_X(x)\widetilde{P}_{Y|X}(y|x) - \lfloor nP_X(x)\widetilde{P}_{Y|X}(y|x)\rfloor$ corresponding to the joint distribution $\widetilde{P}_{XY}$ become feasible, since we trivially have $\lfloor nP_X(x)\widetilde{P}_{Y|X}(y|x)\rfloor + N_{xy} = \widetilde{P}_{XY}(x, y)$, and the assumptions $\widetilde{P}_X = P_X$ and $\widetilde{P}_Y = P_Y$ coincide with the required constraints $\sum_y N_{xy} = N_x$ and $\sum_x N_{xy} = N_y$.

Since $\widetilde{P}_{XY}$ satisfies the metric constraint in (A.6) by assumption, the joint distribution corresponding to the maximizer of the LP must do so as well. As a result, the same will follow for $\widetilde{P}_{XY,n}$ if we can show that the LP and the IP have the same solution. This turns out to be a special case of the following technical yet standard result from the theory of discrete optimization.

Lemma A.1 (Linear Programming and Totally Unimodular Matrices)**.** [67, Sec. III.1] Consider a linear program (LP) of the form

$$\text{maximize}_{\boldsymbol{z}}\ \boldsymbol{c}^T\boldsymbol{z} \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{z} = \boldsymbol{b}, \boldsymbol{z} \succeq 0 \tag{A.9}$$

for some matrix $\boldsymbol{A}$ and vectors $(\boldsymbol{b}, \boldsymbol{c})$, where $\succeq$ denotes element-wise inequality. We have the following:

1. If the matrix $\boldsymbol{A}$ is totally unimodular (i.e., any square sub-matrix has determinant 1, -1, or 0), and the vector $\boldsymbol{b}$ is integer-valued, then any maximizer $\boldsymbol{z}^*$ in (A.9) must be integer valued.

2. A sufficient condition for $\boldsymbol{A}$ to be totally unimodular is the following: The rows of $\boldsymbol{A}$ index the nodes of an undirected bipartite graph, the columns of $\boldsymbol{A}$ index the corresponding edges, and the (i, j)-th entry of $\boldsymbol{A}$ equals $\mathbb{1}\{\text{node } i \text{ is one of the two nodes in edge } j\}$.

We now argue that the LP version of (A.7) (i.e., with $\mathbb{R}_+$ in place of $\mathbb{Z}_+$) satisfies the conditions of Lemma A.1. We first assume that $q(x, y) > 0$ for all (x, y), and then turn to the general case. When $q(x, y) > 0$ for all (x, y), the constraint matrix corresponds to the *complete* bipartite graph with left-nodes $\mathcal{X}$ and right-nodes $\mathcal{Y}$. For instance, in the case $|\mathcal{X}| = |\mathcal{Y}| = 3$, we have

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} \end{bmatrix}, \tag{A.10}$$

where the rows correspond to the 3 elements of $\mathcal{X}$ and then the 3 elements of $\mathcal{Y}$, and the columns correspond to the 9 pairs (x, y). We also see that (A.7) corresponds to an integer vector $\boldsymbol{b}$ in (A.7), containing the values N_x and N_y.

When there exist (x, y) pairs such that $q(x, y) = 0$ (and hence $\log q(x, y) = -\infty$), more care is needed in equating the relaxed variant of (A.7) with (A.9). In particular, any positive value of N_{xy} corresponding to $q(x, y) = 0$ gives an objective value of $-\infty$, which cannot be optimal. Thus, we reformulate (A.7) as optimizing only over $\{N_{xy}\colon q(x, y) > 0\}$, and always setting $N_{xy} = 0$ when $q(x, y) = 0$. In this case, the only difference in the above analysis is that the complete bipartite graph

construction exemplified in (A.10) is replaced by an *incomplete* bipartite graph construction; the edges (x, y) corresponding to $q(x, y) = 0$ are removed.

Hence, Lemma A.1 implies that the IP (A.7) and its LP relaxation have the same solution, implying that $\widetilde{P}_{XY,n}$ is feasible in (A.4), and therefore that (A.8) yields the desired claim (A.5).

A.1.3 Analog of Lemma 2.14 for the GMI

Recall that P_{XY} is assumed to be a joint type. When proving ensemble tightness for the GMI, the analog of (2.115) in Lemma 2.14 is given by

$$\begin{aligned} &\min_{\substack{\widetilde{P}_{XY} \in \mathcal{P}_n(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} D(\widetilde{P}_{XY} \| Q_X \times P_Y) \\ &\quad \leq \min_{\substack{\widetilde{P}_{XY} \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}):\ \widetilde{P}_Y = P_Y, \\ \mathbb{E}_{\widetilde{P}}[\log q(X,Y)] \geq \mathbb{E}_P[\log q(X,Y)]}} D(\widetilde{P}_{XY} \| Q_X \times P_Y) + \delta. \end{aligned} \tag{A.11}$$

Compared to the above analysis, this is relatively simple to prove. First note that since $\widetilde{P}_Y = P_Y$ is fixed on both sides, we can rewrite both sides in terms of $\widetilde{P}_{X|Y}$, with the left-hand side constrained to choices such that $P_Y \times \widetilde{P}_{X|Y}$ is a joint type.

Given any $\widetilde{P}_{X|Y}$ feasible on the right-hand side of (A.11), we construct a feasible $\widetilde{P}^{(n)}_{X|Y}$ on the left-hand side, by performing *rounding* to ensure that each value $nP_Y(y)\widetilde{P}_{X|Y}(x, y)$ is integer-valued. More specifically, for each $y \in \mathcal{Y}$, we sort the values $\{q(x, y)\}_{x\in\mathcal{X}}$ in increasing order, and round the values $\widetilde{P}_{X|Y}(x|y)$ up for the x values with the highest metric, while rounding down for the x values with the lowest metric. This requires choosing the appropriate number of values (between 0 and $|\mathcal{X}|$) to round up vs. down to ensure that the probabilities sum to one. A feasible approach is to work iteratively, rounding another value up (respectively, down) whenever the current sum of rounded and non-rounded values is less than or equal to one (respectively, greater than one).

By doing so, we are guaranteed that $\widetilde{P}^{(n)}_{X|Y}$ yields a higher decoding metric than $\widetilde{P}_{X|Y}$, and hence the metric constraint remains satisfied.

Moreover, the entries of $P_Y \times \widetilde{P}_{X|Y}$ are arbitrarily close to $P_Y \times \widetilde{P}^{(n)}_{X|Y}$ for large n, and hence, the change in the objective function is upper bounded by an arbitrarily small constant $\delta > 0$.

A.1.4 Extensions to Multi-User Coding Techniques

For the most part, the preceding analysis readily extends to the achievable rates obtained via multi-user coding techniques (*cf.*, Sections 5 and 6). However, the analysis of Section A.1.2 requires some non-trivial changes, which we outline here. We focus on the MAC considered in Section 5, since once this is understood, the ensembles of Section 6 can be handled similarly.

Let $g(\widetilde{P}_{X_1X_2Y})$ denote the left-hand side of (5.51), where $\mathcal{S}'_{12,n}$ is defined in (5.23). For the sake of the analysis, we only require the fact that $g(\cdot)$ is a continuous function. We wish to show that

$$\begin{aligned} &\min_{\substack{\widetilde{P}_{X_1X_2Y} \in \mathcal{P}_n(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}): \\ \widetilde{P}_{X_1} = P_{X_1}, \widetilde{P}_{X_2} = P_{X_2}, \widetilde{P}_Y = P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} g(\widetilde{P}_{X_1X_2Y}) \\ &\leq \min_{\substack{\widetilde{P}_{X_1X_2Y} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}): \\ \widetilde{P}_{X_1} = P_{X_1}, \widetilde{P}_{X_2} = P_{X_2}, \widetilde{P}_Y = P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]}} g(\widetilde{P}_{X_1X_2Y}) + \delta \end{aligned} \tag{A.12}$$

for arbitrarily small $\delta > 0$ and sufficiently large n; observe that the only difference in the two minimizations is the presence of $\mathcal{P}$ vs. $\mathcal{P}_n$. To prove (A.12), we apply the arguments of Section A.1.2 in two steps:

1. Write the minimization on the left-hand side of (A.12) as $\min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{P}_n} = \min_{\widetilde{P}_{X_1X_2} \in \mathcal{P}_n} \min_{\widetilde{P}_{Y|X_1X_2} \in \mathcal{P}_n}$, where the decoding metric constraint appears in the second minimization. By the analysis of Section A.1.2, the resulting minimum value is within $\frac{\delta}{2}$ of the corresponding minimization over all conditional distributions (not necessarily yielding a joint type) for sufficiently large n, where $\delta > 0$ is arbitrarily small.

2. Swap the order of the two minimizations from the first step, so that we have $\min_{\widetilde{P}_{Y|X_1X_2} \in \mathcal{P}} \min_{\widetilde{P}_{X_1X_2} \in \mathcal{P}_n}$, and now it is the

minimization over $\widetilde{P}_{X_1X_2}$ that contains the decoding metric constraint. Again applying the analysis of Section A.1.2, the minimum value is within $\frac{\delta}{2}$ of the corresponding minimization over all joint distributions for sufficiently large n.

A.2 Other Omitted Proofs

A.2.1 Condition for Zero Mismatch Capacity (Second Part of Lemma 2.6)

The following proof is taken from [26]. For convenience, the condition (8.5) for the positivity of the GMI and LM rate is repeated here:

$$\mathbb{E}_{Q_X \times W}[\log q(X,Y)] > \mathbb{E}_{Q_X \times P_Y}[\log q(X,Y)]. \tag{A.13}$$

In particular, we will use the fact (from the first part of the lemma) that $C_{\mathrm{LM}} > 0$ (and $C_{\mathrm{GMI}} > 0$) if and only if (A.13) holds for some input distribution Q_X, with P_Y being the resulting output distribution.

We first note that C_{M} is clearly positive whenever the GMI and/or LM rate are positive. Hence, with the first part of the lemma already established, we only need to show that $C_{\mathrm{M}} = 0$ whenever (A.13) fails for all Q_X. We prove the contrapositive statement: If $C_{\mathrm{M}} > 0$, then (A.13) holds for some Q_X.

In fact, we will show a stronger statement, namely, that (A.13) holds for some Q_X whenever the error probability can be made arbitrarily small for a codebook with just two codewords, say $\boldsymbol{x}$ and $\overline{\boldsymbol{x}}$. We write

$$p_{\mathrm{e}}(\boldsymbol{x}) = \mathbb{P}\Big[\sum_{i=1}^{n} \log q(\overline{x}_i, Y_i) \geq \sum_{i=1}^{n} \log q(x_i, Y_i)\Big] \tag{A.14}$$

with $Y_i \sim W(\cdot|x_i)$, and similarly for $p_{\mathrm{e}}(\overline{\boldsymbol{x}})$. By the assumption that the error probability can be made arbitrarily small, we have for any $\epsilon > 0$ and sufficiently large n that

$$p_{\mathrm{e}}(\boldsymbol{x}) \leq \epsilon, \text{ and } p_{\mathrm{e}}(\overline{\boldsymbol{x}}) \leq \epsilon, \tag{A.15}$$

since otherwise, the average error probability would be at least $\frac{\epsilon}{2}$.

To simplify (A.15), we use the following technical result [26, Lemma 3]: Given a finite set $\mathcal{Z} \subseteq \mathbb{R}$, for sufficiently small $\epsilon > 0$,

it holds for any n and any independent random variables $\{Z_i\}_{i=1}^n$ only taking values on $\mathcal{Z}$ that

$$\mathbb{P}\Big[\sum_{i=1}^n Z_i \geq 0\Big] \leq \epsilon \implies \mathbb{E}\Big[\sum_{i=1}^n Z_i\Big] < 0. \tag{A.16}$$

We can assume that $q(x_i, Y_i) > 0$ with probability one in (A.14), since otherwise we would have $p_{\mathrm{e}}(\boldsymbol{x}) \geq W(y|x) > 0$ for the corresponding (x, y) pair that gives $q(x, y) = 0$, contradicting the assumption of arbitrarily small error probability. If it is also the case that $q(\overline{x}_i, Y_i) > 0$ with probability one in (A.14), then we deduce from (A.16) (with $Z_i = \log q(\overline{x}_i, Y_i) - \log q(x_i, Y_i)$) that

$$\mathbb{E}\Big[\sum_{i=1}^n \log q(\overline{x}_i, Y_i) - \sum_{i=1}^n \log q(x_i, Y_i)\Big] < 0, \tag{A.17}$$

where $Y_i \sim W(\,\cdot|x_i)$. Moreover, if the preceding assumption on $q(\overline{x}_i, Y_i) > 0$ fails, then (A.17) holds trivially, since $\log q(\overline{x}, y) = -\infty$ whenever $q(\overline{x}, y) = 0$. By an analogous argument with the roles of $\boldsymbol{x}$ and $\overline{\boldsymbol{x}}$ reversed, we have

$$\mathbb{E}\Big[\sum_{i=1}^n \log q(x_i, \overline{Y}_i) - \sum_{i=1}^n \log q(\overline{x}_i, \overline{Y}_i)\Big] < 0, \tag{A.18}$$

where $\overline{Y}_i \sim W(\,\cdot|\overline{x}_i)$. Adding (A.17)–(A.18) together, we conclude that the sum of the quantities

$$\mathbb{E}\big[\log q(\overline{x}_i, Y_i) + \log q(x_i, \overline{Y}_i) - \log q(x_i, Y_i) - \log q(\overline{x}_i, \overline{Y}_i)\big] \tag{A.19}$$

over $i = 1, \ldots, n$ must be negative, and hence one of the individual terms (i.e., (A.19) for a given choice of i) must also be negative. That is, there exists a pair $(x, \overline{x})$ such that

$$\mathbb{E}\big[\log q(x, Y) + \log q(\overline{x}, \overline{Y})\big] > \mathbb{E}\big[\log q(\overline{x}, Y) + \log q(x, \overline{Y})\big] \tag{A.20}$$

with $Y \sim W(\,\cdot|x)$ and $\overline{Y} \sim W(\,\cdot|\overline{x})$. We claim that this condition implies that (A.13) holds with $Q_X(x) = \frac{1}{2}$ for $x \in \{x, \overline{x}\}$, and $Q_X(x) = 0$ elsewhere. To see this, observe that $\mathbb{E}_{Q_X \times W}[\log q(X, Y)]$ equals $\frac{1}{2}$ times the left-hand side of (A.20), whereas $\mathbb{E}_{Q_X \times P_Y}[\log q(X, Y)]$ equals $\frac{1}{4}$ times the sum of both sides of (A.20).

A.2.2 Equivalence of the GMI and LM Rate Under Output-Symmetry and Equiprobable Inputs (Lemma 2.4)

Recall that we are considering (W, q) exhibiting matching output symmetry according to Definition 2.3, along with the equiprobable input distribution $Q_X(x) = \frac{1}{|\mathcal{X}|}$. Since the channel is output-symmetric, the resulting output probability $P_Y(y)$ is the same for all y in a given subset $\mathcal{Y}_j$ from Definition 2.3 (but may vary with j).

In order to show that $I_{\text{GMI}}(Q_X) = I_{\text{LM}}(Q_X)$, we will show that the objective function in (2.16) is maximized by $a(x) = 0$ for any value of $s \geq 0$. The optimality of $a(x) = 0$ also implies the optimality of any $a(x)$ equaling a constant value not depending on x, but the value of zero is the most convenient to work with.

First note that by expanding the logarithm in (2.16), substituting the equiprobable input distribution, and omitting the term $\mathbb{E}[\log q(X, Y)^s]$ not depending on $a(\cdot)$, we are left with

$$f(s, a) \triangleq \frac{1}{|\mathcal{X}|} \sum_x a(x) - \sum_y P_Y(y) \log \left(\frac{1}{|\mathcal{X}|} \sum_{\overline{x}} q(\overline{x}, y)^s e^{a(\overline{x})} \right). \quad \text{(A.21)}$$

The partial derivative with respect to $a(x)$ is given by

$$\frac{\partial f}{\partial a(x)} = \frac{1}{|\mathcal{X}|} - \sum_y P_Y(y) \frac{q(x, y)^s e^{a(x)}}{\sum_{\overline{x}} q(\overline{x}, y)^s e^{a(\overline{x})}}. \quad \text{(A.22)}$$

We will show that the choice $a(x) = 0$ for all x makes all of these partial derivatives equal to zero. This will establish the optimality of $a(x) = 0$, since the objective function is concave in $a(\cdot)$ for fixed s, and thus any stationary point must be a global maximum. It is clear from (2.10) and (2.16) that if $a(x) = 0$ is optimal, then the GMI and LM rate are equal.

Defining $\eta_s(y) \triangleq \sum_{\overline{x}} q(\overline{x}, y)^s$, we can write (A.22) as

$$\left. \frac{\partial f}{\partial a(x)} \right|_{a(\cdot)=0} = \frac{1}{|\mathcal{X}|} - \sum_y \frac{P_Y(y)}{\eta_s(y)} q(x, y)^s. \quad \text{(A.23)}$$

It will be useful to establish that the second term in (A.23) sums to one when summed over $x \in \mathcal{X}$:

$$\sum_x \sum_y \frac{P_Y(y)}{\eta_s(y)} q(x, y)^s = \sum_y \frac{P_Y(y)}{\eta_s(y)} \sum_x q(x, y)^s = \sum_y \frac{P_Y(y)}{\eta_s(y)} \eta_s(y) = 1, \quad \text{(A.24)}$$

since $\sum_y P_Y(y) = 1$.

Since W and q are output-symmetric with the same partitions in Definition 2.3, both $P_Y(y)$ and $\eta_s(y)$ depend on y only through the subset $\mathcal{Y}_j$ it lies in. Hence, we can also rewrite (A.23) as

$$\left.\frac{\partial f}{\partial a(x)}\right|_{a(\cdot)=0} = \frac{1}{|\mathcal{X}|} - \sum_{j=1}^{k} \frac{P_Y(y_j)}{\eta_s(y_j)} \sum_{y\in\mathcal{Y}_j} q(x,y)^s, \tag{A.25}$$

where y_j is an arbitrary element of $\mathcal{Y}_j$. Again exploiting the symmetry structure of q, we see that the second term in (A.25) is the same for all x; in particular, recall that the $|\mathcal{X}| \times |\mathcal{Y}_j|$ sub-matrix corresponding to $\mathcal{Y}_j$ contains rows that are permutations of each other. By (A.24), we know that the sum of the second term in (A.23) over all $x \in \mathcal{X}$ equals one. Combining these two findings, we conclude that each term is equal to $\frac{1}{|\mathcal{X}|}$, and we attain $\left.\frac{\partial f}{\partial a(x)}\right|_{a(\cdot)=0} = 0$, as desired. As mentioned above, this establishes that $I_{\mathrm{GMI}}(Q_X) = I_{\mathrm{LM}}(Q_X)$.

A.2.3 Non-Concavity of the GMI (Lemma 2.12)

Recall the multi-letter extension $C_{\mathrm{GMI}}^{(k)}$ of the GMI introduced in Section 2.7. It was shown in [21] that, in the special case of the zero-undetected error capacity (*cf.*, Section 1.2.7), $C_{\mathrm{GMI}}^{(k)}$ approaches the mismatch capacity C_{M} in the limit as $k \to \infty$. Here we argue that this result implies that $I_{\mathrm{GMI}}(Q_X)$ is, in general, non-concave in Q_X. To our knowledge, this argument has not appeared in the existing literature.

For concreteness, let (W, q) be the mismatched DMC of Section 2.4.3, in which $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$, and the channel and decoding metric are given by

$$\boldsymbol{W} = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.75 & 0.25 \\ 0.25 & 0 & 0.75 \end{bmatrix}, \qquad \boldsymbol{q} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \tag{A.26}$$

where x indexes the rows and y indexes the columns. Observe that the channel and metric are strongly symmetric in the sense of exhibiting the output-symmetry condition of Definition 2.3 with just a single partition. We also know from Section 2.4.3 that the GMI with $Q_X = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is strictly smaller than C_{M}.

Fix k sufficiently large and Q_{X^k} such that $I_{\mathrm{GMI}}(Q_{X^k}) \geq C_{\mathrm{M}} - \delta$ for some small $\delta > 0$; this is possible due to the fact that $C_{\mathrm{GMI}}^{(k)} \to C_{\mathrm{M}}$ as $k \to \infty$. Writing $Q_{X^k} = Q_{X_1,\ldots,X_k}$ explicitly, we claim that the symmetry of (W, q) implies that

$$I_{\mathrm{GMI}}(Q_{X_1,\ldots,X_k}) = I_{\mathrm{GMI}}(Q_{\pi_1(X_1),\ldots,\pi_k(X_k)}), \tag{A.27}$$

where $Q_{\pi_1(X_1),\ldots,\pi_k(X_k)}$ is the joint distribution obtained by applying a cyclic shift (*not* an arbitrary permutation) π_j to each ternary variable X_j. For instance, if we set $k = 1$ and consider the input distribution $Q_X = (0.2, 0.3, 0.5)$, then the GMI is unchanged when we switch to $Q_X = (0.5, 0.2, 0.3)$ or $Q_X = (0.3, 0.5, 0.2)$. This can be seen via the dual expression (2.10) and the fact that applying the same cyclic shift to X and Y in (A.26) leaves each matrix unchanged.

Now, letting Π_k denote the set of all possible cyclic shifts $(\pi_1, \ldots, \pi_k)$, of which there are 3^k combinations, we have

$$\frac{1}{3^k} \sum_{(\pi_1,\ldots,\pi_k)\in\Pi_k} Q_{\pi_1(X_1),\ldots,\pi_k(X_k)} = Q_{X^k}^{\mathrm{U}}, \tag{A.28}$$

where $Q_{X^k}^{\mathrm{U}}$ is the uniform distribution over all 3^k possible X^k sequences. This is because for any given length-k input sequence (e.g., $(0, \ldots, 0)$), when we sum $Q_{\pi_1(X_1),\ldots,\pi_k(X_k)}(0, \ldots, 0)$ over all $\pi_1, \ldots, \pi_k$, each value of $Q_{X_1\ldots X_k}(x_1, \ldots, x_k)$ is included exactly once, so the total sum is one.

If the GMI were concave, then we could use (A.28) and Jensen's inequality to deduce that the input distribution $Q_{X^k}^{\mathrm{U}}$ also attains a rate of $C_{\mathrm{M}} - \delta$. However, this is impossible, because i.i.d. random coding over the product channel with input distribution $Q_{X^k}^{\mathrm{U}}$ is exactly the same as i.i.d. random coding over the original channel (W, q) with $Q_X = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. We know from Section 2.4.3 that $I_{\mathrm{GMI}}(Q_X)$ is strictly smaller than the LM rate in this case, and we know from Lemma 2.9 that the GMI is tight with respect to the ensemble average for i.i.d. random coding. Hence, attaining a rate of $C_{\mathrm{M}} - \delta$ via the multi-letter GMI with distribution $Q_{X^k}^{\mathrm{U}}$ (for arbitrarily small δ) is not possible, meaning that the GMI cannot be concave in general.

Acknowledgments

We are grateful to Ehsan Asadi Kangarshahi for helpful discussions regarding the single-letter converse bound of Section 8.6, and to Ioannis Kontoyiannis for helpful pointers to the relevant literature on mismatched random coding in rate-distortion theory surveyed in Section 4.7.

We would also like to thank the reviewers for their helpful feedback and suggestions, including Amos Lapidoth for handling our submission and providing detailed comments. In addition, we are grateful to the editorial board and staff of the *Foundations and Trends in Communications and Information Theory* for their support.

J. Scarlett was supported by the Singapore National Research Foundation (NRF) under grant number R-252-000-A74-281. A. Guillén i Fàbregas and A. Martinez were supported by the European Research Council under ERC grant agreement 725411, and by the Spanish Ministry of Economy and Competitiveness under grant TEC2016-78434-C3-1-R. A. Somekh-Baruch was supported by the Israel Science Foundation (ISF) under grant number 631/17.

References

[1] Abou-Faycal, I. and A. Lapidoth (2000). "On the capacity of reduced-complexity receivers for intersymbol interference channels". In: *IEEE Conv. Elec. Eng. in Israel*. 263–266.

[2] Ahlswede, R., N. Cai, and Z. Zhang (1996). "Erasure, list, and detection zero-error capacities for low noise and a relation to identification". *IEEE Trans. Inf. Theory*. 42(1): 55–62.

[3] Ahlswede, R. (1973). "Multi-way communication channels". In: *IEEE Int. Symp. Inf. Theory*. 23–52.

[4] Ahlswede, R. and G. Dueck (1989). "Identification via channels". *IEEE Trans. Inf. Theory*. 35(1): 15–29.

[5] Alsan, M. (2013). "A lower bound on achievable rates by polar codes with mismatch polar decoding". In: *IEEE Inf. Theory Workshop*. 1–5.

[6] Alsan, M. (2012). "Performance of mismatched polar codes over BSCs". In: *Int. Symp. Inf. Theory Apps.* 750–755.

[7] Alsan, M. and E. Telatar (2014). "Polarization as a novel architecture to boost the classical mismatched capacity of B-DMCs". In: *IEEE Inf. Theory Workshop*. 366–370.

[8] Alvarado, A., T. Fehenberger, B. Chen, and F. Willems (2018). "Achievable information rates for fiber optics: Applications and computations". *J. Lightwave Tech.* 36(2): 424–439.

[9] Asadi Kangarshahi, E. and A. Guillén i Fàbregas (2019). "An upper bound to the mismatch capacity". In: *IEEE Int. Symp. Inf. Theory*. 2873–2877.

[10] Asadi Kangarshahi, E. and A. Guillén i Fàbregas (2019). "Properties of a recent upper bound to the mismatch capacity". In: *Int. Zürich Sem. Inf. Comm.* 115–119.

[11] Asadi Kangarshahi, E. and A. Guillén i Fàbregas (2020). "A single-letter upper bound to the mismatch capacity". URL: https://arxiv.org/abs/2004.01785.

[12] Asyhari, A. T. and A. Guillén i Fàbregas (2014). "MIMO block-fading channels with mismatched CSI". *IEEE Trans. Inf. Theory*. 60(11): 7166–7185.

[13] Asyhari, A. and A. Guillén i Fàbregas (2012). "Nearest neighbor decoding in MIMO block-fading channels with imperfect CSIR". *IEEE Trans. Inf. Theory*. 58(3): 1483–1517.

[14] Balakirsky, V. (1995). "A converse coding theorem for mismatched decoding at the output of binary-input memoryless channels". *IEEE Trans. Inf. Theory*. 41(6): 1889–1902.

[15] Binshtok, N. and S. Shamai (1999). "Integer metrics for binary input symmetric output memoryless channels". *IEEE Trans. Comms.* 47(11): 1636–1645.

[16] Blackwell, D., L. Breiman, and A. J. Thomasian (1959). "The capacity of a class of channels". *Ann. Math. Stats.* 30(4): 1229–1241.

[17] Blackwell, D., L. Breiman, and A. J. Thomasian (1960). "The capacities of certain channel classes under random coding". *Ann. Math. Stats.* 31(3): 558–567.

[18] Boyd, S. and L. Vandenberghe (2004). *Convex Optimization*. Cambridge University Press.

[19] Bross, S. I. and A. Lapidoth (2019). "The additive noise channel with a helper". In: *IEEE Inf. Theory Workshop*. 1–5.

[20] Bunte, C. and A. Lapidoth (2012). "The zero-undetected-error capacity of discrete memoryless channels with feedback". In: *Allerton Conf. Comm., Control, and Comp.* 1838–1842.

[21] Bunte, C., A. Lapidoth, and A. Samorodnitsky (2014). "The zero-undetected-error capacity approaches the Sperner capacity". *IEEE Trans. Inf. Theory.* 60(7): 3825–3833.

[22] Cover, T. M. and J. A. Thomas (2006). *Elements of Information Theory.* John Wiley & Sons, Inc.

[23] Csiszár, I. and J. Körner (1981). "Graph decomposition: A new key to coding theorems". *IEEE Trans. Inf. Theory.* 27(1): 5–12.

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

[25] Csiszar, I. and P. Narayan (1988). "The capacity of the arbitrarily varying channel revisited: Positivity, constraints". *IEEE Trans. Inf. Theory.* 34(2): 181–193.

[26] Csiszár, I. and P. Narayan (1995). "Channel capacity for a given decoding metric". *IEEE Trans. Inf. Theory.* 45(1): 35–43.

[27] de Caen, D. (1997). "A lower bound on the probability of a union". *Discrete Math.* 169: 217–220.

[28] Dembo, A. and L. Kontoyiannis (2002). "Source coding, large deviations, and approximate pattern matching". *IEEE Trans. Inf. Theory.* 48(6): 1590–1615.

[29] Divsalar, D. (1978). "Performance of mismatched receivers on bandlimited channels". *PhD Thesis.* University of California, LA.

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

[31] Elias, P. (1955). "Coding for two noisy channels". In: *London Symp. Inf. Theory.* 61–74.

[32] Fehr, D., J. Scarlett, and A. Martinez (2016). "Fixed-energy random coding with rescaled codewords at the transmitter". In: *Int. Zürich Sem. Comms.* 170–174.

[33] Feldman, Y. and A. Somekh-Baruch (2016). "Channels with state information and mismatched decoding". In: *IEEE Inf. Theory Workshop.* 419–423.

[34] Fischer, T. R. M. (1971). "Some remarks on the role of inaccuracy in Shannon's theory of information transmission". In: *Trans. 8th Prague Conf. Inf. Theory.* 211–226.

[35] Forney, G. (1973). "The Viterbi algorithm". *Proc. IEEE.* 61(3): 268–278.

[36] Gallager, R. (1965). "A simple derivation of the coding theorem and some applications". *IEEE Trans. Inf. Theory.* 11(1): 3–18.

[37] Gallager, R. (1992). "Fixed composition arguments and lower bounds to error probability". URL: http://web.mit.edu/gallager/www/notes/notes5.pdf.

[38] Gallager, R. (1968). *Information Theory and Reliable Communication.* John Wiley & Sons.

[39] Ganti, A., A. Lapidoth, and E. Telatar (2000). "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit". *IEEE Trans. Inf. Theory.* 46(7): 2315–2328.

[40] Ghozlan, H. and G. Kramer (2017). "Models and information rates for Wiener phase noise channels". *IEEE Trans. Inf. Theory.* 63(4): 2376–2393.

[41] Gilbert, E. N. (1952). "A comparison of signalling alphabets". *Bell Labs Tech. J.* 31(3): 504–522.

[42] Gray, R. M. and T. Linder (2003). "Mismatch in high-rate entropy-constrained vector quantization". *IEEE Trans. Inf. Theory.* 49(5): 1204–1217.

[43] Gray, R. and L. Davisson (1975). "Quantizer mismatch". *IEEE Trans. Comms.* 23(4): 439–443.

[44] Guillén i Fàbregas, A., A. Martinez, and G. Caire (2008). "Bit-interleaved coded modulation". *Found. Trends Comms. Inf. Theory.* 5(1–2): 1–153.

[45] Han, T. S. (2002). *Information-Spectrum Methods in Information Theory.* Springer.

[46] Hui, J. (1983). "Fundamental issues of multiple accessing". *PhD Thesis.* MIT.

[47] Huleihel, W., S. Salamatian, N. Merhav, and M. Médard (2019). "Gaussian intersymbol interference channels with mismatch". *IEEE Trans. Inf. Theory.* 65(7): 4499–4517.

[48] Kaplan, G. and S. Shamai (1993). "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment". *Arch. Elek. Über.* 47(4): 228–239.

[49] Kazakos, D. (1981). "Upper and lower bounds for noisy channel coding under mismatch". In: *Proc. Conf. Info. Sci. Sys.* 37–42.

[50] Kontoyiannis, I. and R. Zamir (2006). "Mismatched codebooks and the role of entropy coding in lossy data compression". *IEEE Trans. Inf. Theory.* 52(5): 1922–1938.

[51] Körner, J. and A. Orlitsky (1998). "Zero-error information theory". *IEEE Trans. Inf. Theory.* 44(6): 2207–2229.

[52] Kostina, V. and S. Verdú (2012). "Fixed-length lossy compression in the finite blocklength regime". *IEEE Trans. Inf. Theory.* 58(6): 3309–3338.

[53] Lapidoth, A. (1996). "Nearest neighbor decoding for additive non-Gaussian noise channels". *IEEE Trans. Inf. Theory.* 42(5): 1520–1529.

[54] Lapidoth, A. (1997). "On the role of mismatch in rate distortion theory". *IEEE Trans. Inf. Theory.* 43(1): 38–47.

[55] Lapidoth, A. (2016). "A note on feedback communication with mismatched decoding". In: *IEEE Int. Conf. Sci. Elec. Eng.* 1–5.

[56] Lapidoth, A. and P. Narayan (1998). "Reliable communication under channel uncertainty". *IEEE Trans. Inf. Theory.* 44(6): 2148–2177.

[57] Lapidoth, A. and S. Shamai (2002). "Fading channels: How perfect need "perfect side information" be?" *IEEE Trans. Inf. Theory.* 48(5): 1118–1134.

[58] Lapidoth, A. (1996). "Mismatched decoding and the multiple-access channel". *IEEE Trans. Inf. Theory.* 42(5): 1439–1452.

[59] Lapidoth, A. and S. Shamai (1998). "A lower bound on the bit-error-rate resulting from mismatched Viterbi decoding". *European Trans. Telecom.* 9(6): 473–482.

[60] Martinez, A. and A. Guillén i Fàbregas (2011). "Random-coding bounds for threshold decoders: Error exponent and saddlepoint approximation". In: *IEEE Int. Symp. Inf. Theory.* 2899–2903.

[61] Martinez, A. and A. Guillén i Fàbregas (2011). "Saddlepoint approximation of random-coding bounds". In: *Inf. Theory Apps. Workshop*. 1–6.

[62] Martinez, A., A. Guillén i Fàbregas, G. Caire, and F. Willems (2009). "Bit-interleaved coded modulation revisited: A mismatched decoding perspective". *IEEE Trans. Inf. Theory*. 55(6): 2756–2765.

[63] Merhav, N. (2013). "Universal decoding for arbitrary channels relative to a given class of decoding metrics". *IEEE Trans. Inf. Theory*. 59(9): 5566–5576.

[64] Merhav, N. (2017). "The generalized stochastic likelihood decoder: Random coding and expurgated bounds". *IEEE Trans. Inf. Theory*. 63(8): 5039–5051.

[65] Merhav, N., G. Kaplan, A. Lapidoth, and S. Shamai (1994). "On information rates for mismatched decoders". *IEEE Trans. Inf. Theory*. 40(6): 1953–1967.

[66] Motahari, A. S. and A. K. Khandani (2009). "Capacity bounds for the Gaussian interference channel". *IEEE Trans. Inf. Theory*. 55(2): 620–643.

[67] Nemhauser, G. L. and L. A. Wolsey (1999). *Integer Programming and Combinatorial Optimization*. Wiley New York.

[68] Omura, J. and B. Levitt (1982). "Coded error probability evaluation for antijam communication systems". *IEEE Trans. Comms.* 30(5): 896–903.

[69] Polyanskiy, Y., H. V. Poor, and S. Verdú (2010). "Channel coding rate in the finite blocklength regime". *IEEE Trans. Inf. Theory*. 56(5): 2307–2359.

[70] Rusek, F. and D. Fertonani (2012). "Bounds on the information rate of intersymbol interference channels based on mismatched receivers". *IEEE Trans. Inf. Theory*. 58(3): 1470–1482.

[71] Rusek, F. and A. Prlja (2012). "Optimal channel shortening for MIMO and ISI channels". *IEEE Trans. Wireless Comms.* 11(2): 810–818.

[72] Sadeghi, P., P. O. Vontobel, and R. Shams (2009). "Optimization of information rate upper and lower bounds for channels with memory". *IEEE Trans. Inf. Theory*. 55(2): 663–688.

[73] Sakrison, D. (1969). "The rate distortion function for a class of sources". *Information and Control*. 15(2): 165–195.

[74] Sakrison, D. (1970). "The rate of a class of random processes". *IEEE Trans. Inf. Theory*. 16(1): 10–16.

[75] Salz, J. and E. Zehavi (1995). "Decoding under integer metrics constraints". *IEEE Trans. Comms.* 43(234): 307–317.

[76] Scarlett, J., A. Somekh-Baruch, A. Martinez, and A. Guillén i Fàbregas (2015). "A counter-example to the mismatched decoding converse for binary-input discrete memoryless channels". *IEEE Trans. Inf. Theory*. 61(10): 5387–5395.

[77] Scarlett, J. (2014). "Reliable communication under mismatched decoding". [Online: http://itc.upf.edu/biblio/1061]. *PhD Thesis*. University of Cambridge.

[78] Scarlett, J., A. Martinez, and A. Guillén i Fàbregas (2013). "Cost-constrained random coding and applications". In: *Inf. Theory and Apps. Workshop*. San Diego, CA. 1–7.

[79] Scarlett, J., A. Martinez, and A. Guillén i Fàbregas (2014). "Mismatched decoding: Error exponents, second-order rates and saddlepoint approximations". *IEEE Trans. Inf. Theory*. 60(5): 2647–2666.

[80] Scarlett, J., A. Martinez, and A. Guillén i Fàbregas (2015). "The likelihood decoder: Error exponents and mismatch". In: *IEEE Int. Symp. Inf. Theory*. 86–90.

[81] Scarlett, J., A. Martinez, and A. Guillén i Fàbregas (2016). "Multiuser random coding techniques for mismatched decoding". *IEEE Trans. Inf. Theory*. 62(7): 3950–3970.

[82] Scarlett, J., A. Martinez, and A. Guillén i Fàbregas (2018). "Mismatched multi-letter successive decoding for the multiple-access channel". *IEEE Trans. Inf. Theory*. 64(4): 2253–2266.

[83] Scarlett, J., L. Peng, N. Merhav, A. Martinez, and A. Guillén i Fàbregas (2014). "Expurgated random-coding ensembles: Exponents, refinements and connections". *IEEE Trans. Inf. Theory*. 60(8): 4449–4462.

[84] Scarlett, J., V. Y. F. Tan, and G. Durisi (2017). "The dispersion of nearest-neighbor decoding for additive non-Gaussian channels". *IEEE Trans. Inf. Theory*. 63(1): 81–92.

[85] Secondini, M. and E. Forestieri (2017). "Scope and limitations of the nonlinear Shannon limit". *J. Lightwave Tech.* 35(4): 893–902.

[86] Shamai, S. and I. Sason (2002). "Variations on the Gallager bounds, connections, and applications". *IEEE Trans. Inf. Theory.* 48(12): 3029–3051.

[87] Shannon, C. E. (1948). "A mathematical theory of communication". *Bell Sys. Tech. Journal.* 27(July): 379–423.

[88] Shannon, C. E. (1956). "The zero error capacity of a noisy channel". *IRE Trans. Inf. Theory.* 2(3): 8–19.

[89] Shannon, C. E. (1959). "Probability of error for optimal codes in a Gaussian channel". *Bell Sys. Tech. Journal.* 38(3): 611–656.

[90] Shulman, N. (2003). "Communication over an unknown channel via common broadcasting". *PhD Thesis.* Tel Aviv University.

[91] Somekh-Baruch, A. (2015). "A general formula for the mismatch capacity". *IEEE Trans. Inf. Theory.* 61(9): 4554–4568.

[92] Somekh-Baruch, A. (2015). "Multi-letter converse bounds for the mismatched discrete memoryless channel with an additive metric". In: *IEEE Int. Symp. Inf. Theory.* 531–535.

[93] Somekh-Baruch, A. (2015). "On achievable rates and error exponents for channels with mismatched decoding". *IEEE Trans. Inf. Theory.* 61(2): 727–740.

[94] Somekh-Baruch, A. (2015). "On mismatched list decoding". In: *IEEE Int. Symp. Inf. Theory.* 526–530.

[95] Somekh-Baruch, A. (2017). "Mismatched identification via channels". In: *IEEE Int. Symp. Inf. Theory.* 2751–2755.

[96] Somekh-Baruch, A. (2018). "Converse theorems for the DMC with mismatched decoding". *IEEE Trans. Inf. Theory.* 64(9): 6196–6207.

[97] Somekh-Baruch, A., J. Scarlett, and A. Guillén i Fàbregas (2019). "Generalized random Gilbert-Varshamov codes". *IEEE Trans. Inf. Theory.* 65(6): 3452–3469.

[98] Steinberg, Y. and M. Gutman (1993). "An algorithm for source coding subject to a fidelity criterion, based on string matching". *IEEE Trans. Inf. Theory.* 39(3): 877–886.

[99] Stiglitz, I. (1966). "Coding for a class of unknown channels". *IEEE Trans. Inf. Theory.* 12(2): 189–195.

[100] Tan, V. Y. F. (2014). "Asymptotic estimates in information theory with non-vanishing error probabilities". *Found. Trends Comms. Inf. Theory.* 11(1-2): 1–184.

[101] Telatar, E. (1992). "Multi-access communications with decision feedback decoding". *PhD Thesis.* MIT.

[102] Tse, D. and P. Viswanath (2005). *Fundamentals of Wireless Communication.* Cambridge University Press.

[103] Varshamov, R. R. (1957). "Estimate of the number of signals in error correcting codes". *Dokl. Akad. Nauk SSSR.* 117(5): 739–741.

[104] Verdú, S. and T. S. Han (1994). "A general formula for channel capacity". *IEEE Trans. Inf. Theory.* 40(4): 1147–1157.

[105] Weingarten, H., Y. Steinberg, and S. Shamai (2004). "Gaussian codes and weighted nearest neighbor decoding in fading multiple-antenna channels". *IEEE Trans. Inf. Theory.* 50(8): 1665–1686.

[106] Yang, E. and J. C. Kieffer (1998). "On the performance of data compression algorithms based upon string matching". *IEEE Trans. Inf. Theory.* 44(1): 47–65.

[107] Yang, E. and Z. Zhang (1999). "On the redundancy of lossy source coding with abstract alphabets". *IEEE Trans. Inf. Theory.* 45(4): 1092–1110.

[108] Yassaee, M. H., M. R. Aref, and A. Gohari (2013). "A technique for deriving one-shot achievability results in network information theory". URL: http://arxiv.org/abs/1303.0696.

[109] Zhang, W. (2012). "A general framework for transmission with transceiver distortion and some applications". *IEEE Trans. Comms.* 60(2): 384–399.

[110] Zhang, Z. and V. K. Wei (1996). "An on-line universal lossy data compression algorithm via continuous codebook refinement. I. Basic results". *IEEE Trans. Inf. Theory.* 42(3): 803–821.

[111] Zhou, L., V. Y. F. Tan, and M. Motani (2019). "Refined asymptotics for rate-distortion using Gaussian codebooks for arbitrary sources". *IEEE Trans. Inf. Theory.* 65(5): 3145–3159.

[112] Zhou, L., V. Y. F. Tan, and M. Motani (2018). "Second-order asymptotics of universal JSCC for arbitrary sources and additive channels". In: *IEEE Int. Symp. Inf. Theory.* 1500–1504.